Improving the efficiency of Monte Carlo simulations of systems that undergo temperature-driven phase transitions

L. Velazquez[1] and J. C. Castro-Palacio[2]
[1]*Departamento de Física, Universidad Católica del Norte, Av. Angamos 0610, Antofagasta, Chile*
[2]*Department of Chemistry, University of Basel, Klingelbergstr. 80, 4056 Basel, Switzerland*
(Received 13 April 2013; revised manuscript received 17 June 2013; published 29 July 2013)

Recently, Velazquez and Curilef proposed a methodology to extend Monte Carlo algorithms based on a canonical ensemble which aims to overcome slow sampling problems associated with temperature-driven discontinuous phase transitions. We show in this work that Monte Carlo algorithms extended with this methodology also exhibit a remarkable efficiency near a critical point. Our study is performed for the particular case of a two-dimensional four-state Potts model on a square lattice with periodic boundary conditions. This analysis reveals that the extended version of Metropolis importance sampling is more efficient than the usual Swendsen-Wang and Wolff cluster algorithms. These results demonstrate the effectiveness of this methodology to improve the efficiency of MC simulations of systems that undergo any type of temperature-driven phase transition.

DOI: 10.1103/PhysRevE.88.013311 PACS number(s): 02.70.Tt, 05.20.Gg

I. INTRODUCTION

Many different algorithms have been proposed to overcome slow sampling problems in large-scale Monte Carlo (MC) simulations. Most of them are based on two types of strategies: (1) the substitution of local MC moves by a simultaneous update of a large number of degrees of freedom, the so-called *cluster MC methods* [1–10], and (2) the use of histograms to extract information from MC simulations combined with *reweighting techniques* to improve the statistics, such as the multicanonical method and its variants [11–15]. Cluster MC methods are useful to overcome slow sampling problems associated with a continuous phase transition (PT). However, the application of nonlocal moves, by itself, does not help so much in the presence of a discontinuous PT. For such cases, the consideration of a reweighting technique as the multicanonical method is more appropriate, which reduces the size dependence of the decorrelation time from exponential to a power-law behavior.

Not one of the above strategies seem to be sufficiently general to overcome any type of sampling problem of MC simulations. The success of cluster MC algorithms is not universal because the proper cluster moves seem to be highly dependent on the system. In fact, efficient cluster MC methods have been found for only a reduced number of models [1–10]. Multicanonical method and its variants have a general applicability. However, the efficiency of these algorithms is not so significant to justify their application to overcome slow sampling problems associated with continuous PTs [14].

Recently, Velazquez and Curilef introduced a different methodology to overcome slow sampling problems associated with a temperature-driven discontinuous PT [16,17]. Their proposal is based on the general equilibrium situation associated with the fluctuation relation [18,19]

$$C = \beta^2 \langle \delta U^2 \rangle + C \langle \delta\beta_\omega \delta U \rangle, \tag{1}$$

which generalizes the canonical relation [20,21]

$$C = \beta^2 \langle \delta U^2 \rangle, \tag{2}$$

between the heat capacity C and the energy fluctuations [22]. This relation describes the existence of a *feedback perturbation* of the environment during its thermodynamic interaction with the system. This mechanism is characterized by the correlation function $\langle \delta\beta_\omega \delta U \rangle$ between the system internal energy U and the environmental inverse temperature $\beta_\omega = 1/T^\omega$. A relevant feature of fluctuation relation (1) is its compatibility with the existence of *negative heat capacities* $C < 0$ [23–25].

The consideration of the above arguments in MC simulations enables a considerable reduction of the dependence of the decorrelation times on system size N, from exponential $\tau(N) \propto \exp(\gamma N)$ to a very weak power-law behavior $\tau(N) \propto N^\alpha$ [26]. For example, the dynamic critical exponent α ranges from 0.14 to 0.2 in the case of two-dimensional (2D) 7- and 10-state Potts models regardless of if one employs local or nonlocal MC moves [16,17]. Such an improvement is significantly better than the one achieved by applying the multicanonical method and its variants to the same model systems, whose typical exponent α ranges from 2 to 2.5 [11,12]. We shall show in this work that canonical MC algorithms extended with this methodology also exhibit a good performance near a critical point. This claim is illustrated for the particular case of 2D four-state Potts model [14]. Our results evidence the effectiveness of the present methodology to overcome slow sampling problems associated with temperature-driven PT regardless of its continuous or discontinuous character.

II. OVERVIEW OF METHODOLOGY

A. Theoretical background

The methodology reviewed in this section is based on the consideration of *generalized ensembles*. Many generalized ensembles that are employed in MC simulations have no physical meaning, e.g., multicanonical ensemble [11–15]. However, this is not necessarily the case of the equilibrium situation considered in Ref. [18] to derive fluctuation relation (1): a closed system composed of two systems, A and B, with *finite heat capacities* C_A and C_B, which are put in thermal contact among them and isolated from any external

influence. This situation can be implemented in a physical laboratory with an acceptable accuracy. Moreover, this is the arrangement considered in statistical mechanics to discuss thermal equilibrium conditions [20,21]. Curiously, standard textbooks of statistical mechanics never refer to implications of equilibrium conditions concerning to states with negative heat capacities [23–25]. For its own importance in this work, let us start this section by clarifying this question. See Ref. [18] for further details.

As usual, the total energy U_T and entropy S_T of the closed system are assumed as additive quantities, $U_T = U_A + U_B$ and $S = S_A + S_B$. Maximization of entropy S_T at constant energy U_T demands the stationary condition,

$$\frac{\partial S_T}{\partial U_A} = \frac{\partial S_A}{\partial U_A} - \frac{\partial S_B}{\partial U_B} = 0 \Rightarrow \frac{1}{T_A} = \frac{1}{T_B} (\equiv \beta), \tag{3}$$

as well as the stability condition,

$$\frac{\partial^2 S_T}{\partial U_A^2} = \frac{\partial^2 S_A}{\partial U_A^2} + \frac{\partial^2 S_B}{\partial U_B^2} < 0 \Rightarrow \frac{C_A C_B}{C_A + C_B} > 0. \tag{4}$$

We have considered here the microcanonical expressions

$$\frac{1}{T_\alpha} = \frac{\partial S_\alpha}{\partial U_\alpha} \quad \text{and} \quad \frac{\partial^2 S_\alpha}{\partial U_\alpha^2} = -\frac{1}{T_\alpha^2 C_\alpha}, \tag{5}$$

with $\alpha = (A, B)$. Accordingly, systems A and B can be found in thermal equilibrium if they exhibit the same temperature and their heat capacities satisfy one of the following stability conditions: (i) both systems exhibit positive heat capacities or (ii) a system exhibits a negative heat capacity, e.g., $C_A < 0$, and the other a positive heat capacity $C_B > 0$ that satisfies the following inequality:

$$C_B < |C_A|. \tag{6}$$

Condition (6) was obtained by Thirring almost 40 years ago [27]. Accordingly, a system with a negative heat capacity cannot be found in thermal equilibrium with an environment that exhibits an infinite heat capacity, e.g., under thermodynamic influence of the natural environment. In other words, a canonical ensemble is unable to study systems with negative heat capacities. However, these systems can be analyzed considering the thermal contact with an environment that exhibits *a finite heat capacity*. This conclusion is especially relevant for MC simulations. An unexpected consequence of the above analysis concerns the so-called *zeroth law of thermodynamics* [28], which states that if two systems are both in thermal equilibrium with a third system, then they are in thermal equilibrium with each other. Although this law helps to define the notion of temperature, its validity is restricted to systems exhibiting positive heat capacities. Two identical systems that are initially prepared in the same macroscopic state cannot be in thermal equilibrium if such a macroscopic state exhibits a negative heat capacity. It is noteworthy that each system can remain in thermal equilibrium with a third system exhibiting a positive heat capacity whenever it obeys Thirring inequality (6). This violation of the zeroth law of thermodynamics was recently discussed in the literature [29].

Let us now regard system B as an *environment* in order to study thermodynamic properties of system A. Since the heat capacity of this environment is finite, its temperature T_B will be affected by the energy interchange with system A. Considering $\delta T_B = -\delta U_A / C_B \Rightarrow \delta\beta_B = \beta^2 \delta U_A / C_A$, one obtains from (1) the following result:

$$\frac{C_A C_B}{C_A + C_B} = \beta^2 \langle \delta U_A^2 \rangle. \tag{7}$$

Stability condition (4) is derived from the positivity of the right-hand side of Eq. (7) but this time from a *fluctuational viewpoint*. Relation (7) drops to canonical fluctuation relation (2) in the limit $C_B \to +\infty$, as well as the microcanonical result $\langle \delta U_A^2 \rangle \to 0$ when $C_B \to 0^+$. Accordingly, this type of equilibrium situation can be associated with a family of *generalized ensembles* that contains microcanonical and canonical ensembles as particular cases. As expected, the exact mathematical form of each ensemble depends on system B acting as the environment.

The phenomenon of negative heat capacity has been regarded as an anomalous behavior. In particular, their existence is incompatible with results of classical fluctuation theory [20,21], e.g., fluctuation relation (2). However, this incompatibility arises because of the restricted applicability of some conventional assumptions. Specifically, the macroscopic state of the environment can be affected by influence of the system under study. Such an *environmental feedback perturbation* is systematically omitted when one employs traditional ensembles, such as Boltzmann-Gibbs distributions [20,21],

$$\omega_{\mathrm{BG}}(U, X | T, Y) = \frac{1}{Z} \exp[-\beta(U + XY)], \tag{8}$$

where the environmental inverse temperature $\beta = 1/T$ and the external generalized forces Y (e.g., pressure p, magnetic and electric fields $\mathbf{H}$ and $\mathbf{E}$, etc.) are assumed as *constant control parameters* for the energy U and the generalized displacements X (e.g., volume V, magnetization $\mathbf{M}$, and electric polarization $\mathbf{P}$, etc.). As expected, more general equilibrium situations involve nonvanishing correlations such as $\langle \delta\beta \delta U \rangle$ or $\langle \delta Y \delta X \rangle$. Fluctuation theorems associated with these equilibrium situations provide a suitable treatment for states with negative heat capacities as well as other anomalies in response functions [30].

B. Application to MC simulations

The inclusion of a feedback effect $\langle \delta\beta_\omega \delta U \rangle$ to extend any canonical MC algorithm is achieved replacing the constant inverse temperature β of the canonical ensemble by an *effective inverse temperature* $\beta_\omega(U)$,

$$\beta_\omega(U) = -\frac{\partial}{\partial U} \log \omega(U), \tag{9}$$

which depends on the energy U of system under study. This effective inverse temperature corresponds to an environmental influence whose probability weight $\omega(U)$ differs from the one associated with canonical ensemble,

$$\omega(U) \neq \omega_c(U|\beta) = \frac{1}{Z(\beta)} \exp(-\beta U). \tag{10}$$

This type of arguments were employed by Gerling and Hüller to propose *dynamic ensemble* MC method [31]. These

authors considered as environment an ideal gas with N degrees of freedom. Analysis of detailed balance led them to introduce an effective inverse temperature $\beta_\omega = (N-2)/2Nk_b$. Here, $k_b = (U_T - U)/N$ is the mean kinetic energy per particle for the ideal gas, while U_T is the total energy. Effective inverse temperature β_ω is *adjusted dynamically* during the course of MC simulation. Data are then obtained by computing the mean value of the energy $\langle U\rangle$ and the mean value of the temperature from $2\langle k_b\rangle$. This method allows to detect the presence of states with negative heat capacities whenever Thirring inequality (6) is fulfilled.

Methodology proposed by Velazquez and Curilef includes three important modifications for dynamic ensemble MC method of Gerling and Hüller [16,17]: (i) the consideration of a more suitable *generalized ensemble*, (ii) the employment of a *point statistical estimation* to obtain the relevant microcanonical dependencies and reduce incidence of finite-size effects, and, finally, (iii) the *optimization of efficiency* considering a more active control on the system fluctuating behavior. In the rest of this section, let us explain these modifications in detail.

C. Gaussian ensemble and its implementation

Let us consider the power expansion of environmental inverse temperature $\beta_\omega(U)$ around the mean value of the energy $U_e = \langle U\rangle$,

$$\beta_\omega(U) = \beta + \sum_{n=1}^{\infty} a_n (U - U_e)^n. \tag{11}$$

Assuming that energy fluctuations $\delta U = U - U_e$ are sufficiently small, power expansion (11) can be restricted to first-order approximation, $\beta_\omega = \beta + \lambda\delta U/N$. Coupling constant $\lambda = Na_1$ is the control parameter in conjunction with the expectation value of inverse temperature $\beta = \langle\beta_\omega\rangle$. Substituting this ansatz into Eq. (1), one obtains the fluctuation relations

$$\langle\delta U^2\rangle = \frac{N}{\beta^2 N/C + \lambda} \quad \text{and} \quad \langle\delta\beta_\omega^2\rangle = \frac{1}{N}\frac{\lambda^2}{\beta^2 N/C + \lambda}, \tag{12}$$

as well as the stability condition,

$$\beta^2 N/C + \lambda > 0. \tag{13}$$

It is noteworthy that the size dependencies of the energy ΔU and inverse temperature $\Delta\beta_\omega$ dispersions ($\Delta x \equiv \sqrt{\langle\delta x^2\rangle}$) behave as $\Delta U \propto \sqrt{N}$ and $\Delta\beta_\omega \propto 1/\sqrt{N}$ for a short-range interacting system. As expected, the linear approximation $\beta_\omega = \beta + \lambda\delta U/N$ is good as long as the system size N be sufficiently large. If coupling constant λ obeys the stability condition (13), the statistical ensemble associated with the present equilibrium situation becomes equivalent to the microcanonical ensemble in the thermodynamic limit,

$$\lim_{N\to\infty}\frac{\Delta U}{U} = \lim_{N\to\infty}\Delta\beta_\omega = 0, \tag{14}$$

regardless of the positive or negative character of heat capacity C of the system under study. The origin of the exponential dependence with N of the decorrelation time in MC simulations $\tau(N) \propto \exp(\gamma N)$ is due to the *multimodal character* of the energy distribution function within the canonical ensemble [14]. Such a bimodal character of energy distributions is associated with the existence of macrostates with negative heat capacities. Since the ensemble equivalence ensures the existence of only one peak, MC simulations based on the present equilibrium situation cannot undergo this type of slow sampling problem.

The equilibrium situation previously described is implemented assuming a linear dependence of the environmental inverse temperature on the system energy, $\beta_\omega(U) = \beta_s + \lambda_s(U - U_s)/N$, with (U_s,β_s,λ_s) being three seed parameters, where U_s and β_s are roughly estimates of the expectation values $\langle U_\omega\rangle$ and $\langle\beta_\omega\rangle$. According to Eq. (9), this choice corresponds to the *Gaussian ensemble* [32,33],

$$\omega_G(U) = \frac{1}{Z_\lambda(\beta_s)}\exp\left[-\beta_s U - \frac{1}{2N}\lambda_s(U - U_s)^2\right], \tag{15}$$

with parameter $\lambda_s \geqslant 0$. The Gaussian ensemble describes intermediate equilibrium situations between the usual thermal contact (canonical ensemble) when $\lambda_s \to 0^+$ and energy isolation (microcanonical ensemble) when $\lambda_s \to +\infty$. The bath associated with this ensemble corresponds to a hypothetical substance whose heat capacity decreases with temperature as $C_B \propto 1/T^2$ [37] . Gaussian ensemble (15) provides several advantages to improve canonical MC simulations. In particular, its mathematical form makes the analysis of the detailed balance and the point statistical estimation easier.

Let $W^c(U_i \to U_j;\beta)$ be the transition probability of a given canonical MC algorithm, which satisfies the detailed balance condition

$$\frac{W^c(U_i \to U_j;\beta)}{W^c(U_j \to U_i;\beta)} = \exp(-\beta\delta U_{ij}), \tag{16}$$

where $\delta U_{ij} = U_j - U_i$ is energy change after transition. The detailed balance condition corresponding to the Gaussian ensemble (15) can be satisfied considering the following transition probability $W(U_i \to U_j)$:

$$W(U_i \to U_j) = W^c\left(U_i \to U_j;\beta_\omega^t\right), \tag{17}$$

where $\beta_\omega^t = (\beta_\omega^i + \beta_\omega^j)/2$ is the *transition inverse temperature* [17], with β_ω^i and β_ω^j being the environmental inverse temperatures at the initial and the final configurations, respectively, $\beta_\omega^i = \beta_\omega(U_i)$ and $\beta_\omega^j = \beta_\omega(U_j)$. This result follows from the identity

$$\frac{W(U_i \to U_j)}{W(U_j \to U_i)} = \frac{\omega_G(U_j)}{\omega_G(U_i)} \equiv \exp\left(-\beta_\omega^t\delta U_{ij}\right), \tag{18}$$

which is obtained from the mathematical form of Gaussian ensemble. Accordingly, one should replace the constant inverse temperature β of any canonical MC algorithm with the transition inverse temperature β_ω^t. Unfortunately, the application of this result requires, *a priori*, the final configuration X_j of the system with energy U_j. Therefore, this method can be applied only to extend local MC algorithms such as Metropolis importance sampling [34,35] or Glauber dynamics [36]. Extending cluster canonical MC algorithms is also possible, but their implementation is carried out by dividing each MC moves into two steps as follows [17]:

(1) To obtain a virtual configuration X_j with energy U_j through a canonical cluster MC method using the inverse

temperature β_ω^i of the initial configuration X_i with energy U_i;

(2) To accept the virtual configuration X_j using the acceptance probability $w_{i\to j}$:

$$w_{i\to j} = \min\left\{1, \frac{W_{j\to i}^j}{W_{i\to j}^i}\exp\left(-\beta_\omega^t \delta U_{ij}\right)\right\}. \tag{19}$$

The terms $W_{i\to j}^i = W^c[U_i \to U_j; \beta_\omega^i]$ and $W_{j\to i}^j = W^c[U_j \to U_i; \beta_\omega^j]$ represent the transition probabilities of the direct and the reverse process, respectively. Thus, the transition probability of the global process can be expressed as

$$W(U_i \to U_j) = W^c\left(U_i \to U_j; \beta_\omega^i\right) w_{i\to j}. \tag{20}$$

In general, values of the acceptance probability $w_{i\to j}$ are close to the unity because of the change of the inverse temperature $\delta\beta_\omega^{ij} = \beta_\omega^j - \beta_\omega^i$ and the energy change δU_{ij} are very small if the system size N is sufficiently large.

D. Point statistical estimation

By definition, statistical expectation values of macroscopic observables are *ensemble dependent*, that is, they depend on the concrete equilibrium situation associated with a given statistical ensemble. To avoid this arbitrariness, one should perform the calculation of microcanonical quantities derived from the system entropy $S(U)$, such as the microcanonical caloric curve $\beta(U) = \partial S(U)/\partial U$ and the curvature curve $\kappa(U) = -N\partial^2 S(U)/\partial U^2$. Notice that this second quantity is directly related to the microcanonical heat capacity C as $\kappa = \beta^2 N/C$.

In multicanonical algorithms and other reweighting techniques, the microcanonical dependencies $\beta(U)$ and $\kappa(U)$ can be obtained by direct numerical differentiation of the entropy $S(U)$, which was previously estimated using energy histograms. However, this procedure increases the statistical errors associated with any MC calculations, whose incidence is more significant with a larger order of differentiation [17]. A more precise calculation is performed using the *point statistical estimation* at the equilibrium energy U_e, which is related to the thermal equilibrium condition $\beta_\omega(U_e) = \beta(U_e) = \beta_e$. The estimation of microcanonical quantities (U_e, β_e, κ_e) is based on the asymptotic tendency of the energy distribution to adopt a *Gaussian form* in the thermodynamic limit $N \to +\infty$. Analogously to the dynamic ensemble MC method [31], *estimation of microcanonical dependencies is exact only in the thermodynamic limit*. However, the incidence of finite-size effects is considerably reduced using the following expressions [17]:

$$\begin{aligned} U_e &= \langle U\rangle - \frac{1-\psi_1}{2\langle\delta U^2\rangle}\langle\delta U^3\rangle + O\left(\frac{1}{N^3}\right), \\ \beta_e &= \langle\beta_\omega\rangle - \lambda\frac{1-\psi_1}{2N\langle\delta U^2\rangle}\langle\delta U^3\rangle + O\left(\frac{1}{N^3}\right), \\ \kappa_e &= \frac{1-\psi_1-\lambda\langle\delta U^2\rangle/N}{\langle\delta U^2\rangle/N} + O\left(\frac{1}{N^2}\right), \end{aligned} \tag{21}$$

where $\psi_1 = \frac{6}{5}\epsilon_2 + \frac{11}{30}\epsilon_1$ is a second-order correction term defined from the cumulants ϵ_1 and ϵ_2,

$$\epsilon_1 = \frac{\langle\delta U^3\rangle^2}{\langle\delta U^2\rangle^3}, \quad \epsilon_2 = 1 - \frac{\langle\delta U^4\rangle}{3\langle\delta U^2\rangle^2}. \tag{22}$$

Accordingly, one should proceed with the MC calculation of the statistical expectation values $\langle U\rangle$ and $\langle\beta_\omega\rangle$, as well as the n moments of energy $\langle\delta U^n\rangle$ with $n = (2,3,4)$. It is noteworthy that the expression for curvature $\kappa_e = \beta_e^2 N/C_e$ represents a second-order improvement of the energy fluctuations considered in Eq. (12). These same calculations enable us to obtain a rough estimate for the third- and the four-order derivatives of the entropy:

$$\begin{aligned} \zeta_e^3 &= N^2\frac{\partial^3 S(U_e)}{\partial U^3} = N^2\frac{\langle\delta U^3\rangle}{\langle\delta U^2\rangle^3}(1-3\psi_1) + O\left(\frac{1}{N^2}\right), \\ \zeta_e^4 &= N^3\frac{\partial^4 S(U_e)}{\partial U^4} = -\psi_2\frac{N^3}{\langle\delta U^2\rangle^3} + O\left(\frac{1}{N}\right), \end{aligned} \tag{23}$$

where $\psi_2 = \frac{12}{5}\epsilon_2 + \frac{41}{15}\epsilon_1$. Derivation of the above formulas was discussed in the appendix in Ref. [17]. The same ones were obtained for the particular case of Gaussian ensemble (15), and their applicability is associated to the licitness of the Gaussian approximation for describing system fluctuating behavior. This means that the seed parameters $(U_s, \beta_s, \lambda_s)$ of Gaussian ensemble (15) should be carefully chosen to guarantee applicability of Gaussian approximation. The way to achieve this goal will be explained at the end of the next subsection.

E. Efficiency factor and its optimization

The efficiency of MC methods is commonly characterized by the *decorrelation time* τ, that is, the minimum number of MC steps needed to generate effectively independent, identically distributed samples in the Markov chain [14]. This quantity will be calculated as follows:

$$\tau = \lim_{k\to\infty}\tau_k = \lim_{k\to\infty}\frac{k\,\mathrm{var}(u_k)}{\mathrm{var}(u_1)}, \tag{24}$$

where $\mathrm{var}(u_k) = \langle u_k^2\rangle - \langle u_k\rangle^2$ is the variance of u_k, which is defined as the arithmetic mean of the energy per particle $u = U/N$ over k samples (consecutive MC steps),

$$u_k = \frac{1}{k}\sum_{i=1}^{k} u_i. \tag{25}$$

However, the decorrelation time τ provides a partial characterization about the efficiency in the case of the extended canonical MC methods. To clarify this idea, let us consider the number of MC steps S needed to obtain a point of the caloric curve $\beta(u)$ with a precision $\delta u^2 + \delta\beta^2 < a^2$. This quantity can be estimated in terms of the total dispersion $\Delta_T^2 = \langle\delta U^2\rangle/N + N\langle\delta\beta_\omega^2\rangle$ and the decorrelation time τ as follows:

$$S = \tau\Delta_T^2/Na^2. \tag{26}$$

The total dispersion Δ_T^2 is kept fixed for canonical ensemble, and, hence, a canonical MC algorithm is more efficient as its decorrelation time τ decreases. However, the total dispersion Δ_T^2 is *ensemble dependent*, e.g., this quantity depends on the control parameters $(U_s, \beta_s, \lambda_s)$ of Gaussian ensemble (15).

According to expression (26), an extended canonical MC algorithm is more efficient as its *efficiency factor* decreases:

$$\eta = \tau \Delta_T^2. \tag{27}$$

The efficiency factor (27) depends on both decorrelation time τ and the system fluctuating behavior. Moreover, decorrelation time τ depends on both the statistical ensemble and on the concrete canonical MC algorithm. The explicit mathematical form of the decorrelation time τ in terms of control parameters of a given MC calculation is difficult to determine precisely. The simplest criterion to reduce the efficiency factor η is to minimize the total dispersion Δ_T^2, that is, to introduce a more active control onto the system fluctuating behavior [16,17]. Using the expressions of Eq. (12), the lower bound of the total dispersion Δ_T^2 and the optimal value of the control parameter λ_s are as follows:

$$\lambda_s = \lambda_\Delta(\kappa_e) = \sqrt{1+\kappa_e^2} - \kappa_e \quad \text{and} \quad \min\left(\Delta_T^2\right) = 2\lambda_\Delta, \tag{28}$$

where κ_e is the curvature at the energy point U_e. Accordingly, the optimal value for the parameter λ_s requires a rough estimation of the curvature κ_e.

Seeds parameters (U_s,β_s,λ_s) for a given MC run can be specified using the microcanonical estimates (U_e,β_e,κ_e) obtained from a previous simulation run. We have employed in this work the following iterative scheme:

$$U_s^{j+1} = U_e^j + \varepsilon; \quad \beta_s^{j+1} = \beta_e^j - \kappa_e^j \varepsilon \quad \text{and} \quad \lambda_s^{j+1} = \lambda_\Delta\left(\kappa_e^j\right), \tag{29}$$

with ε being a small energy step. It can be noticed that the scheme for β_s^{j+1} is simply a first-order power expansion of the microcanonical inverse temperature, $\beta_s^{j+1} = \beta(U_e^j + \varepsilon) = \beta(U_e^j) + \beta'(U_e^j)\varepsilon + O(\varepsilon^2)$. Moreover, we have assumed a zero-order approximation for the curvature $\kappa_e^{j+1} = \kappa(U_e^j + \varepsilon) = \kappa_e^j + O(\varepsilon)$. The initial values of the seed parameters (U_s,β_s,λ_s) could be estimated from any canonical MC algorithm far from the region of temperature-driven PT. Sometimes, it is recommendable to consider a *variable energy step* ε, overall, in those energy regions where the absolute values of microcanonical curvature curve $\kappa(U)$ are sufficiently large. We have employed in this work the following rule $\varepsilon = \varepsilon_0/\sqrt{1+\kappa_e^2}$, with ε_0 being the energy step near the critical point where $\kappa_e \simeq 0$. Notice that this rule guarantees, approximately, *a constant arc length* between neighboring

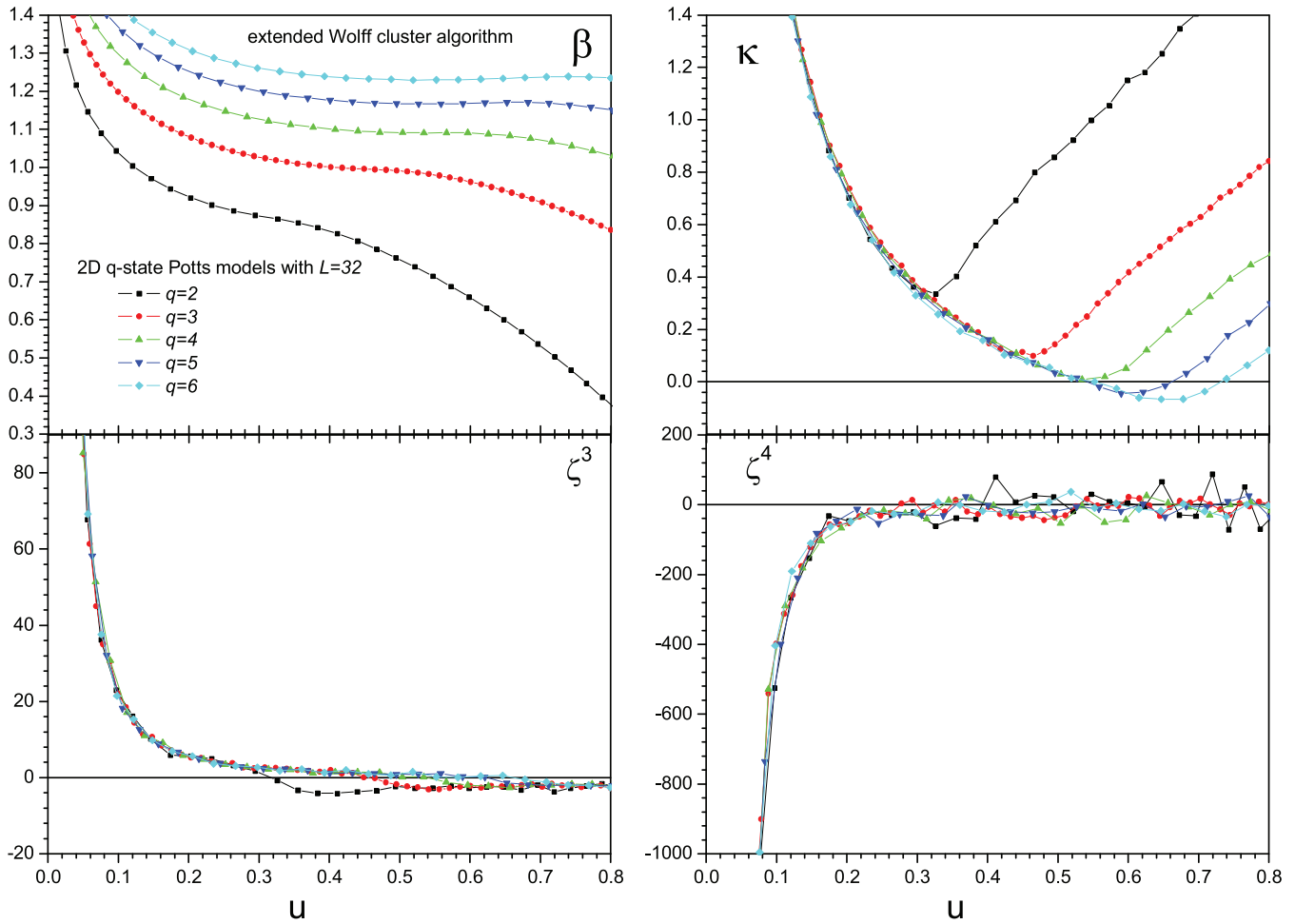

FIG. 1. (Color online) Microcanonical estimates of a 2D q-state Potts model for $q = 2$–6 with $L = 32$, which were obtained from extended version of Wolff cluster algorithm: the inverse temperature $\beta(u) = \partial s(u)/\partial u$, the curvature curve $\kappa(u) = -\partial^2 s(u)/\partial u^2$, and third and four partial derivatives, $\zeta^3(u) = \partial^3 s(u)/\partial u^3$ and $\zeta^4(u) = \partial^4 s(u)/\partial u^4$, with $s(u)$ and u being the entropy and the energy per site, respectively.

points of the microcanonical caloric curve β versus U. This feature is shown in Fig. 1.

III. EFFICIENCY NEAR A CRITICAL POINT

A. Temperature-driven continuous PT

Ensemble equivalence is always ensured in the case of a temperature-driven continuous PT. Slow sampling problems in systems that undergo this type of PT are a consequence of the large increases of the energy fluctuations and the heat capacity C when the inverse temperature β of the canonical ensemble approaches the critical point β_c. As discussed elsewhere [20], the fluctuating behavior observed here can be associated with the existence of *large correlation length* ξ among the system constituents.

The incidence of slow sampling problems could be significantly reduced if such strong correlations could be *avoided by some external influence*. If possible, the relaxation times of the averages of physical observables could be good enough even using local MC moves. Such a reduction of correlation length ξ can also be achieved considering the feedback perturbation of the environment. According to Eqs. (12) and (13), the coupling constant λ acts as a control parameter of the system thermodynamic stability and fluctuating behavior. Canonical fluctuation relation (2) predicts that the energy dispersion diverges $\Delta U \to \infty$ when $C \to \infty$. However, *the quantity* ΔU *remains finite whenever the stability condition* (13) *is applied*, that is, if the coupling constant $\lambda > 0$ when $C \longrightarrow \infty$. Since the energy fluctuations are kept finite at the critical point, the underlying correlation length ξ among the system constituents should be reduced.

B. Potts model and its MC algorithms

For the sake of convenience, let us consider the q-state Potts model [14],

$$H = -\sum_{(i,j)} \delta_{\sigma_i \sigma_j}, \tag{30}$$

defined over a square lattice $L \times L$ with periodic boundary conditions, where $\sigma_i = (1,2,\ldots,q)$ is the spin variable of the i-th site and the sum in (30) involves all nearest neighbors. This family of toy models undergoes both continuous and discontinuous PT at $\beta_c = \ln(1+\sqrt{q})$ in the thermodynamic limit $L \to \infty$. Their MC study can be performed using different canonical MC algorithms. Specifically, we will consider Metropolis importance sampling [34] as a local MC method, as well as Swendsen-Wang and Wolff cluster algorithms [1–3] as examples of nonlocal MC methods. These cluster MC methods are easily extended with the application of the present methodology. First, we need to obtain the transition probabilities $W^c_{i\to j}$ and $W^c_{j\to i}$, which appear in the acceptance probability (19). Denoting by $p_i = 1 - e^{-\beta^i_\omega}$ and $p_j = 1 - e^{-\beta^j_\omega}$ the acceptance probabilities of bonds for the direct and reverse processes, the transition probabilities $W^i_{i\to j}$ and $W^j_{j\to i}$ are expressed as follows:

$$W^i_{i\to j} = p_i^{b_a}(1-p_i)^{b_p+b_d}, \quad W^j_{j\to i} = p_j^{b_a}(1-p_j)^{b_p+b_c}. \tag{31}$$

Here, b_a and $b_p + b_d$ are the numbers of inspected bonds which have been accepted and rejected in the direct process, respectively. Moreover, b_d is the number of rejected bonds which have been destroyed in the final configuration X_f, while b_c is number of created bonds. Note that the energy change is given by $\delta U_{if} = b_d - b_c$. The integer numbers (b_a,b_d,b_c,b_p) should be obtained for each cluster move.

C. Numerical simulations

Figure 1 shows several microcanonical dependencies of the q-state Potts model with $L = 32$ and $q = 2$–6, which were estimated using the extended version of the Wolff cluster algorithm and the point statistical estimation (21). Each point of these curves was obtained considering $M = 4 \times 10^4 \tau$ iterations for each MC run, with τ being its associated decorrelation time. The convergence of the four-order derivative $\zeta^4(u)$ is less significant than the other microcanonical dependencies. However, this is a reasonable result, taking into consideration that $\zeta^4(u)$ is associated with high-order fluctuating behavior beyond the Gaussian approximation.

According to the minimal total dispersion, $\min(\Delta^2_T) = 2\lambda_\Delta(\kappa_e)$, the system exhibits its largest energy fluctuations when the curvature κ_e reaches its minimum value $\kappa_{\min}$. The character of the PT depends on the signature of the curvature $\kappa_{\min}$. It is continuous for $\kappa_{\min} \geqslant 0$ ($q = 2$–4), while discontinuous for $\kappa_{\min} < 0$ ($q > 4$). The extended version of the Wolff algorithm is able to describe both continuous and discontinuous temperature-driven PTs. Since the 2D four-state Potts model exhibits the largest fluctuating behavior near critical point, this particular case will be considered to analyze the impact of the present methodology on the efficiency of MC simulations.

For comparison purposes, the microcanonical quantities will be estimated using the entropy $S(U)$ derived from the Wang-Landau method [13]. To avoid statistical errors associated with numerical differentiation of the entropy $S(U)$, we shall consider the formulas (21) of the point statistical estimation. Statistical expectation values can be evaluated as follows:

$$\langle a(U)\rangle = \frac{\sum_i a(U_i)\exp[-\phi_G(U_i)+S(U_i)]}{\sum_i \exp[-\phi_G(U_i)+S(U_i)]}, \tag{32}$$

where $\phi_G(U) = \beta_s(U-U_s) + \lambda_s(U-U_s)^2/2N$. The estimates of the entropy per site $s = S/N$ and the inverse temperature β versus energy per site $u = U/N$ are shown in Fig. 2 for the case of a 2D four-state Potts model with $L = 32$. As clearly illustrated, results obtained from a direct numerical differentiation of entropy $S(U)$ are strongly affected by the statistical errors associated with MC calculations of energy histograms. Fortunately, the point statistical estimation overcomes this difficulty providing a smooth dependence for the microcanonical caloric curve β versus u. Results from the Wang-Landau method are considered as a reference in Fig. 3, which illustrates microcanonical estimates derived from three canonical MC algorithms and their extended versions. We have considered a variable number of steps $M = 4 \times 10^4 \tau$ for each simulation run, with τ being its decorrelation time. Dependencies associated with Wang-Landau method were

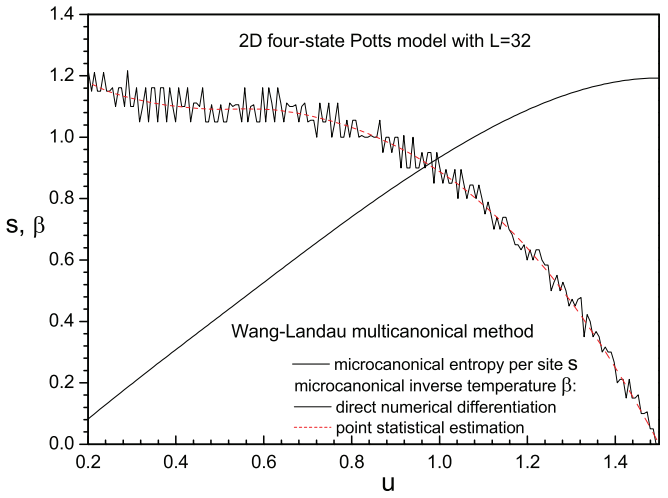

FIG. 2. (Color online) Entropy per site $s(u)$ and microcanonical inverse temperature $\beta(u)$ of a 2D four-state Potts model estimated from the Wang-Landau multicanonical algorithm. Here the variable u denotes the energy per site, $u = U/N$, with $N = L^2$.

obtained from two simulation runs with $M = 2 \times 10^7$ and $M = 1.1 \times 10^8$ steps.

Results derived from the extended versions of canonical MC algorithms and the Wang-Landau method exhibit a great agreement among them. Discrepancy among these MC methods is observed only for estimation of the curvature curve $\kappa(u)$ near the critical region [see Fig. 3(b)]. This discrepancy was also observed in Fig. 2 of Ref. [17]. In principle, extended canonical MC algorithms and the Wang-Landau method should provide same microcanonical estimates when the number of steps M is sufficiently large. However, the entropy per site $s(u)$ obtained from the Wang-Landau method is not sufficiently equilibrated to perform a more precise estimation of the curvature curve $\kappa(u) = -\partial^2 s(u)/\partial u^2$ near the critical point. The convergence of results obtained from estimation formulas (21) is not uniform everywhere. Even using the optimal values for the seed parameters $(U_s, \beta_s, \lambda_s)$ of the Gaussian ensemble (15), the largest fluctuating behavior is always observed near the critical point. This fact evidences a particular advantage of extended canonical MC algorithms. These methods enable the study of a small energy region in a given simulation run. Thus, the number of steps M of each run can be locally extended as large as needed to guarantee the convergence of microcanonical estimates. On the contrary, the Wang-Landau method sweeps the whole energy range in a single run. Although this feature is regarded as an advantage in many applications, this is not the case of calculations of partial derivatives $\partial^n s(u)/\partial u^n$. Statistical errors of entropy per site $s(u)$ are only reduced, increasing the number of steps of the Wang-Landau method for whole energy range. According to the results shown in Fig. 3(b), there exists a certain convergence of results of the Wang-Landau method towards the results of extended canonical MC methods when the number of steps M is increased from 2×10^7 to 1.1×10^8 [38]. However, the full convergence requires many more calculations. This exigence contrasts with the high performance of the extended Wolff cluster algorithm. Using this last MC method, we have employed a total of only

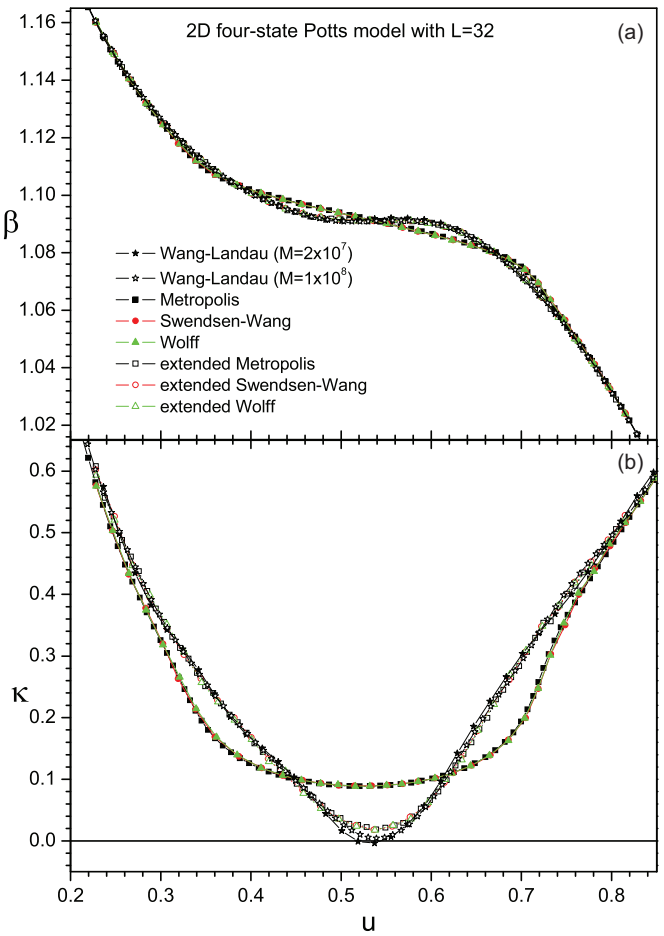

FIG. 3. (Color online) Energy dependence of inverse temperature β and curvature κ obtained from three different canonical MC algorithms and their extended versions in the case of a 2D four-state Potts model with $L = 32$. Each point estimated with these MC algorithms was obtained from a simulation run with a number of steps $M = 4 \times 10^4 \tau$, with τ being the decorrelation time of this simulation run. Results of the Wang-Landau multicanonical method are employed here as a reference, which were obtained from two different simulation runs with $M = 2 \times 10^7$ and $M = 1.1 \times 10^8$ steps.

$M = 7.3 \times 10^6$ steps, with an average of $M = 2.2 \times 10^5$ steps for each calculated point.

Microcanonical estimates derived from usual canonical MC algorithms undergo large systematic deviations. This behavior is not associated with a poor equilibration of the MC averages but with the large energy fluctuations experienced by this model system near the critical point within the canonical ensemble. The canonical ensemble is a particular case of a Gaussian ensemble with $\lambda_s = 0$, so formulas (21) of the point statistical estimation are applicable to this ensemble whenever the associated energy distribution satisfies the Gaussian approximation. This requirement cannot be satisfied near the critical point, which is is illustrated in Fig. 4. We show here the energy distributions near the critical point obtained from the MC simulations using both the Wolff cluster algorithm and its extended version for $\beta \simeq 1.098$. The distribution obtained from the usual Wolff cluster algorithm (canonical ensemble) cannot be described by a Gaussian approximation. On the

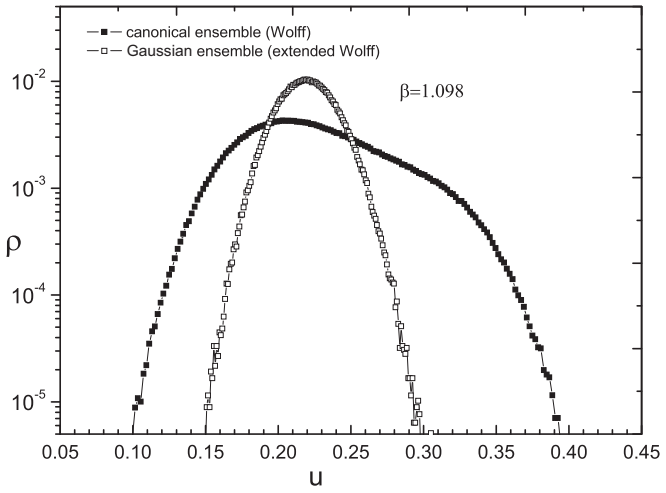

FIG. 4. Energy distributions associated with the canonical ensemble and the Gaussian ensemble with optimal values of the seed parameters (U_s,β_s,λ_s) near the critical point. These results were obtained from the Wolff cluster algorithm and its extended version, respectively.

contrary, the Gaussian approximation is fully licit for the distribution obtained from the extended Wolff cluster algorithm,

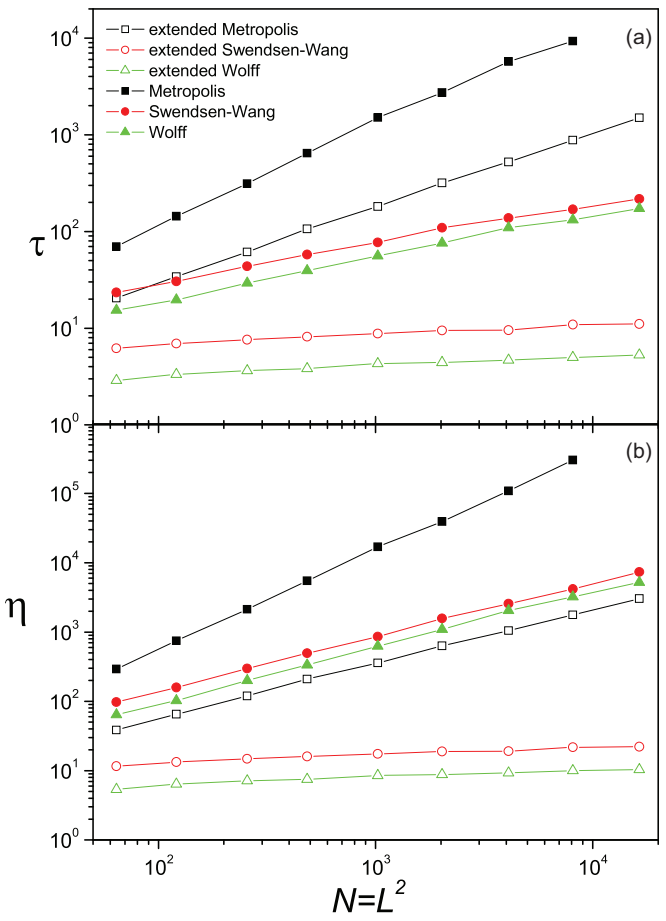

FIG. 5. (Color online) Size dependence of decorrelation time τ and efficiency factor η for three different canonical MC algorithms and their extended versions at the critical point of a 2D four-state Potts model.

TABLE I. Dynamic critical exponents α_τ and α_η associated with the size dependencies of decorrelation time τ and efficiency factor η shown in Fig. 5.

MC Method	α_τ	α_η
Metropolis	1.06 ± 0.01	1.42 ± 0.01
Extended Metropolis	0.777 ± 0.006	0.790 ± 0.008
Swendsen-Wang	0.432 ± 0.007	0.792 ± 0.008
Extended Swendsen-Wang	0.098 ± 0.004	0.117 ± 0.004
Wolff	0.474 ± 0.005	0.833 ± 0.007
Extended Wolff	0.094 ± 0.006	0.103 ± 0.006

which considers a Gaussian ensemble with optimal values of the seed parameters (U_s,β_s,λ_s). A way to reduce the incidence of finite-size effects of microcanonical estimates derived from the canonical ensemble is to consider higher-order correction terms in formulas (21). This exigency presupposes calculation of energy moments $\langle \delta U^n \rangle$ with $n > 4$, which demands larger simulation runs to achieve their convergence.

The size dependencies of the decorrelation time τ and the efficiency factor η at the critical point are shown in Fig. 5 for canonical and extended versions of three different MC algorithms for lattice sizes L ranging from 8 to 128. For all extended versions, size dependency of decorrelation time τ and the efficiency factor η exhibit power-law behaviors $\tau(N) = C_\tau N^{\alpha_\tau}$ and $\eta(N) = C_\eta N^{\alpha_\eta}$, which are weaker than their canonical counterparts. For a better quantitative characterization, estimates of dynamic critical exponents α_τ and α_η are shown in Table I. The size dependency associated with Metropolis importance sampling is reduced, but the improvement of its decorrelation time τ is less significant than the one achieved by cluster algorithms. Greater impact of the present methodology is manifested when the efficiency is described in terms of the efficiency factor η. The efficiency factor η precisely determines the number of iterations needed to achieve the convergence of the microcanonical caloric curve $\beta(u)$. All extended MC algorithms exhibit a better efficiency factor η than their original canonical counterparts. The extended version of Metropolis importance sampling, in particular, is slightly more efficient than canonical Swendsen-Wang and Wolff cluster algorithms. The exponents for extended cluster algorithms near critical point $\alpha_\eta \simeq 0.1$, which are very similar to the typical values of systems that undergo temperature-driven discontinuous PT. Dynamic critical exponents α_τ and α_η are practically the same for extended canonical MC algorithms. On the contrary, dynamic critical exponents of canonical MC algorithms exhibit a constant difference $\delta = \alpha_\eta - \alpha_\tau \simeq 0.36$ that is directly associated with the incidence of size effects in the total dispersion Δ_T^2.

IV. FINAL REMARKS

The methodology proposed by Velazquez and Curilef [16,17] leads to a significant improvement of the efficiency of MC simulations in the presence of any type of temperature-driven phase transition. Although extended canonical cluster algorithms exhibit the highest efficiencies, a local MC method, such as extended Metropolis importance sampling, has universal applicability and very good efficiency. For the particular case of a 2D four-state Potts model, this extended

local MC methods exhibits an efficiency comparable to the canonical cluster algorithms of Swendsen-Wang and Wolff. Consequently, this extended local MC algorithm can be especially useful in MC simulations of systems whose canonical cluster algorithms are still unavailable in the literature.

Before ending this section, let us refer to some open problems. First, the present methodology should be extended to those MC algorithms based on Boltzmann-Gibbs distributions [20]. An important antecedent of this problem was considered by Velazquez and Curilef in Ref. [30], where a general equilibrium fluctuation theorem (1) was generalized for the case of many thermodynamic variables. However, some relevant factors are still missing, such as extending of formulas (21) for the point statistical estimation. On the other hand, the present methodology can be combined with reweighting techniques, such as the multihistograms method, to improve statistics [14], which can provide a better estimation for the higher-order derivatives of the entropy $S(U)$.

ACKNOWLEDGMENT

L.V. thanks the financial support of CONICyT/Programa Bicentenario de Ciencia y Tecnología PSD 65 (Chilean agency).

[1] R. H. Swendsen and J.-S. Wang, Phys. Rev. Lett. **58**, 86 (1987).
[2] J.-S. Wang, R. H. Swendsen, and R. Kotecký, Phys. Rev. Lett. **63**, 109 (1989).
[3] U. Wolff, Phys. Rev. Lett. **62**, 361 (1989).
[4] R. G. Edwards and A. D. Sokal, Phys. Rev. D **38**, 2009 (1988).
[5] F. Niedermayer, Phys. Rev. Lett. **61**, 2026 (1988).
[6] H. G. Evertz, M. Hasenbusch, M. Marcu, K. Pinn, and S. Solomon, Phys. Lett. B **254**, 185 (1991).
[7] M. Hasenbusch, M. Marcu, and K. Pinn, Physica A **211**, 255 (1994).
[8] C. Dress and W. Krauth, J. Phys. A **28**, L597 (1995).
[9] J. W. Liu and E. Luijten, Phys. Rev. Lett. **92**, 035504 (2004).
[10] C. H. Mak and Arun K. Sharma, Phys. Rev. Lett. **98**, 180602 (2007).
[11] B. A. Berg and T. Neuhaus, Phys. Lett. B **267**, 249 (1991); Phys. Rev. Lett. **68**, 9 (1992).
[12] B. A. Berg, in *Proceedings of the International Conference on Multiscale Phenomena and Their Simulations*, *Bielefeld, Oct. 1996*, edited by F. Karsch, B. Monien, and H. Satz (World Scientific, Singapore, 1997).
[13] F. Wang and D. P. Landau, Phys. Rev. Lett. **86**, 2050 (2001); Phys. Rev. E **64**, 056101 (2001).
[14] P. D. Landau and K. Binder, *A Guide to Monte Carlo Simulations in Statistical Physics* (Cambridge University Press, Cambridge, UK, 2000).
[15] J. S. Wang and R. H. Swendsen, J. Stat. Phys. **106**, 245 (2001).
[16] L. Velazquez and S. Curilef, J. Stat. Mech. (2010) P02002.
[17] L. Velazquez and S. Curilef, J. Stat. Mech. (2010) P04026.
[18] L. Velazquez and S. Curilef, J. Stat. Mech. (2009) P03027.
[19] L. Velazquez and S. Curilef, J. Phys. A: Math. Theor. **42**, 095006 (2009); **42**, 335003 (2009).
[20] L. E. Reichl, *A Modern Course in Statistical Mechanics* (Wiley, New York, 1998).
[21] L. D. Landau and E. M. Lifshitz, *Statistical Physics* (Pergamon Press, London, 1980).
[22] Boltzmann constant k is assumed as the unity.
[23] T. Padmanabhan, Phys. Rep. **188**, 285 (1990).
[24] D. Lynden-Bell, Physica A **263**, 293 (1999).
[25] D. H. E. Gross, *Microcanonical Thermodynamics: Phase Transitions in Small Systems*, Vol. 66 in Lectures Notes in Physics (World Scientific, Singapore, 2001).
[26] Dependencies of decorrelation times considered in this work are expressed in terms of system size (or the number of lattice sites) $N = L^d$ instead of lattice linear size L. Consequently, dynamic critical exponents α are related to exponents z of dependence $\tau(L) \propto L^z$ as $z = D * \alpha$, with D being the lattice dimensionality.
[27] W. Thirring, Z. Phys. **235**, 339 (1970); W. Thirring, Essays in Physics **4**, 125 (1970).
[28] E. A. Guggenheim, *Thermodynamics: An Advanced Treatment for Chemists and Physicists*, 7th ed. (North Holland, Amsterdam, 1985).
[29] A. Ramirez-Hernandez, H. Larralde, and F. Leyvraz, Phys. Rev. Lett. **100**, 120601 (2008).
[30] L. Velazquez and S. Curilef, J. Stat. Mech. (2010) P12031; (2011) P06021.
[31] A. Gerling and R. W. Hüller, Z. Phys. B **90**, 207 (1993).
[32] M. S. S. Challa and J. H. Hetherington, in *Computer Simulation Studies in Condensed Matter Physics I*, edited by D. P. Landau, K. K. Mon, and H.-B. Schüttler (Springer, Heidelberg, 1988).
[33] J. H. Hetherington, J. Low Temp. Phys. **66**, 145 (1987).
[34] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).
[35] W. K. Hastings, Biometrika **57**, 97 (1970).
[36] R. J. Glauber, J. Math. Phys. **4**, 294 (1963).
[37] Dependence of the form $C \propto 1/T^2$ is observed in the hight temperature limit of a paramagnetic ideal gas.
[38] For implementing Wang-Landau multicanonical method described in Ref. [14], we have considered a minimum entry of 95% of the mean value for histogram of energies visited. First simulation run with $M = 2 \times 10^7$ steps was extended until parameter f reaches the value $f = \exp(10^{-7})$. Second simulation run with $M = 1.1 \times 10^8$ steps was extended until parameter f reaches the value $f = \exp(10^{-8})$.