PHYSICAL REVIEW E **82**, 031104 (2010)

Comparing parallel- and simulated-tempering-enhanced sampling algorithms at phase-transition regimes

Carlos E. Fiore[*] and M. G. E. da Luz[†]
Departamento de Física, Universidade Federal do Paraná, CP 19044, 81531-980 Curitiba, PR, Brazil

(Received 3 May 2010; revised manuscript received 5 July 2010; published 1 September 2010)

Two important enhanced sampling algorithms, simulated (ST) and parallel (PT) tempering, are commonly used when ergodic simulations may be hard to achieve, e.g., due to a phase space separated by large free-energy barriers. This is so for systems around first-order phase transitions, a case still not fully explored with such approaches in the literature. In this contribution we make a comparative study between the PT and ST for the Ising (a lattice gas in the fluid language) and the Blume–Emery–Griffiths (a lattice gas with vacancies) models at phase-transition regimes. We show that although the two methods are equivalent in the limit of sufficiently long simulations, the PT is more advantageous than the ST with respect to all the analysis performed: convergence toward the stationarity; frequency of tunneling between phases at the coexistence; and decay of time-displaced correlation functions of thermodynamic quantities. Qualitative arguments for why one may expect better results from the PT than the ST near phase-transitions conditions are also presented.

DOI: 10.1103/PhysRevE.82.031104 PACS number(s): 05.70.Fh, 05.10.Ln, 05.50.+q

I. INTRODUCTION

A keystone procedure to obtain macroscopic thermodynamics quantities (e.g., energy, specific heat, magnetization, phase-transition points, etc.) of statistical systems is to perform appropriate averages over their microscopic configurations. In practice, however, such systems usually have a prohibitive number of states for a full covering. Therefore, approaches relying on proper representative samplings must be considered and so Monte Carlo tools become fundamental for calculations. By a proper sampling we mean that for a given instance a method should satisfactorily: (i) represent the way the system actually evolves throughout the different microstates (among the whole set $\mathcal{S}$ of microstates in the system); and (ii) generate a set Ω of visited microstates that indeed gives a good picture of all the relevant microstates which describe the problem at that particular situation.

Within this framework, an important issue is to know under what conditions the above criteria are fulfilled. For example, biased values for physical quantities may arise when the system displays local free-energy minima and the dynamics used to generate the microscopic configurations either is not able to cross such barriers or it does so, but only after too long times. Consequently, we have broken ergodicity for finite (even large) simulations [1,2], leading to metastability and thus to poor estimates for the system properties due to a nonrepresentative Ω. Metastability and broken ergodicity appear in several problems such as spin glasses, protein folding, biomolecules, and random search, to name just a few [3]. Moreover, they are not restricted only to complex systems, also being present in simpler contexts like in lattice-gas models displaying first-order phase transitions [4–6]. As noted, in such case the sampling dynamics may present difficulties to cross the energetic barriers. Then, the system can develop hysteresis by passing back and forth the phase frontiers as we change the parameter control [4].

Different alternative ideas have been considered to overcome [7] or even circumvent [5,6] entropic barriers, thus restoring the ergodic behavior. In particular, enhanced sampling algorithms, such as parallel tempering (PT) [8–10]—also known as multiple replica exchange—and simulated tempering (ST) [11–13], have recently attracted a lot of attention, specially due to their simplicity and generality compared to other Monte Carlo algorithms [4,5]. Briefly, in the PT method, microscopic configurations in higher temperatures are used to assure an ergodic free walk in lower temperatures: one simulates replicas of the same system at distinct T's, allowing the exchange of temperature between the replicas. For the ST, on the other hand, a unique replica is considered, however, the system occasionally undergoes temperature changes along its evolution.

Given the different tempering implementation in the two approaches, a natural question is how they compare to each other [14–16]. For example, the rate of temperatures switching is higher for the ST [14–16]. So, usually one could expect a larger number of distinct phase space regions visited when using the ST, thus a possible advantage over the PT. But as we discuss in Sec. II C, near phase-transition conditions this is not always the case. Therefore, it still an open query if indeed one method is systematically superior in all situations.

With the above in mind, here we compare the PT and ST efficiencies when applied to phase transitions, specially to the first-order case. In this respect, we should observe the following. In principle, for a true first-order transition, i.e., for systems in the thermodynamic limit, the energy discontinuous gap would lead to a small probability of accepting exchanges between the PT replicas [8]. But in concrete calculations, one is always dealing with finite sizes L, where the actual thermodynamics properties are described by continuous functions. Also, these functions are smooth and tend to the correct asymptotic behavior (for $L \to \infty$) only if the state space is properly sampled [6,7], what has been shown to be the case for the PT [4]. Thus, in practice the above mentioned difficulty for the PT is not an issue and the method is

---

[*] fiore@fisica.ufpr.br
[†] luz@fisica.ufpr.br

indeed an appropriate tool to study first-order transitions, as discussed and exemplified in different works [4,17,18]. Hence, the PT and ST (this latter rarely considered in such regime, few exceptions being Ref. [19]) can be analyzed at the same footing. So, possible convergence differences can be associated just to the way the algorithms generate the sets Ω, and not to the approaches eventual intrinsic distinctions (recall that conceptually they are similar [20]).

In this contribution we first revisit the simplest Ising spin model displaying a well understood second order phase transition. This is an instructive example because in a recent work [15], it has been shown that through an improved version of the ST, the frequency of successful exchanges (measured in terms of transition decay rates) is higher for the ST than for the PT method. However, the comparison was not carried near the critical temperature. By analyzing time correlation functions, defined as (with $\langle\ \rangle$ denoting time averages [21])

$$C_w(\tau) = \langle [w(t) - \bar{w}][w(t+\tau) - \bar{w}] \rangle / \sigma_w^2, \tag{1}$$

for w relevant thermodynamic quantities (like energy and magnetization) of mean $\bar{w}$ and variance σ_w, one no longer gets a better performance of the ST around T_c. In fact, we find that the PT leads to faster decaying C's.

Then, we move to the main focus of this contribution: the harder situation of strong first-order phase transitions, where the use of one-flip algorithms such as Metropolis often gives rise to poor numerical simulations. As the specific case study, we consider the lattice gas model with vacancies (a spin-1 model in the magnetic systems jargon) [22]. This class of problems has been extensively studied under different alternative methods [4–6,23,24]. Hence, the many available results can help to benchmark those obtained from the PT and ST. We show that although both, PT and ST, lead to equivalent good results in the limit of long simulations, the PT displays a faster convergence toward stationarity. Moreover, for the PT, the tunneling between different phases at the coexistence is more frequent and the generated microscopic configurations uncorrelate faster.

The work is organized as the following. In Sec. II we review the PT and ST methods, discussing distinct implementations. We also give reasons why the PT may outperform ST near phase-transition conditions. In Sec. III we consider a spin system displaying a second-order phase transition. The lattice-gas model and its comparative study with the PT and ST methods—addressing a first-order phase transition—are presented in Sec. IV. Finally, in Sec. V we draw our last remarks and the conclusion.

II. PT AND ST SAMPLING ALGORITHMS

The central idea behind a tempering enhanced sampling algorithm is try to guarantee ergodicity by means of appropriate temperature changes during the simulations, thus allowing efficient and uniform visits to a fragmented multiple regions phase space [20]. Suppose we shall study a system at a given T_0. We assume $T_1 = T_0$ and define a set of N distinct temperatures $T_1 < T_2 < \ldots < T_N$, with $\Delta T = T_N - T_1$. There are different ways to implement tempering [25], two important ones being the PT and ST, which we describe next.

A. Parallel tempering

The PT approach combines a standard algorithm (e.g., Metropolis) with the simultaneous evolution of N copies of the system (each at a different T_n), occasionally allowing the replicas to exchange their temperatures. Fixing relevant parameters, the method is implemented by first running M_{eq} times (to assure equilibration of all the N copies) a two parts procedure, (a) and (b), discussed below. After that, for each (a)-(b) composite MC step (repeated $M_{a,b}$ times) we calculate the thermodynamics quantities at the temperature of interest $T = T_1$. The average over the $M_{a,b}$ partial values give the final results. In fact, we further improve the calculations and estimate the statistical deviations by performing this procedure (after relaxation) M_{rep} times, so that in total the number of (a)-(b) MC steps is $M_{tot} = M_{eq} + M_{a,b} \times M_{rep}$.

In (a), for each replica (at a distinct T_n), a site lattice l is chosen randomly. Then, its occupation variable ω_l may change to a new value ω_l' according to the Metropolis prescription $P = \min\{1, \exp[-\beta \Delta \mathcal{H}]\}$ [26], where $\Delta \mathcal{H} = \mathcal{H}(\omega') - \mathcal{H}(\omega)$ is the energy variation due to the occupation change. This is done until a full lattice covering and the process is repeated all over again M times. (b) In the second part, arbitrary pairs of replicas (say, at $T_{n'}$ and $T_{n''}$ and with microscopy configurations ω' and ω'') can undergo temperatures switchings, with probability $[\beta_n = (k_B T_n)^{-1}]$

$$p_{n' \leftrightarrow n''} = \min\{1, \exp\{(\beta_{n'} - \beta_{n''})[\mathcal{H}(\omega') - \mathcal{H}(\omega'')]\}\}. \tag{2}$$

The PT algorithm is schematic represented in Fig. 1(a).

Although the above prescription is rather simple, few technical aspects should be observed. First, it is necessary to find a good compromise between the p's values (which increase with $\Delta T/N$ decreasing) and the replicas number N. This is so to guarantee relatively frequent exchanges, while keeping the computational efforts low. Hence, extra procedures have been proposed [18,27–30]. Here we use only the ones explained above. However we mention that for our present systems, one of us has tested some of these extra implementations [4] [always assuming arbitrary n''s and n'''s for the step (b) above], not finding any significant difference. Second, the system size (L) also imposes restrictions on the N's. For small systems, a few number of replicas is enough to assure rapid convergence. On the other hand, by increasing L the exchange probabilities [Eq. (2)] decreases, so the inclusion of extra copies becomes necessary. Such care has been explicit taken in our simulations. Finally, we observe that most works that use the PT method implement the switching attempts only between adjacent replicas (i.e., at $T_{n'}$ and $T_{n''=n'+1}$), in principle because the probability of exchanges decreases for increasing $T_{n''} - T_{n'}$. Nevertheless, it has been shown [4] that nonadjacent exchanges are essential to speed up the crossing of high free-energy barriers (what we discuss in more details in Sec. II C). Therefore, here we will allow exchanges between first ($\delta = 1$), second ($\delta = 2$), etc, neighbor replicas, meaning those between T_n and $T_{n+\delta}$.

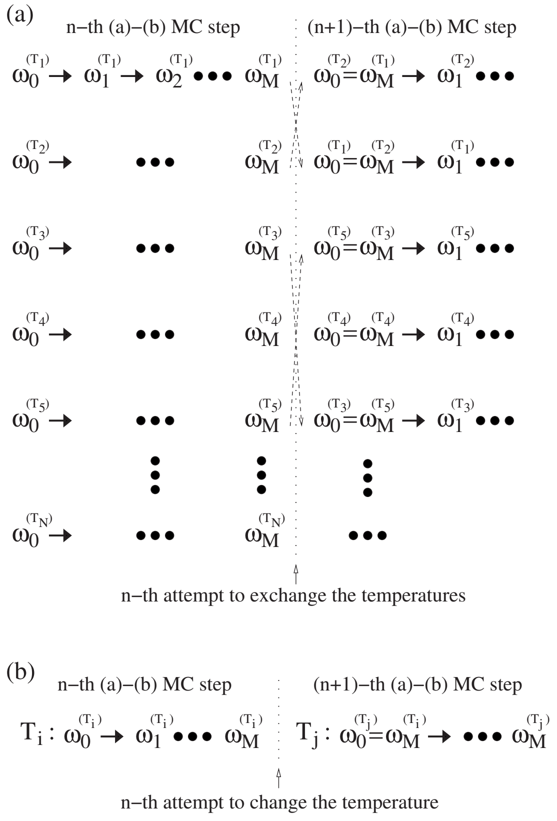

FIG. 1. Schematics of the (a) PT and (b) ST implementations. In this example, there have been two temperatures exchanges for the PT ($T_1 \leftrightarrow T_2$ and $T_3 \leftrightarrow T_5$) and one temperature change for the ST ($T_i \rightarrow T_j$).

B. Simulated tempering

For the ST, a single realization of the model is considered, however, during the dynamics its temperature can assume the different values T_n's. The implementation is similar to that for the PT in Sec. II A, but applied only to one copy of the system. Therefore, the previous step (b) now reads: A change $T_{n'} \rightarrow T_{n''}$ may take place for the system according to the probability (with ω its configuration)

$$p_{n' \rightarrow n''} = \min\{1, \exp[(\beta_{n'} - \beta_{n''})\mathcal{H}(\omega) + (g_{n''} - g_{n'})]\}. \quad (3)$$

Note that like for the PT, for the ST we will also allow nonadjacent temperatures changes. The ST algorithm is illustrated in Fig. 1(b).

In Eq. (3), $p_{n' \rightarrow n''}$ depends on the weights g's. Moreover, for a better sampling, the evolution should uniformly visit all the established temperatures. This is just the case when $g_n = \beta_n f_n$, with f_n the system free energy at T_n [13,14,16]. To obtain f is not an easy task. For instance, in Ref. [15] its exact (numerical) values follows from $f_n = -\ln[Z_n]/(V\beta_n)$, with the partition function Z_n computed by an involving recursive procedure [42]. Here, V is the system volume, which in a regular square lattice reads $V = L^2$. In our examples we will consider this same protocol, but using a simpler numerical implementation for Z_n. Indeed, in the thermodynamic limit

$$Z_n = (\lambda_n^{(0)})^L, \quad (4)$$

where $\lambda_n^{(0)}$ is the largest eigenvalue of the transfer matrix $\mathcal{T}$ at T_n (for details see, e.g., Ref. [31]). By its turn, $\lambda^{(0)} = \langle \mathcal{T}(S_k, S_k) \rangle / \langle \delta_{S_k, S_{k+1}} \rangle$ can be calculated from straightforward Monte Carlo simulations [31], where S_k is the lattice k-layer configuration $\omega_{1,k}, \omega_{2,k}, \ldots, \omega_{L,k}$ and $\delta_{S_k, S_{k+1}} = 1$ (=0) if the k and $k+1$ layers are equal (different). A central point is that in principle Eq. (4) would hold true only for infinite size systems. However, if L is not too small, the above relation is extremely accurate and for any practical purpose gives the correct Z_n, as we show in the next Section. Such way to determine $p_{n' \rightarrow n''}$ will be named the ST (exact) free-energy method, ST-FEM.

Finally, we observe that approximations for g are equally possible. One implementation being [13]

$$g_{n+1} - g_n \approx (\beta_{n+1} - \beta_n)(U_{n+1} + U_n)/2, \quad (5)$$

with $U_n = \langle \mathcal{H}_n \rangle$ ($n = 1, 2, \ldots, N$) the average energy at T_n. The U's can be evaluated from direct auxiliary simulations. For completeness we will also consider this ST approximated method, which we call ST-AM.

C. PT and ST methods near phase-transition regimes

The sampling of a statistical system when the phase space has a complicated landscape full of free-energy valleys and hills [32] is particularly delicate: one needs to uniformly visit different regions of $\mathcal{S}$ [33] (those more important for the given parameters), but which are separated by many entropic barriers [27]. In this case, the particular way in which a method evolves throughout the microstates space to generate Ω—even with the use of enhanced procedures—may crucially determine the final outcome of sampling. For instance, nonergodic "probing" of the multiple domains [34] can prevent the proper relaxation to equilibrium.

The previous comments fit perfectly well first-order phase transitions, where the minima of the free energy are separated by large barriers. Nevertheless, we observe that for second-order phase transitions, the divergence of time and spatial length correlations creates strongly correlated configurations [35]. It leads to a certain clusterization of relevant parts of $\mathcal{S}$ at the critical point, with independent and unbiased Ω difficult to obtain. So, although associated to different mechanisms, near both first and second order transitions we can expect a "fragmented" phase space. Hence, even if the PT and ST are not crucially distinct in usual situations (in fact, the ST being slight better than the PT in few instances [15]), here we argue qualitatively that in such cases the PT can outperform the ST.

Thus, for the above contexts of multiple basins [36], the Fig. 2 schematically represents "stretches" of typical dynamical paths generated by the ST and PT algorithms. The successively visited ω's until leaving the domain—delimited by high local free-energy barriers (or cluster walls)—can form, due to a complex topography, a very sinuous trajectory on that particular region of $\mathcal{S}$.

Thus, consider first the ST, Fig. 2(a). The initial microstate ω_0 evolves (for $T = T_1$) in a very tortuous path, but

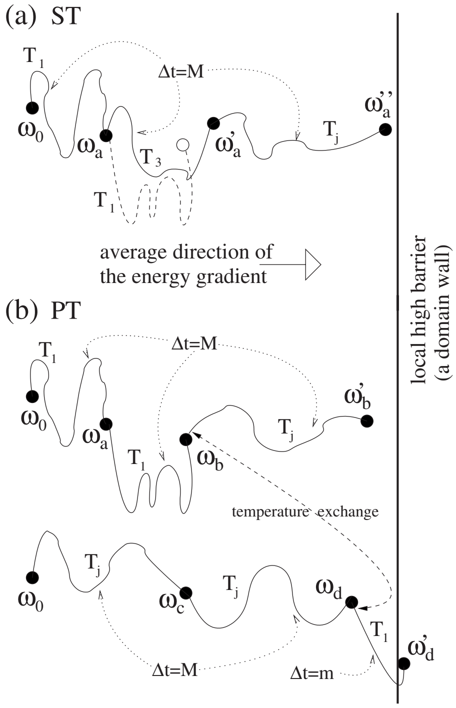

FIG. 2. Schematic illustration of the trajectories—succession of ω's—generated by the PT and ST algorithms in the case of a complex topography for the relevant microstate space. A higher sinuosity (usually associated to smaller T's) represents a higher difficulty to leave the particular region of $\mathcal{S}$, full of energetic valleys and hills. The length of the paths is proportional to the number of algorithms steps.

toward the border of the domain, reaching ω_a after M steps. Then, it undergoes a temperature change $T_1 \rightarrow T_3$ and again evolves M steps getting to ω'_a, but in a more straight trajectory because the higher T (note if there was no temperature change, the path would follow the dashed line displayed in the plot). Finally, there is a second successful attempt to change T, $T_3 \rightarrow T_j > T_3$, and after M steps the system ends up very close to the barrier separating the basins.

In Fig. 2(b) we observe the PT evolution, where just one successful temperature exchange takes place (between the only two replicas depicted). The microstate ω_b (ω_d) is obtained from ω_0 after $2M$ steps at $T=T_1$ ($T=T_j$). Obviously, ω'_a in the ST must be in average closer to (farther from) the domain border than ω_b (ω_d) in the PT implementation. Then, there is an exchange of temperatures and the evolution of ω_d at T_1, after $m<M$ steps, already allows the replica to cross the basin barrier to the microstate ω'_d. Furthermore, after $\Delta t=M$ the state ω_b at T_j leads to a ω'_b close to the border.

The above illustrated dynamics—although certainly not extinguishing all the possibilities—is already representative of why the PT can be more efficient in sampling a space full of energetic valleys and hills (e.g., at phase-transition regimes). The main reasons can be summarized as the following: (i) in the PT, the existence of replicas at all the T_n's in the interval ΔT generates paths which more quickly can approach the domain borders, e.g., $\omega_0 \rightarrow \omega_d$ at T_j in Fig. 2(b). Moreover, the microstates along such trajectories at higher T's are of course usually more energetic. (ii) So, when finally there is a temperature exchange, a microstate of high energy, even if now at lower T's, will demand a smaller number of steps to cross a barrier [like $\omega_d \rightarrow \omega'_d$ in Fig. 2(b)], and thus to start visiting other basins. On the other hand, trajectories which during a certain Δt are generated under small values of T's, thus constituted of low energy microstates, e.g., $\omega_0 \rightarrow \omega_b$ in Fig. 2(b), when shifted to higher temperatures will speed up their ways toward the barrier ($\omega_b \rightarrow \omega'_b$). Note, nevertheless, that this is possible *only* if nonadjacent exchanges are allowed, the case we are assuming here. (iii) The above collective dynamics makes possible many of the replicas successfully leave a domain after fairly similar number of steps. Hence, once in another basin region, this "parallel" process will continue in the same fashion. (iv) By its turn, we can face the ST as a "serial" process, a faster drift toward the domain walls takes place only when T increases. As a consequence, the eventual more frequent temperature exchange for the ST [14–16] not necessarily becomes an advantage in complex $\mathcal{S}$ landscapes (as illustrated in Fig. 2). (v) Also associated to (iv), we further observe that the PT (ST) is based on a multipoint (single point) searching, in an interesting analogy with collective random search theory [37]. Therefore, generally the PT with N replicas is superior to N independent ST's. (vi) Lastly, a not critical issue but which may also give some small advantage for the PT over the ST is that in the former, often the replicas (even at smaller T's) cross the domain high barriers more or less at the same time. Thus, once leaving a certain basin we already have a sample of microstates at $T=T_1$ to make averages for the PT. As displayed in the Fig. 2(a), for the ST it may happen that when the system reaches a microstate configuration able to cross the barrier, it is not at T_1. Hence, extra time is necessary for the system (naturally from the algorithm dynamics) to come back to T_1 and so the averages to be performed.

We finally observe that when the relevant space is more homogeneous in energy (e.g., far away from phase-transitions), one should not expect so high increase of the trajectories sinuosity as we diminish T. Then, it is not difficult to realize that the above mentioned differences between the PT and ST methods may become unimportant.

The previous discussion is based on qualitative arguments. Of course, they should be corroborated by concrete quantitative studies. Next we analyze two systems near phase-transition conditions. We will explicit show through detailed numerical simulations that indeed the PT algorithm is more efficient, specially in the case of first-order phase transitions.

III. ISING MODEL

The model is defined by the following Hamiltonian

$$\mathcal{H} = -J\sum_{\langle i,j\rangle} \sigma_i\sigma_j - H\sum_{i=1}^{V} \sigma_i, \tag{6}$$

where $\langle i,j\rangle$ denotes nearest-neighbors pairs i and j of a d-dimensional lattice of $V=L^d$ sites. At each site i, the spin variable assumes the values $\sigma_i = \pm 1$. J is the interaction energy and H is the magnetic field. The Ising model displays a

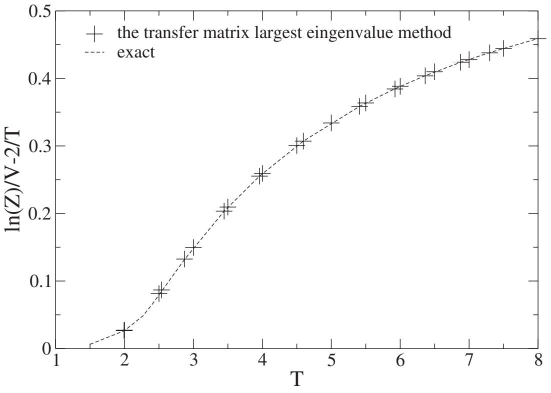

FIG. 3. For the Ising model with $H=0$, $L=32$, and units of J/k_B, comparison between the partition function versus T calculated exactly [38] and from Eq. (4).

second-order phase transition (ferromagnetic-paramagnetic) at $T_c \approx 2.269$ and $H=0$. For a square lattice ($d=2$), the transfer matrix diagonal elements are

$$\mathcal{T}(S_k,S_k) = \exp\left[\beta\left(\sum_{l=1}^{L} J(1+\sigma_{l,k}\sigma_{l+1,k}) + H\sigma_{l,k}\right)\right]. \qquad (7)$$

Our interest are in the energy $u=\langle\mathcal{H}\rangle/V$ and in the modulus of the magnetization (which is the order parameter) $m=\langle|\Sigma_{i=1}^{V}\sigma_i|\rangle/V$ per volume. For their autocorrelation functions, we just set $w=u$ and $w=m$ in Eq. (1). Regarding the parameters, we choose $H=0$ and a square lattice of $L=32$. All the results are given in units of J/k_B. To test the accuracy of the transfer matrix largest eigenvalue method in obtaining Z, in Fig. 3 we compare the exact partition function (obtained from the solution in Ref. [38]) with that calculated from Eq. (4) for the Ising model and the above parameters. The agreement is indeed remarkable, indicating that even for $L=32$, Z and consequently f is already very close to the thermodynamic limit value.

Hereafter, in all our simulations (including those in Sec. IV) we will set $M=1$ for the procedure (a) described in Sec. II A. In this case, (a) is just the Monte Carlo step per site in the parlance of Ref. [35]. Moreover, for brevity will refer to a "MC step" as being a composite implementation (a)-(b), recalling that (b) corresponds to one attempt to exchange the N replicas temperatures in the PT or to change the unique system copy temperature in the ST.

Figure 4 displays C_m and C_u for $T_1=T_c$, where we use only two temperatures, with $T_2=2.4$ (in fact, we also have considered larger numbers of replicas, not finding any significant change). From the plots we see that the autocorrelations decay faster when calculated by the PT than by both the ST-AM and ST-FEM methods. In Fig. 5 we compare the time evolution of the thermodynamic quantities (u and the modulus of the magnetization m) starting from a "hard" initial condition ω_0, i.e., a configuration very different from those representative of the steady state. This is a way of testing how efficient is a certain approach to drive the system to the stationary state. Thus, we choose a fully ordered ω_0, which obviously is not typical at $T=T_c$. The Ising model at the transition temperature evolves to the equilibrium basically in the same fashion either when simulated by the PT or by both the ST's.

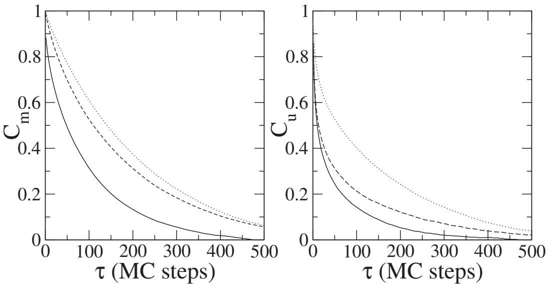

FIG. 4. For the Ising model at T_c, the autocorrelation functions versus τ (in MC steps unities), simulated from the PT (continuous), ST-FEM (dashed), and ST-AM (dotted).

So, we have that for a continuous phase transition (at least for the Ising model) the performances of the two tempering methods are essentially equivalent. Although at T_c the PT shows a somehow faster autocorrelation decays (in contrast with the results of Ref. [15] for the same model, however calculated far away from the critical temperature), the stationary state is characterized by equivalent values of m and u for all methods.

IV. LATTICE-GAS MODEL WITH VACANCIES (BEG)

A. Model

The lattice-gas model (of size $V=L^d$) with vacancies is characterized by the Hamiltonian

$$\mathcal{H} = -\sum_{\langle i,j\rangle}\sum_{r,s}\epsilon_{r,s}N_{r,i}N_{s,j} - \sum_{r}\sum_{i}\mu_r N_{r,i}. \qquad (8)$$

Here, r and s run over the species labels A and B, the ϵ_{rs}'s are the coupling energies (ϵ_{AA}, ϵ_{BB}, ϵ_{AB}, and ϵ_{BA}), $N_{r,i}=0,1$ is the occupation numbers at site i for species r, and μ_r is the species r chemical potential. The above model is equivalent to the Blume-Emery-Griffiths (BEG) spin-1 $\mathcal{H}$ [22] Indeed, defining (with $\sigma_i=0,\pm 1$ the possible values for the spin variable)

$$N_{A,i} = (\sigma_i^2+\sigma_i)/2, \quad N_{B,i} = (\sigma_i^2-\sigma_i)/2, \qquad (9)$$

associating $\sigma_i=1(-1)$ with the species $A(B)$ and $\sigma_i=0$ with a vacancy, and setting $\epsilon_{AA}=\epsilon_{BB}$ and $\epsilon_{AB}=\epsilon_{BA}$, we get the BEG Hamiltonian

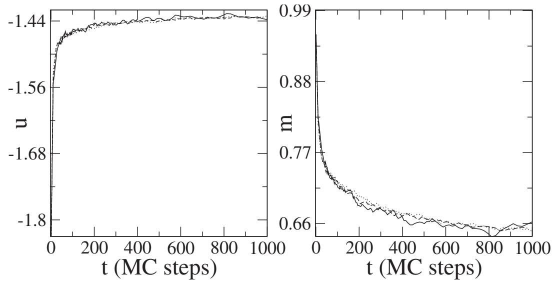

FIG. 5. For the Ising model at T_c, the time evolution of u and m from a nontypical initial configuration simulated by the PT (continuous), ST-FEM (dashed) and ST-AM (dotted).

$$\mathcal{H} = -\sum_{\langle i,j\rangle} (J\sigma_i\sigma_j + K\sigma_i^2\sigma_j^2) - \sum_i (H\sigma_i - D\sigma_i^2), \qquad (10)$$

for

$$H = (\mu_A - \mu_B)/2, \quad D = -(\mu_A + \mu_B)/2,$$

$$J = (\epsilon_{AA} - \epsilon_{AB})/2, \quad K = (\epsilon_{AA} + \epsilon_{AB})/2. \qquad (11)$$

We will consider a square lattice with periodic boundary conditions. In this case, the transfer matrix diagonal elements read

$$\mathcal{T}(S_k,S_k) = \exp\left[\beta\sum_{l=1}^{L}\{(H + J\sigma_{l+1,k})\sigma_{l,k} + [J - D + K(1 + \sigma_{l+1,k}^2)]\sigma_{l,k}^2\}\right]. \qquad (12)$$

The model has two order parameters, q and m, defined by $q = \langle\Sigma_{i=1}^{V}(N_{A,i} + N_{B,i})\rangle/V$ and $m = \langle\Sigma_{i=1}^{V}(N_{A,i} - N_{B,i})\rangle/V$. Also important is the quantity energy per volume, given by $u = \langle\mathcal{H}\rangle/V$. The autocorrelation are then obtained from $w = q$, $w = m$ and $w = u$ in Eq. (1).

B. Results

For fixed K/J, H and T, the characteristic of the phase space is determined by D. In the regime we are interested, there are two phases if D is small, one rich in species A and the other in species B. For high values of D, the model displays a single gas phase, rich in vacancies. A strong first-order phase transition between these two situations takes place at $D = D^*$, which obviously depends on K/J, H and T. For definiteness, in the following we study the BEG Hamiltonian assuming $K/J = 3$, $H = 0$ and $T = T_1 = 1.4$ (for other parameter values, see Sec. V). In such case, $D^* = 8.000$ in the thermodynamic limit [4]. All the results will be presented in units of J/k_B.

It is well known that for different lattice-gas systems, approaches based on cluster algorithms [5] are very appropriate to deal with metastability arising in first-order phase transitions. So, next we will compare results obtained from both tempering with those available from the cluster method [5]. Regarding the parameters, unless otherwise explicit mentioned, we consider $L = 20$, $D = 8.000$ and the replicas in the temperature interval $\Delta T = 0.6$. Also, whenever necessary we perform in total up to $M_{tot} = 8 \times 10^7$ simulation steps (see Sec. II A) to evaluate the sought quantities. Finally as a test, we have compared the partition function (necessary for the ST) obtained from the transfer matrix with the quite accurate calculations in Ref. [23] for the Hamiltonian Eq. (10), in the case of the Blume-Capel model ($K/J = 0$). The results were for any practical purpose identical, even for $L < 20$. So, the ST-FEM here again uses very precise numerical values for the weights g.

As the first comparative analysis, in Fig. 6 we plot the order parameter q probability distribution histogram for a long simulation run of 10^7 MC steps. As the chemical potential we set $D = 8.004$, instead of $D = 8.000$, since it leads to a same high for the two peaks of the bimodal order-parameter probability distribution (we mention, nevertheless, that $D = 8.000$ gives the same qualitative results). The agreement of the two tempering with the cluster method [6] is similar (in fact, a little better for the PT case). Such calculations show that for a long enough time, both the PT and ST are able to circumvent the metastable states, allowing the system to cross the free-energy barriers separating the different phases at the coexistence.

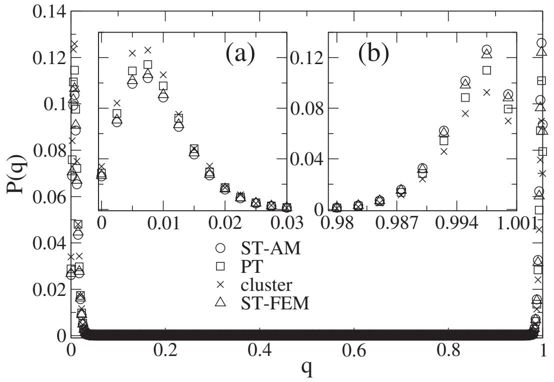

FIG. 6. For the BEG model, the histograms of the order parameter q from a long simulation using the PT, ST, and cluster algorithms. The insets are blow-ups of the (a) low and (b) high densities regions, $q \approx 0$ and $q \approx 1$, respectively.

Despite the previous agreement, the PT and ST do present differences when other aspects are analyzed. For instance, we show in Fig. 7 the time evolution of q toward the steady state, starting from a fully random initial configuration. We also consider distinct number of replicas N and temperature intervals ΔT. We find that under the same simulation conditions, generally the PT converges faster, being closer to the cluster results than the ST (ST-FEM and ST-AM). However, for the lower value of $\Delta T = 0.25$, in all cases the system (up to 10^4 MC steps) cannot even escape the region near the initial random configuration. On the other hand, by increasing $\Delta T = 0.6$—although the probability for temperature ex-

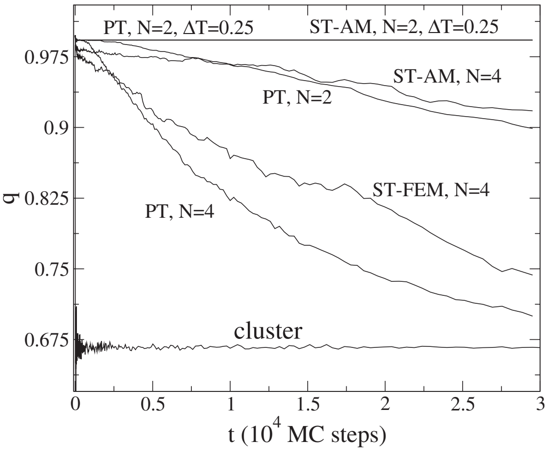

FIG. 7. For the BEG model, the time evolution of q from a fully random initial configuration, simulated from the PT, ST, and cluster. N denotes the number of replicas and $\Delta T = 0.6$ if not otherwise specified in the curves.

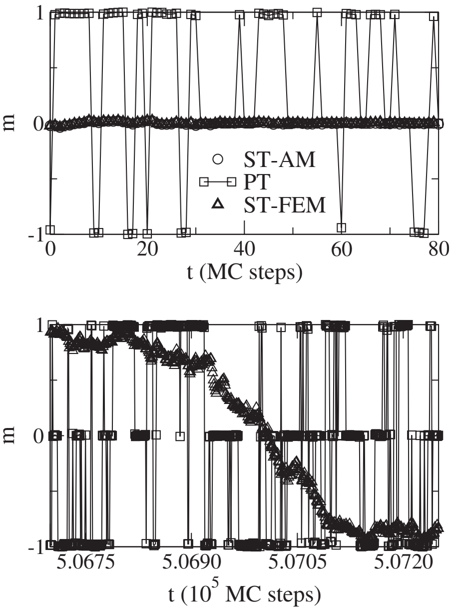

FIG. 8. For the BEG model, m versus t in two distinct time intervals at the steady state (after M_{eq}), calculated with the PT, ST-FEM and ST-AM algorithm.

changing decreases—the system starts to move toward the stationary regime. Furthermore, the larger the number of replicas N, the faster the convergence. Finally we mention that the steady value of $q=2/3$ at $D=D^{*}=8.000$ can be understood recalling that at the phase coexistence, two liquid phases ($q\approx 1$) coexist with one gas phase ($q\approx 0$). Since their weights are equal (1/3), we have $q\approx 2/3$ for any system size.

Another interesting test is to perform the numerical simulations when the system is already at the steady state. In Fig. 8 we show the time evolution of the "magnetization" m for both tempering methods at the phase coexistence. In the plots the time is shifted so to discard the M_{eq} initial MC steps necessary for equilibration. We see that the tunneling between the three different phases is substantially more frequent for the PT than for the ST. It being true along the whole evolution, as we have checked for an interval of 10^7 MC steps (in the Fig. 8 we show only two distinct simulation stretches). Actually, the PT tunneling pattern presents the same behavior than that observed in the notorious accurate cluster algorithm [5], Fig. 9.

The above results concrete exemplify some of the qualitative arguments given in Sec. II C to explain why the PT should be more efficient than the ST around first-order phase transitions. Indeed, recall that in the present situation (already in equilibrium), the coexistence takes place at a relatively low $T_1=1.4$. Therefore, the states basically belong only to one of the phases: rich in species A ($m=1$), rich in species B ($m=-1$), and rich in vacancies ($m=0$). It explains the sharp jumps between these three cases seen in the Figs. 8 and 9 for the PT (always calculated for the replica at T_1) and cluster algorithms. On the other hand, due to the way the ST changes the system temperatures along the evolution, at certain time intervals $m\times t$ must be calculated at $T_n>T_1$. This artifact leads to the observed smoother transition among the three phases. A consequence of such "slower" dynamics in visiting the different phases is that the ST needs a longer time to yield a proper Ω for the thermodynamic averages. This is exemplified in Fig. 10, where in contrast to the PT and cluster algorithms, for the number of MC steps considered the ST has not yet reached a balanced number of microstates representing the coexisting phases. For instance, note a very small numbers of states from the ST for m around -1 and a larger than expected for m around 0.

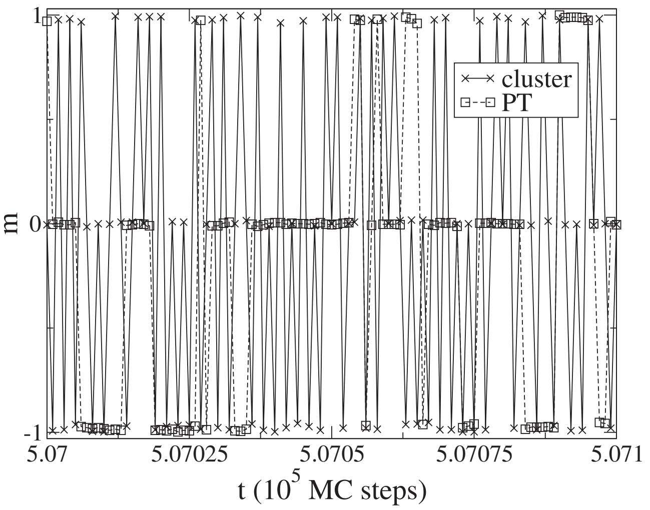

FIG. 9. Similar to Fig. 8, but comparing PT and cluster.

A different efficiency for the methods is observed not just at the phase coexistence, but also for other values of the chemical potential D around D^{*}. Figure 11 plots the order parameter q versus D for the PT and ST implementations, evaluating the averages at each $M_{a,b}=10^4$ MC steps. Note that overall the PT is already quite close to the values obtained from the cluster algorithm, whereas both ST still show some discrepancy, specially for $D>D^{*}$. If now the averages are calculate each $M_{a,b}=5\times 10^4$ MC steps, the ST also becomes closer to the cluster's (inset of Fig. 11). Once more such results can be understood in terms of the tunneling between the phases. For $D\sim D^{*}$, we still can expect high free-energy barriers. With the ST, the system does not cross such barriers a sufficient number of times if $M_{a,b}=10^4$. By increasing the number of MC steps for the averages, we generate a more representative Ω and thus a better estimation for m.

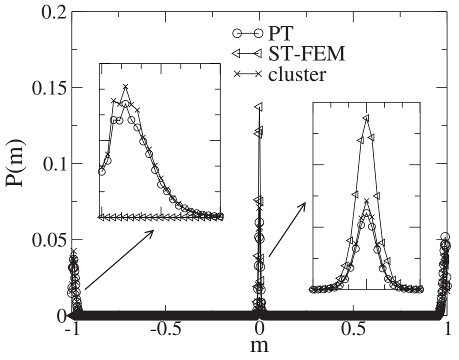

FIG. 10. For the BEG model, the histogram of m calculated with 10^5 MC steps after M_{eq}. The insets show the peaks around $m=-1$ and $m=0$ (around $m=1$, the three methods give similar peaks).

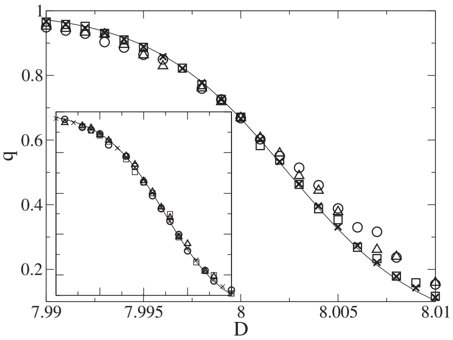

FIG. 11. For the BEG model, q versus the chemical potential, simulated from the PT (square), ST-FEM (triangle), ST-AM (circle), and cluster ($\times$). The averages are taken at each 10^4 MC steps. In the inset, exactly the same curve but for the averages at each 5×10^4 MC steps.

As a last efficiency measure, we consider the two relevant auto-correlation functions, $C_q(\tau)$ and $C_u(\tau)$, shown in Fig. 12. We should note that although time displaced correlation functions are more commonly studied in the context of continuous phase transitions, in the present case they are an interesting auxiliary tool to compare the PT and ST performances. As it should be, the ST-FEM uncorrelates faster than the ST-AM. Nevertheless, we see that the C's decay even faster for the PT method [in fact, with a very drastic difference in the case of $C_u(\tau)$].

Usually, the frequency (measured in terms of a probability p^*) in which a given tempering method changes the system temperature is taken as a good indication of its efficiency. For the PT and ST algorithms, such quantity respectively reads [39] $p^*=\langle\min\{1,\exp[(\beta_i-\beta_j)(\mathcal{H}(\sigma_i)-\mathcal{H}(\sigma_j)]\}\rangle$ and $p^*=\langle\min\{1,\exp[(\beta_i-\beta_j)\mathcal{H}(\sigma)+g_j-g_i]\}\rangle$. The averages are over $T_1,\ldots,T_N$, such that p^* of order δ is the mean from all the exchanges among T_n and $T_{n+\delta}$ (see Sec. II A).

In Fig. 13 we display p^* as function of $T=T_1$ for the PT and ST-FEM (the ST-AM being similar to the latter), with $N=12$ and $\Delta T=0.55$. As it can be seen, for any δ the ST always presents a higher probability of acceptance than the PT, in agreement with previous studies [15,16]. Such findings are in contrast with our results here. Indeed, larger p^*'s do not translate into a better performance of the ST, at least in the case of phase transitions as argued in Sec. II C. Therefore, exchange probabilities alone should be faced with care when trying to characterize the best tempering method for a certain context.

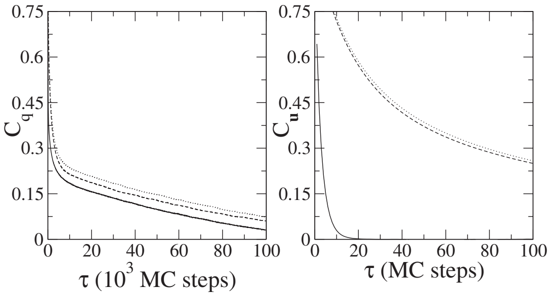

FIG. 12. Autocorrelation functions versus τ from the PT (continuous), ST-FEM (dashed), and ST-AM (dotted).

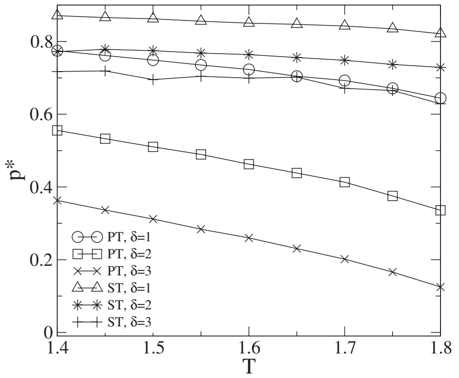

FIG. 13. Mean probability of exchange versus the temperature $T=T_1$ for the PT and ST-FEM. The symbols $\delta=1,2,3$ refers, respectively, to exchanges allowed between first, second and third neighbors (see main text).

Finally, we show in Figs. 14 and 15 finite-size analysis for the total density q and the isothermal susceptibility $\chi_T=\beta L^2(\langle q^2\rangle-\langle q\rangle^2)$ from the PT and ST-FEM. Continuous lines correspond to fitting curves by a method proposed in Ref. [6]. At the phase coexistence, thermodynamic quantities scale with the system volume [40,41]. A discontinuous phase transition is characterized by a jump in the order parameter or even a delta functionlike singularity for the susceptibility or specific heat. But this is so only at the thermodynamic limit. For finite systems not only the order parameter, but also other quantities are described by continuous functions [4,6,7], thus justifying the use of the PT method as previously mentioned. We should emphasizes that smooth curves are obtained only when one uses a simulation dynamics which yields an appropriate sampling. For instance, from simple Metropolis algorithms, neither the crossing among isotherms nor accurate finite-size analysis for smooth curves is possible. It is due to the presence of hysteresis effects [4–6], which hence demand tempering enhanced algorithm. From the plots we see that both the PT and ST give fairly good results. However, the cluster continuous curve [6] is smoother and better fitted in the PT case, specially for the larger $L=30$ value.

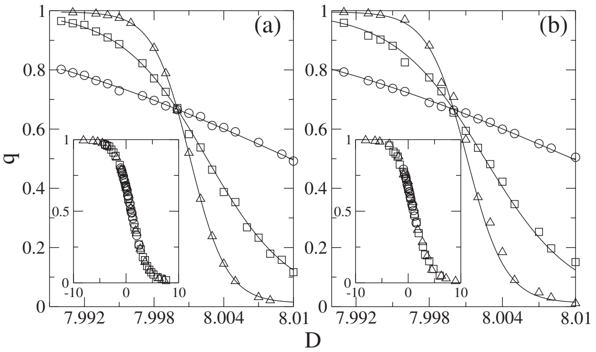

FIG. 14. q versus D for L equal to 10 (circle), 20 (square), and 30 (triangle), calculated from the (a) PT and (b) ST-FEM. Continuous lines are fitting results [4,6]. The curves collapse if plotted as $q\times(D-D^*)L^2$ (insets).

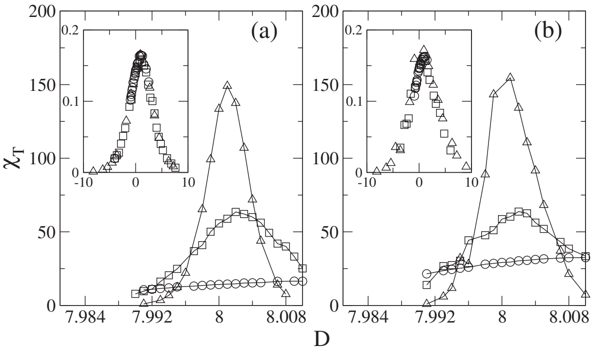

FIG. 15. Susceptibility versus D for L equal to 10 (circle), 20 (square) and 30 (triangle), calculated from the (a) PT and (b) ST-FEM. The curves collapse if plotted as $\chi_T/L^2 \times (D-D^*)L^2$ (insets).

V. REMARKS AND CONCLUSION

In this paper we have presented a comparative study between two important enhanced sampling methods, namely, simulated (ST) and parallel (PT) tempering, considering spin-lattice models at phase-transition conditions. Special attention has been paid to first-order phase transitions at low temperatures (for the BEG model). In such regimes, more standard algorithms often give poor results because their difficulties to overcome the large free-energy barriers in the phase space, leading, e.g., to ergodicity breaking and artificial algorithm-induced hysteresis. We also have investigated the less critical case of second-order phase transition—for which no free-energy barriers exist but there is the formation of strongly correlated clusters (basin regions) [35]—for the well understood Ising model.

As for the tempering implementations, we have followed the usual PT procedure, but allowing temperature exchanges between nonadjacent replicas. For the weights g in the ST, we have assumed a recent proposed approximation [13] (ST-AM) as well as the exact formulation in Ref. [15], nevertheless considering the eigenvalues of the transfer matrix [31] to calculate the partition function (ST-FEM). It avoids the necessity to implement more complicated recursive procedures to estimate Z, while still leading to very accurate values for g in practice. Finally, for the ST we also have allowed nonadjacent temperatures changes.

Different comparative studies, both at the transient regime and already at the steady state, have been carried out. Despite the facts that: (i) after long times (thus demanding large computational effort) the final results from the PT and ST are similar; and (ii) the PT displays a smaller exchange probability than the ST; we have found that for discontinuous phase transitions the PT is always more efficient in any verified aspect. The main reason for this is basically that the PT enables the system to cross free-energy barriers more frequently than the ST: either at or near phase coexistence conditions (as explicit illustrated, e.g., in Figs. 8 and 9). Furthermore, besides the quantitative numerical analysis, we also have presented heuristic arguments for why it should be expected.

Results for the instructive Ising model at the critical temperature (second-order phase transition) have also agreed with our qualitative predictions. Indeed, far away from T_c it has been reported a faster convergence for the ST [15]. We have shown that for $T \sim T_c$ just the opposite takes place, with the autocorrelations decaying faster for the PT.

For completeness, we also have analyzed other values of K/J for the BEG model (not shown), in particular for $K/J = 0$, the so called Blume-Capel model. The calculations at the first-order transition ($T_1 = 0.4$ and $D = 1.9968$) have corroborated the higher efficiency of the PT over the ST. More specifically, until $M_{tot} = 3 \times 10^7$, the system when simulated with the ST-AM has not reached the steady state, whose values for the thermodynamic quantities were different from those obtained by the ST-FEM, PT and cluster algorithms. Furthermore, the ST-FEM have agreed with the PT and cluster only for long M_{tot}'s. Time-displaced correlation functions decays and actual thermodynamic quantities convergence were always faster for the PT.

As mentioned in the Introduction, a possible difficulty in tempering approaches arises when the energy gap and/or the order-parameter jump at the transition become significant. Thus, as we increase the system size, it may start to display a rapid decreasing in the accepting probabilities for the temperature exchanges. The finite-size analyzes overcomes this problem, but in practice one should use L's only to a certain maximum value L_m, for which the considered simulation method still works well. The crucial point is then if such L_m is large enough to allow a correct extrapolation to the thermodynamic limit. We have seen here that for a certain set of parameters for the BEG model ($K/J = 3$), this is achieved already for L up to 30 (Figs. 14 and 15). Moreover, for different K/J's it has been shown [4] that at other regions of the phase space—where the first-order transition is stronger, displaying a much sharper variation of the curve order $\times$ control parameters as one increases L—again the PT leads to good results, even thought the L_m's are smaller (about 25–30 for $K/J = 3.3$ and 20–25 for the Blume-Capel model [4]). However, for those systems for which the present simpler tempering implementation would result in L_m's below the values necessary for a proper finite-size analysis, a possibility may be to consider more sophisticated procedures [27,29] to select the set $\{T_1 < T_2 < \ldots < T_N\}$, trying to reduce the effect of large latent heat at the transition point. How it can improve the efficiency of tempering methods for strong first-order transitions is presently under investigation and hopefully will be reported in the due course.

A second contribution of this work has been an (numerically simpler) alternative way to calculate the exact g in the ST method. When comparing the ST-AM with the ST-FEM, we have found that the ST-FEM allows the system to converge to steady regime quicker than the ST-AM (see above). In addition, at the steady state, configurations generated by ST-FEM uncorrelate faster than those by the ST-AM. On the other hand, with respect to the frequency in which the system tunnels between different phases at the coexistence and the final sough thermodynamic quantities values, both implementations are similar, but the ST–AM only for long M_{tot}'s.

Summarizing, at phase-transition regimes the PT and ST provide the same results for long (sometimes even costly) simulations. However, we find that for all the tested measures, the parallel converges faster than the simulated tempering. Also, even in such situation of a better performance from the PT, still the rate of temperature switching is higher for the ST. Thus, another message from our work is that alone, the switching rates are not sufficient to characterize the efficiency of a tempering enhanced sampling algorithm. It should be considered with other tests like, for instance, the ergodic (round-trip) time between the lowest and highest temperatures [8] (which even can be minimized by different ways [27,43]).

ACKNOWLEDGMENTS

We acknowledge researcher grants by CNPq. Financial support is also provided by CNPq-Edital Universal, Fundação Araucária and Finep/CT-Infra.

[1] R. G. Palmer, Adv. Phys. **31**, 669 (1982).

[2] J. P. Neirotti, D. L. Freeman, and J. D. Doll, Phys. Rev. E **62**, 7445 (2000).

[3] J. D. Bryngelson and P. G. Wolynes, Proc. Natl. Acad. Sci. U.S.A. **84**, 7524 (1987); K. H. Fisher and J. A. Hertz, *Spin Glasses* (Cambridge University Press, Cambridge, England, 1993); U. H. E. Hansmann, Chem. Phys. Lett. **281**, 140 (1997); C. L. Faustino, L. R. da Silva, M. G. E. da Luz, E. P. Raposo, and G. M. Viswanathan, EPL **77**, 30002 (2007); W. Nadler, J. H. Meinke, and U. H. E. Hansmann, Phys. Rev. E **78**, 061905 (2008).

[4] C. E. Fiore, Phys. Rev. E **78**, 041109 (2008).

[5] W. Janke and S. Kappler, Phys. Rev. Lett. **74**, 212 (1995); M. B. Bouabci and C. E. I. Carneiro, Phys. Rev. B **54**, 359 (1996); A. Rachadi and A. Benyoussef, *ibid.* **68**, 064113 (2003).

[6] C. E. Fiore and C. E. I. Carneiro, Phys. Rev. E **76**, 021118 (2007).

[7] F. Wang and D. P. Landau, Phys. Rev. Lett. **86**, 2050 (2001); Phys. Rev. E **64**, 056101 (2001).

[8] K. Hukushima and K. Nemoto, J. Phys. Soc. Jpn. **65**, 1604 (1996).

[9] C. J. Geyer, *Markov-Chain Monte Carlo Maximum Likelyhood*, Computing Science and Statistics: Proceedings of the 23rd Symposium on the Interface, edited by E. M. Keramidas (Interface Foundation, Fairfax Station, VA, 1991), p. 156-163.

[10] W. Nadler and U. H. E. Hansmann, Phys. Rev. E **76**, 057102 (2007).

[11] E. Marinari and G. Parisi, EPL **19**, 451 (1992).

[12] A. P. Lyubartsev, A. A. Martsinovski, S. V. Shevkunov, and P. N. Vorontsov-Velyaminov, J. Chem. Phys. **96**, 1776 (1992).

[13] S. Park and V. S. Pande, Phys. Rev. E **76**, 016703 (2007).

[14] S. Park, Phys. Rev. E **77**, 016709 (2008).

[15] C. Zhang and J. P. Ma, J. Chem. Phys. **129**, 134112 (2008).

[16] X. Huang, G. R. Bowmann, and V. S. Pande, J. Chem. Phys. **128**, 205106 (2008).

[17] P. Sengupta, A. W. Sandvik, and D. K. Campbell, Phys. Rev. B **65**, 155113 (2002); E. Bittner and W. Janke, J. Phys. A **41**, 395001 (2008); A. P. Young, S. Knysh, and V. N. Smelyanskiy, Phys. Rev. Lett. **104**, 020502 (2010).

[18] T. Neuhaus, M. P. Magiera, and U. H. E. Hansmann, Phys. Rev. E **76**, 045701(R) (2007).

[19] U. H. E. Hansmann and Y. Okamoto, J. Comput. Chem. **18**, 920 (1997); G. Besold, J. Risbo, and O. G. Mouritsen, Comput. Mater. Sci. **15**, 311 (1999); G. Döge, K. Mecke, J. Moller, D. Stoyan, and R. P. Waagepetersen, Int. J. Mod. Phys. C **15**, 129 (2004).

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

[21] In practice, for a discrete set $\{w_n\}$ obtained at the time instants $n=1,2,\ldots,N$, Eq. (1) reads $C_w(m)=[(N-m)^{-1}\sum_{n=1}^{n=N-m}(w_n-\bar{w})(w_{n+m}-\bar{w})]/\sigma_w^2$.

[22] M. Blume, V. J. Emery, and R. B. Griffiths, Phys. Rev. A **4**, 1071 (1971); W. Hoston and A. N. Berker, Phys. Rev. Lett. **67**, 1027 (1991).

[23] C. J. Silva, A. A. Caparica, and J. A. Plascak, Phys. Rev. E **73**, 036702 (2006).

[24] C. E. Fiore, V. B. Henriques, and M. J. de Oliveira, J. Chem. Phys. **125**, 164509 (2006).

[25] U. H. E. Hansmann and Y. Okamoto, Curr. Opin. Struct. Biol. **9**, 177 (1999); R. Denschlag, M. Lingenheil, P. Tavan, and G. Mathias, J. Chem. Theory Comput. **5**, 2847 (2009); A. Mitsutake and Y. Okamoto, Phys. Rev. E **79**, 047701 (2009).

[26] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, and A. H. Teller, J. Chem. Phys. **21**, 1087 (1953); R. J. Glauber, J. Math. Phys. **4**, 294 (1963).

[27] H. G. Katzgraber, S. Trebst, D. A. Huse, and M. Troyer, J. Stat. Mech. (2006) P03018.

[28] C. Predescu, M. Predescu, and C. Ciobanu, J. Chem. Phys. **120**, 4119 (2004); J. Phys. Chem. B **109**, 4189 (2005).

[29] D. Sabo, M. Meuwly, D. L. Freeman, and J. D. Doll, J. Chem. Phys. **128**, 174109 (2008).

[30] E. Bittner, A. Nussbaumer, and W. Janke, Phys. Rev. Lett. **101**, 130603 (2008).

[31] R. A. Sauerwein and M. J. de Oliveira, Phys. Rev. B **52**, 3060 (1995).

[32] J. C. Mauro, P. K. Gupta, and R. J. Loucks, J. Chem. Phys. **126**, 184511 (2007).

[33] P. G. Debenedetti, F. H. Stillinger, T. M. Truskett, and C. J. Roberts, J. Phys. Chem. B **103**, 7390 (1999).

[34] S. R. Williams and D. J. Evans, J. Chem. Phys. **127**, 184101 (2007).

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

[36] J. C. Mauro and R. J. Loucks, J. Non-Cryst. Solids **355**, 676 (2009).

[37] M. C. Santos, E. P. Raposo, G. M. Viswanathan, and M. G. E. da Luz, J. Phys. A **42**, 434017 (2009).

[38] B. Kaufman, Phys. Rev. **76**, 1232 (1949); A. E. Ferdinand and M. E. Fisher, *ibid.* **185**, 832 (1969).

[39] F. Calvo, J. Chem. Phys. **123**, 124106 (2005).

[40] M. S. S. Challa, D. P. Landau, and K. Binder, Phys. Rev. B **34**, 1841 (1986).

[41] C. Borgs and R. Kotecký, Phys. Rev. Lett. **68**, 1734 (1992); J. Stat. Phys. **61**, 79 (1990).

[42] C. Zhang and J. Ma, Phys. Rev. E **76**, 036708 (2007).

[43] S. Trebst, M. Troyer, and U. H. E. Hansmann, J. Chem. Phys. **124**, 174903 (2006).