Extending canonical Monte Carlo methods

L Velazquez$^{1,2}$ and S Curilef$^2$

1 Departamento de Física, Universidad de Pinar del Río, Martí 270, Esq. 27 de Noviembre, Pinar del Río, Cuba
2 Departamento de Física, Universidad Católica del Norte, Av. Angamos 0610, Antofagasta, Chile
E-mail: lvelazquez@ucn.cl and scurilef@ucn.cl

Received 26 September 2009
Accepted 12 January 2010
Published 5 February 2010

Online at stacks.iop.org/JSTAT/2010/P02002
doi:10.1088/1742-5468/2010/02/P02002

Abstract. In this paper, we discuss the implications of a recently obtained equilibrium fluctuation-dissipation relation for the extension of the available Monte Carlo methods on the basis of the consideration of the Gibbs canonical ensemble to account for the existence of an anomalous regime with negative heat capacities $C < 0$. The resulting framework appears to be a suitable generalization of the methodology associated with the so-called dynamical ensemble, which is applied to the extension of two well-known Monte Carlo methods: the Metropolis importance sampling and the Swendsen–Wang cluster algorithm. These Monte Carlo algorithms are employed to study the anomalous thermodynamic behavior of the Potts models with many spin states $q$ defined on a $d$-dimensional hypercubic lattice with periodic boundary conditions, which successfully reduce the exponential divergence of the decorrelation time $\tau$ with increase of the system size $N$ to a weak power-law divergence $\tau \propto N^\alpha$ with $\alpha \approx 0.2$ for the particular case of the 2D ten-state Potts model.

Keywords: rigorous results in statistical mechanics, classical Monte Carlo simulations
1. Introduction

In the present paper, we shall not propose new Monte Carlo (MC) methods based on the equilibrium distributions of statistical mechanics. On the contrary, we shall discuss how the available MC methods based on the consideration of the Gibbs canonical ensemble:

\[
dp(E | \beta_B) = Z (\beta_B)^{-1} \exp(-\beta_B E) \Omega(E) dE
\]  

(1)
could be extended by using a minimal but crucial modification in their schemes to account for the existence of an anomalous regime with negative heat capacities \(C < 0\) \([1]-[6]\). This avoids the occurrence of the so-called super-critical slowing down, a dynamical anomaly that significantly affects the efficiency of large-scale canonical MC simulations \([7]\).

Our proposal follows as a direct application of the recently obtained fluctuation-dissipation relation \([8,9]\)

\[
C = \beta^2 \langle \delta E^2 \rangle + C \langle \delta \beta \delta E \rangle,
\]  

(2)

which involves the heat capacity \(C\) of a given system in an equilibrium situation where the inverse temperature \(\beta\) of a certain environment exhibits correlated fluctuations with the system internal energy \(E\) as a consequence of their mutual thermodynamic interaction\(^3\). Equation (2) accounts for the realistic possibility that the internal state of the system acting as the environment could be affected by the presence of the system under study, which is a fact a priori disregarded by the consideration of the Gibbs canonical

---

\(^3\) Throughout this paper, the Boltzmann constant is assumed to be \(k_B \equiv 1\).
ensemble (1), where the inverse temperature $\beta_{\omega}$ of the environment exhibits a constant value $\beta_B$ because its heat capacity $C_{\omega}$ is practically infinite. Obviously, equation (2) is just a suitable extension of the well-known relation

$$C = \beta^2 \langle \delta E^2 \rangle$$

(3)

between the heat capacity $C$ and the energy fluctuations derived from the canonical ensemble (1). While the canonical result (3) only admits macrostates with positive heat capacities $C > 0$, it is easy to verify that the fluctuation relation (2) is compatible with the presence of macrostates having negative heat capacities $C < 0$. This last conclusion is the fundamental ingredient considered in this work for allowing a direct extension of some MC algorithms based on the Gibbs canonical ensemble (3) in order to account for a regime with $C < 0$. As discussed elsewhere [1]–[6], this kind of anomaly in the caloric curve appears to be associated with the occurrence of a discontinuous (first-order) phase transition in finite short-range interacting systems, as well as systems with long-range interactions such as astrophysical systems.

This paper is organized into sections as follows. First, we shall discuss in section 2 how the present ideas can be considered to extend the available canonical MC methods; afterwards, we shall apply these arguments in section 3 to extend two well-known canonical MC algorithms: the Metropolis importance sampling [10,11] and the Swendsen–Wang cluster algorithm [12]–[14], in their application in the study of anomalous macrostates present in the thermodynamic description of the $q$-state Potts models defined on a $d$-dimensional hypercubic lattice; finally, concluding remarks are presented in section 4.

2. The proposal

2.1. Overview

For convenience, let us begin the present discussion by reviewing the most important results related to the fluctuation theorem (2). Our analysis starts from the consideration of the following generic energy distribution function [8]:

$$d\rho_{\omega}(E) = \omega(E)\Omega(E)\,dE,$$

(4)

where $\omega(E)$ is a probabilistic weight that considers the thermodynamic influence of a certain environment. The above working hypothesis admits the canonical weight

$$\omega_c(E) = Z (\beta_B)^{-1} \exp (-\beta_B E)$$

(5)

as a relevant but particular case when the environment is just a thermal bath having an infinite heat capacity. This general situation could be implemented with the help of a Metropolis Monte Carlo simulation by using the transition probability

$$W_{\omega}(E \to E + \delta E) = \min \left\{ \frac{\omega(E + \delta E)}{\omega(E)} , 1 \right\}.$$  

(6)

Since the energy thermal fluctuations $|\delta E|$ are small when the system size is sufficiently large, equation (6) can be rewritten in a canonical fashion as follows:

$$W_{\omega}(E \to E + \delta E) \simeq \min \{ \exp [-\beta_{\omega}(E)\delta E] , 1 \},$$

(7)
where $\beta_{\omega}(E)$ is hereafter referred to as the inverse temperature of the environment:

$$\beta_{\omega}(E) = \frac{1}{T_{\omega}(E)} = -\frac{\partial}{\partial E} \log \omega(E).$$

(8)

The density of states $\Omega(E)$ is related to the system entropy $S(E)$ as $\Omega(E) \equiv C \exp[S(E)]$, which allows us to obtain the system temperature $T$ by using the thermodynamic relation

$$\beta(E) = \frac{1}{T(E)} = -\frac{\partial S(E)}{\partial E}. \quad (9)$$

Equations (8) and (9) can be combined to express the inverse temperature difference $\eta$ as follows:

$$\eta(E) = \beta_{\omega}(E) - \beta(E) \equiv -\frac{\partial}{\partial E} \log \rho(E), \quad (10)$$

with $\rho(E) = \omega(E)\Omega(E)$ being the density of probability. This last representation (10) is very useful for obtaining two remarkable thermodynamic relations. The first one involves the statistical expectation value $\langle \eta \rangle$, and its calculation reads as follows:

$$\langle \eta \rangle = \int_{E_{\text{inf}}}^{E_{\text{sup}}} \eta(E) \rho(E) \, dE = -\int_{E_{\text{inf}}}^{E_{\text{sup}}} \frac{\partial}{\partial E} \rho(E) \, dE$$

$$= -\rho(E)|_{E_{\text{sup}}}^{E_{\text{inf}}} = 0, \quad (11)$$

while the second one considers the correlation function $\langle E\eta \rangle$:

$$\langle E\eta \rangle = \int_{E_{\text{inf}}}^{E_{\text{sup}}} E\eta(E) \rho(E) \, dE = -\int_{E_{\text{inf}}}^{E_{\text{sup}}} E \frac{\partial}{\partial E} \rho(E) \, dE$$

$$= \int_{E_{\text{inf}}}^{E_{\text{sup}}} \rho(E) \, dE = 1. \quad (12)$$

Here, we have taken into account the vanishing of the density of probability $\rho(E)$ and its first derivative $\partial \rho(E)/\partial E$ at the maximum $E_{\text{sup}}$ and minimum $E_{\text{inf}}$ values of the system energy, as well as the normalization condition.

The vanishing of the expectation value $\langle \eta \rangle$, equation (11), is simply the known thermal equilibrium condition

$$\langle \frac{1}{T_{\omega}} \rangle \equiv \langle \frac{1}{T} \rangle, \quad (13)$$

whose mathematical form clearly indicates that the equalization of temperature expressed by the Zeroth Principle of Thermodynamics takes place in an average sense. Equation (12) can be rewritten as a rigorous fluctuation relation by using the identity $\langle \delta E \delta \eta \rangle \equiv \langle E\eta \rangle - \langle E \rangle \langle \eta \rangle$:

$$\langle \delta E \delta \left( \frac{1}{T_{\omega}} - \frac{1}{T} \right) \rangle = 1. \quad (14)$$

doi:10.1088/1742-5468/2010/02/P02002
By using the *Schwarz inequality* \( (AB)^2 \leq (A^2)(B^2) \), this last result can be rephrased as
\[
\Delta E \Delta \left( \frac{1}{T_\omega} - \frac{1}{T} \right) \geq 1,
\]
where \( \Delta x \equiv \sqrt{\langle \delta x^2 \rangle} \). Finally, by substituting the first-order approximation:
\[
\delta \left( \frac{1}{T} \right) \simeq -\frac{1}{T^2 C} \delta E
\]
into equation (14), with \( C = dE/dT \) being the heat capacity, one obtains the fluctuation-dissipation relation (2).

The fluctuation-dissipation relation (2) can be rewritten as follows:
\[
C \left( 1 - \langle \delta \beta_\omega \delta E \rangle \right) = \beta^2 \langle \delta E^2 \rangle.
\]
Since the right-hand side of this expression is always nonnegative, it is easy to see that the presence of macrostates with positive heat capacities \( C > 0 \) demands that the correlation function \( \langle \delta \beta_\omega \delta E \rangle \) obey the constraint
\[
\langle \delta \beta_\omega \delta E \rangle < 1.
\]
Clearly, such a condition is fulfilled by the Gibbs canonical ensemble (1), where \( \delta \beta_\omega \equiv 0 \). However, the existence of macrostates having negative heat capacities \( C < 0 \) can only be observed provided that the constraint
\[
\langle \delta \beta_\omega \delta E \rangle > 1
\]
holds. Thus, any attempt to impose the canonical condition \( \delta \beta_\omega \to 0 \) is always accompanied with a progressive increase of the energy fluctuations \( \delta E \to \infty \), which leads to the thermodynamic instability and inaccessibility of such anomalous macrostates.

The simplest way to guarantee the existence of non-vanishing correlated fluctuations \( \langle \delta \beta_\omega \delta E \rangle \neq 0 \) is achieved by considering an environment with a *finite heat capacity* \( C_\omega \). Here, the inverse temperature fluctuations \( \delta \beta_\omega \) can be expressed in terms of the amount of energy \( \delta E \) released or absorbed by the system around its equilibrium value:
\[
\delta \beta_\omega \simeq \beta^2 \frac{1}{C_\omega} \delta E,
\]
where we have considered the thermal equilibrium condition \( \beta_\omega = \beta \). By substituting this last expression into fluctuation-dissipation relation (2), one obtains
\[
\frac{CC_\omega}{C + C_\omega} = \beta^2 \langle \delta E^2 \rangle.
\]
Since the right-hand side of this last expression is always nonnegative, the thermodynamic stability of macrostates with negative heat capacity \( C < 0 \) demands the applicability of the following constraint [9]:
\[
0 < C_\omega < |C|.
\]
Remarkably, this result was also obtained in the past by Thirring [15].

doi:10.1088/1742-5468/2010/02/P02002
Extending canonical Monte Carlo methods

Figure 1. Schematic behavior of the microcanonical caloric curve $\beta(E) = \partial S(E)/\partial E$ of a finite short-range interacting system undergoing a first-order phase transition between a low (lep) and a high (hep) energy phase. The densities of probabilities $\rho_1(E)$ and $\rho_2(E)$ are associated with the thermal contact of this system with certain environments characterized by the inverse temperatures $\beta_1(\omega)(E)$ and $\beta_2(\omega)(E)$ respectively.

The above consequences are illustrated in detail in figure 1. We show here the typical back-bending behavior of the microcanonical caloric curve $\beta(E)$ of a finite short-range interacting system undergoing a first-order phase transition [3], where the points within the energetic region $E_p < E < E_q$ represent anomalous macrostates with negative heat capacities. The density of probability $\rho(E)$ corresponds to a situation where this system is put in thermal contact with a certain environment characterized by the inverse temperature $\beta_\omega(E)$. The maxima and minima of the distribution function $\rho(E)$ are determined from the thermal equilibrium condition $\beta(E) = \beta_\omega(E)$, that is, the points of intersection between these inverse temperature dependences. For convenience, we have shown here two relevant cases.

The first case corresponds to the equilibrium situation associated with the Gibbs canonical ensemble (1), where the inverse temperature dependence $\beta_B(E)$ remains at the constant value $\beta_B$ despite the underlying energy interchange. It should be noticed that the thermal equilibrium condition is fulfilled by three points $(E_a, E_b, E_c)$ when the inverse temperature $\beta_B$ is within the interval $(\beta_p, \beta_q)$. If this is the case, the energy distribution function $\rho_1(E)$ is *bimodal*, where the points $E_a$ and $E_c$ determine the positions of its peaks (local maxima), while the intermediate point $E_b$ determines the position of its local minimum. One can verify that the local minimum $E_b$ always belongs to the anomalous region with $C < 0$, while the maxima $(E_a, E_c)$ are located within the regions with $C > 0$. The accessibility of such points behaves with increase of the system size $N$ as $\rho_1(E_{a,c}) \propto e^{\alpha_{a,c}N}$ and $\rho_1(E_b) \propto e^{-\alpha_bN}$, where $\alpha_{a,b,c} > 0$. Consequently, anomalous macrostates with $C < 0$ become *practically inaccessible* within the canonical ensemble (1) when $N$ is sufficiently large. The existence of such a hidden region is the origin of the latent heat $q_L$ necessary for the conversion of one phase into the other during the phase
transition, as well as for the ensemble inequivalence between the microcanonical and canonical descriptions [3].

A multi-modal character of the density of probability $\rho_1(E)$ in the framework of MC simulations based on the canonical ensemble (1) leads to the occurrence of the super-critical slowing down [7]: an exponential divergence of correlation times with the increasing of the system size, $\tau \propto \exp(\lambda N)$. In other words, this phenomenon manifests as an effective trapping of the system energy within any of the coexisting peaks of the distribution function $\rho_1(E)$, due to the probability $T$ for the occurrence of a large energy fluctuation that allows a transition towards a neighboring peak decreasing exponentially with the increase of the system size

$$T \propto \frac{\rho_1(E_b)}{\rho_1(E_a)} \propto e^{-\lambda N}.$$ 

Thus, the characteristic time scale for such a transition is given by $\tau \propto 1/T \sim \exp(\lambda N)$. Because the canonical averages should consider the contribution of all these coexisting peaks, the relaxation time scales of such expectation values also exhibit an exponential growth with increase of $N$.

The second case shown in figure 1 corresponds to a situation where the system is put in thermal contact with an environment having a finite heat capacity. By choosing appropriately the environment and its internal conditions, in particular, the applicability of Thirring’s constraint (22), the corresponding inverse temperature $\beta_2(E)$ can ensure the existence of only one intersection point with the microcanonical caloric curve $\beta(E)$ of the system under study. Such a point could even be located within the anomalous region with $C < 0$, e.g., the unstable macrostate $E_b$. Since the energy distribution function $\rho_2(E)$ is monomodal, the phenomenon of super-critical slowing down cannot be present when such a physical situation is simulated by using a suitable MC method.

2.2. Application in Monte Carlo methods

In the multicanonical MC method and its variants [7], the main aim is to obtain the density of states $\Omega(E)$, or, equivalently, the microcanonical entropy $S(E)$. Such a goal can be achieved through a direct MC calculation of the energy distribution function $\rho(E) = \omega(E)\Omega(E)$, which can be inverted to express the system entropy and the canonical expectation values $\langle A \rangle_\beta$ as follows:

$$S(E) = \log [\rho(E)] - \log [\omega(E)] + \text{const}, \quad (23)$$

$$\langle A \rangle_\beta = \left[ \int A(E) \frac{\rho(E)}{\omega(E)} e^{-\beta_\omega E} dE \right] \left[ \int \frac{\rho(E)}{\omega(E)} e^{-\beta_\omega E} dE \right]^{-1}. \quad (24)$$

The mathematical form of the probabilistic weight, $\omega(E)$, can be conveniently proposed before performing the MC simulation [16], although the commonest strategy is to carry out its iterative reconstruction through several preliminary MC runs to obtain a flat histogram $\rho(E) \sim \text{const}$ within the energy interval of interest [17,18]. This latter alternative allows one to enhance rare events, which is particularly useful for avoiding the super-critical slowing down near first-order phase transitions [7]. This kind of MC simulation allows the acquisition of thermodynamic information over a wide energy interval by performing a single MC run. A clear disadvantage is that a substantial fraction of the computational resources could be consumed in finding an optimal probabilistic weight $\omega(E)$. Moreover, the acquisition of the microcanonical caloric curve $\beta(E) = \partial S(E)/\partial E$ as well as the heat capacity $C(E) = -\beta^2(E)[\partial^2 S(E)/\partial E^2]^{-1}$ from a numerical differentiation of the entropy

\[ \text{doi:10.1088/1742-5468/2010/02/P02002} \]
Extending canonical Monte Carlo methods

$S(E)$ is a procedure that enhances the unavoidable statistical errors involved in the MC calculation of the energy distribution $\rho(E)$ and/or the probabilistic weight $\omega(E)$.

Clearly, it is desirable to implement some type of methodology that allows a direct estimation of the caloric curve $\beta(E)$ and of the heat capacity $C(E)$ without considering a numerical differentiation of the system entropy $S(E)$. A simple strategy is provided by considering the thermal contact of the system with an environment characterized by a finite heat capacity, as in the second case illustrated in figure 1. It should be noticed here that the system energy $E$ and the inverse temperature $\beta_\omega$ of the environment undergo small thermal fluctuations around their expectation values $\langle E \rangle$ and $\langle \beta_\omega \rangle$, which provide a suitable estimation of the intersection point $(E_e, \beta_e)$ derived from the condition of thermal equilibrium $\beta_\omega(E_e) = \beta(E_e) \equiv \beta_e$:

$$E_e = \langle E \rangle, \quad \beta_e = \langle \beta_\omega \rangle. \quad (25)$$

In essence, equation (25) expresses the procedure associated with the MC method of the so-called dynamical ensemble proposed by Gerling and Hüller [19], which is defined by a probabilistic weight $\omega(E)$ with a power-law shape:

$$\omega_{DE}(E) = N_0 (E_T - E)^B, \quad (26)$$

whose corresponding inverse temperature is given by

$$\beta_\omega(E) = B/(E_T - E). \quad (27)$$

The fluctuation relation (2) enables the introduction of several improvements for this kind of MC calculation for estimating the thermo-statistical properties of a system within the microcanonical ensemble by using modified canonical MC algorithms. In fact, one can also obtain the heat capacity $C_e = C(E_e)$ at the interception point from the fluctuating behavior of the system as follows:

$$C_e = \beta_e^2 \langle \delta E^2 \rangle / (1 - \langle \delta \beta \omega \delta E \rangle). \quad (28)$$

Thus, one is able to obtain a suitable estimation of any point on the microcanonical caloric curve $\beta(E)$ as well as the heat capacity $C(E)$ regardless of whether its character is positive or negative.

2.2.1. The linear ansatz and its optimization. Generally speaking, the specific mathematical form of the probabilistic weight $\omega(E)$ or its corresponding inverse temperature $\beta_\omega(E)$ is unimportant as long as the following conditions apply:

1. The inverse temperature $\beta_\omega(E)$ of the environment must ensure the existence of only one point of intersection with the caloric curve $\beta(E)$, that is, the existence of only one peak in the energy distribution function $\rho(E)$.

2. The size $N$ of the system under analysis should be sufficiently large to guarantee the validity of the first-order approximation (16) considered for deriving the fluctuation-dissipation relation (2) from the rigorous fluctuation relation (14).
As already depicted in figure 1, the energy distribution function $\rho(E)$ has the shape of a bell curve, usually approximated by a Gaussian distribution. Since the system exhibits small thermal fluctuations $\Delta E/E \sim 1/\sqrt{N}$, only local properties of the inverse temperature dependence $\beta_\omega(E)$ close to the equilibrium point $E_e$ are significant; hence, one can obtain the same practical results by using different mathematical dependences for the inverse temperature $\beta_\omega(E)$.

The simplest mathematical form is provided by restricting to the first-order power expansion of the inverse temperature dependence $\beta_\omega(E)$ around the intersection point:

$$\beta_\omega(E) = \beta_e + \lambda \delta E/N,$$

which considers a linear coupling of the environment inverse temperature $\beta_\omega$ with the thermal fluctuations of the system energy $\delta E = E - E_e$. Remarkably, such a linear ansatz (29) is equivalent to the so-called Gaussian ensemble [20]:

$$\omega_{GE}(E) = A \exp \left[ -\beta_e E - \frac{1}{N} \lambda (E - E_e)^2 \right]$$

proposed by Hetherington [21], which approaches in the limit $\lambda \to +\infty$ the microcanonical ensemble $\omega_{ME}(E) = A \delta(E - E_e)$.

Hypothesis (29) ensures the existence of non-vanishing correlated fluctuations $\langle \delta \beta_\omega \delta E \rangle \neq 0$ when the coupling constant $\lambda \neq 0$. By using the fluctuation-dissipation relation (2) and the ansatz (29), one can obtain the following expression for the heat capacity:

$$C_e = \frac{\beta_e^2 (\Delta E)^2}{1 - \lambda (\Delta E)^2/N},$$

which can be rewritten to obtain the energy and inverse temperature dispersions, $\Delta E$ and $\Delta \beta_\omega$, as follows:

$$\langle \Delta E \rangle^2 = \frac{N}{\beta_e^2 N/C_e + \lambda}, \quad \langle \Delta \beta_\omega \rangle^2 = \frac{1}{N} \frac{\lambda^2}{\beta_e^2 N/C_e + \lambda},$$

where $\Delta x \equiv \sqrt{\langle \delta x^2 \rangle}$. The inverse temperature dispersion $\Delta \beta_\omega$ decreases with increase of the system size $N$ as $\Delta \beta_\omega = |\lambda| \Delta E/N \sim 1/\sqrt{N}$. In the limit $N \to +\infty$, the thermal fluctuations disappear and the present equilibrium situation becomes equivalent to the microcanonical ensemble provided that the following condition:

$$\beta_e^2 N/C_e + \lambda > 0$$

holds. According to the first-order approximation (20), the coupling constant $\lambda$ can be related to the heat capacity $C_\omega$ of the environment as $\lambda = \beta_e^2 N/C_\omega$; hence, the stability condition (33) is fully equivalent to Thirring’s constraint (22).

At first glance, it is desirable to maximally reduce the thermal dispersions of the system energy $E$ and its inverse temperature $\beta = 1/T$ indirectly derived from the environment inverse temperature $\beta_\omega = 1/T_\omega$. However, the inequality of equation (15) imposes an important limitation on the precision of such a measuring process: a reduction of the thermal uncertainties $\Delta(1/T_\omega - 1/T)$ affecting the temperature equalization provokes an increasing of the system energy fluctuations $\Delta E$, and vice versa. As previously discussed [8, 9], equation (15) accounts for the existence of some type of complementarity.
of character between the thermodynamic quantities of energy and temperature. In consequence, one must assume the existence of non-vanishing thermal uncertainties for the system energy \( E \) and its inverse temperature \( \beta \).

According to expressions (32), the increase of the coupling constant \( \lambda \) allows us to reduce the energy dispersion \( \Delta E \), but its value should not be excessively large because its increasing also leads to increasing of the inverse temperature dispersion \( \Delta \beta \). A simple criterion for providing the best value for \( \lambda \) is obtained after minimizing the total dispersion 
\[
\Delta^2_T = (\Delta E)^2/N + N(\Delta \beta)^2.
\]
On introducing the microcanonical curvature \( \kappa \):
\[
\kappa = \kappa(E) = \beta^2 N/C \equiv -N\partial^2 S(E)/\partial E^2,
\]
which is defined in terms of the second derivative of the entropy with the opposite sign, the optimal value (opt) for \( \lambda \) is given by
\[
\lambda_{\text{opt}}(\kappa) = \sqrt{\kappa^2 + 1} - \kappa.
\]
The microcanonical caloric curve \( \beta(E) \) and the heat capacity \( C(E) \) derived from equations (25) and (31) should not depend on the coupling constant \( \lambda \) as long as this latter parameter fulfills condition (33) and the system size \( N \) is sufficiently large. Nevertheless, the use of its optimal value (35) minimizes the underlying thermal fluctuations, which should also reduce the number of steps needed to ensure the convergence of a MC run.

2.2.2. Iterative schemes. Given a certain dependence of the environment inverse temperature \( \beta_{\text{env}}(E) \), one can obtain from a MC run suitable estimations of the system inverse temperature \( \beta_i \) and its heat capacity \( C_i \) at the \( i \)th interception point \( E_i \). These estimated values can be employed to provide the following dependence \( \beta_{\text{opt}}^{(i+1)}(E) \) for performing an analogous estimation at the \( i+1 \)st neighboring point \( E_{i+1} \). To fix some ideas, let us denote by \( \varepsilon = E/N \) the system energy per particle. The \( i \)th dependence of the environment inverse temperature (29) considered for the MC calculation of the \( i \)th point \( \varepsilon_i \) is given by
\[
\beta_{\omega}^{(i)}(\varepsilon) = \bar{\beta}_i + \lambda_i (\varepsilon - \bar{\varepsilon}_i),
\]
where \( \bar{\varepsilon}_i \) and \( \bar{\beta}_i \) are some rough estimates of the expectation values \( \bar{\varepsilon}_i = \langle \varepsilon \rangle \) and \( \bar{\beta}_i = \langle \beta_{\omega}^{(i)} \rangle \). The parameters \( (\bar{\varepsilon}_i, \bar{\beta}_i, \lambda_i) \) could be provided in an interactive way by using the curvature
\[
\kappa_i = \beta^2_i N/C_i
\]
derived from the previous MC simulation:
\[
\bar{\varepsilon}_{i+1} = \bar{\varepsilon}_i + \Delta \varepsilon, \quad \bar{\beta}_{i+1} = \bar{\beta}_i - \kappa_i \Delta \varepsilon, \quad \lambda_{i+1} = \lambda_{\text{opt}}(\kappa_i),
\]
where \( \Delta \varepsilon \) is a small energy step. The initial value \( \bar{\varepsilon}_0 \) is assumed to be the expectation value \( \langle \varepsilon \rangle \) of the energy per particle obtained from any canonical MC simulation with inverse temperature \( \beta_B = \bar{\beta}_0 \) sufficiently away from the anomalous region with \( C < 0 \).

The stability of the iterative procedure previously described depends on the precision of the estimated value of the microcanonical curvature \( \kappa \). Of course, alternative iteration schemes are also possible, e.g., the following scheme:
\[
\bar{\varepsilon}_{i+1} = \langle \varepsilon \rangle_i + \Delta \varepsilon, \quad \bar{\beta}_{i+1} = \langle \beta_{\omega} \rangle_i,
\]
which forces with \( \Delta \varepsilon > 0 \) (\( \Delta \varepsilon < 0 \)) a forward (backward) motion of the expectation values \( \langle \varepsilon \rangle \) and \( \langle \beta_{\omega} \rangle \) along the system caloric curve \( \beta(\varepsilon) \) whenever the value of the coupling parameter \( \lambda \) does not change in a significant way. In general, the inverse temperature \( \beta(\varepsilon) \)
of a large enough short-range interacting system shows a plateau within the anomalous region with $C < 0$, which means that the corresponding curvature $\kappa$ does not significantly differ from zero, $\kappa \simeq 0$. In such cases, the value $\lambda \simeq 1$ constitutes a suitable approximation within the energy range containing the anomalous region with $C < 0$, while the value $\lambda = 0$ corresponding to the canonical ensemble is a good choice elsewhere.

2.2.3. Implementation. Since the inverse temperature $\beta_B$ of the Gibbs canonical ensemble (1) appears as a driving parameter in the transition probability $W(\mathbf{X} \rightarrow \mathbf{X}' ; \beta_B)$ of canonical MC methods, the most general idea for extending such algorithms is replacing one can always find a certain value $\beta_B$ of the environment inverse temperature $\beta_B$ by a variable inverse temperature $\beta_w(E)$, $\beta_B \rightarrow \beta_w(E)$. Moreover, it is desirable that the transition probability resulting from such a modification fulfills the so-called detailed balance condition [7]:

$$p(X)W(X \rightarrow X') = p(X')W(X' \rightarrow X),$$

(39)

where $p(X)$ is the system distribution function, which is given by the function $p_\omega(X) = \omega[E(X)]$.

Let us denote the energy variation during the configuration change $X \rightarrow X'$ as $\delta E = E(X') - E(X)$ and its mean value as $E^m = [E(X) + E(X')] / 2$. According to the mean value theorem, one can express the variation of the function $\log p_\omega(X)$ by using the environment inverse temperature $\beta_w(E)$ evaluated at a certain intermediate energy $E_\theta = E^m + \theta \delta E$ as follows:

$$\log p_\omega(X') - \log p_\omega(X) = -\beta_w^* \delta E,$$

(40)

where $\theta$ is a real parameter with $|\theta| \leq 1$ and $\beta_w^* \equiv \beta_w(E_\theta)$. By considering the transition probability of any canonical MC algorithm that obeys the detailed balance condition:

$$\frac{W(X \rightarrow X'; \beta_B)}{W(X' \rightarrow X; \beta_B)} \equiv \exp(-\beta_B \delta E),$$

(41)

as well as equation (40), one can finally obtain

$$p_\omega(X)W(X \rightarrow X'; \beta_w^*) = p_\omega(X')W(X' \rightarrow X; \beta_w^*).$$

(42)

This last result clarifies that given the initial and final system configurations, $X$ and $X'$, one can always find a certain value $\beta_w^* \equiv \beta_w(X, X')$ of the bath inverse temperature that fulfills the detailed balance condition (39) for the present distribution function $p_\omega(X)$. In particular, the exact value of $\beta_w^*$ for the Gaussian ensemble (30) is the one corresponding to $\theta = 0$, $\beta_w^* = \beta_w(E^m)$.

The main obstacle to performing a direct application of equation (42) to extend canonical MC methods rests on the fact that the final system configuration $X'$ is a priori unknown in many non-local MC algorithms [12]–[14], [22]–[27]. Consequently, one should consider some suitable approximation $\beta_w^* = \beta_w(E_{\theta^*})$ of the exact value $\beta_w^*$, e.g., the one corresponding to the environment inverse temperature at the initial system configuration $X$, $\beta_w^* = \beta_w[E(X)]$. To estimate the error involved in this last approximation, it is important to take into account that the environment heat capacity $C_\omega$ and the energy change $\delta E$ behave with increase of the system size $N$ as $C_\omega \sim N$ and $\delta E \sim N^\alpha$. Here, the exponent $\alpha$ ranges from zero for a local algorithm such as the Metropolis importance sampling up to 1/2 for a hypothetical non-local algorithm able to obtain an effective
independent configuration after each MC step\textsuperscript{4}. Consequently, the difference \( \delta_\beta = |\beta_\omega^i - \beta_\omega^*| \) merely constitutes a small size effect:

\[
\delta_\beta \sim (\beta_\omega^m)^2 \frac{1}{C_\omega^m} |\delta E| \sim \frac{1}{N^{1-\alpha}},
\]

(43)

where the index \( m \) indicates that the corresponding quantities \( \beta_\omega \) and \( C_\omega \) have been evaluated at the energy value \( E^m \). Since the estimated inverse temperature \( \beta_\omega^i \) depends on the system energy \( E(X) \), the corresponding distribution function associated with this approximation \( p_\omega^i(X) \) can also be expressed as a certain function of the system energy, \( p_\omega^i(X) = \omega^*[E(X)] \). According to equation (43), the corresponding inverse temperature:

\[
\beta_\omega^*(E) = -\frac{\partial \log \omega^*(E)}{\partial E}
\]

(44)

cannot differ in a significant way from the exact dependence \( \beta_\omega(E) \) as long as the system size \( N \) as sufficiently large.

The extended canonical MC methods based on an estimation of the inverse temperature \( \beta_\omega^* \) do not fulfill the detailed balance condition (39). However, this fact does not represent any fundamental difficulty since one practically obtains the same numerical results for the caloric curve \( \beta(E) \) and the heat capacity \( C(E) \) with the help of equations (25) and (31) by using a slightly different probabilistic weight \( \omega^*(E) \). The only requirement is that the energy distribution function \( \rho(E) \) exhibits a sharp Gaussian profile, which is simply achieved when the size \( N \) of the system under study is large enough. While the super-critical slowing down of canonical MC methods observed in systems undergoing a first-order phase transition becomes more severe as \( N \) increases, the errors associated with all approximations assumed here become more and more negligible. This is the reason for the present methodology being particularly useful for avoiding such slow sampling problems in large-scale MC simulations.

As is naturally expected, finite size effects can be significant when one is also interested in describing systems whose size \( N \) is not so large. If this is the case, it is necessary not only to implement extended MC schemes that fulfill the detailed balance condition (39), but also to include some finite size corrections into equations (25) and (31) employed here to obtain the microcanonical dependences \( \beta(E) \) and \( C(E) \). Although the complete analysis of these questions is beyond the scope of the present paper, we would like to clarify that the most general way to fulfill the detailed balance condition (39) after the consideration of an estimated value for the transition inverse temperature \( \beta_\omega^* \) is to introduce \textit{a posteriori} an acceptance probability \( w \):

\[
w = \min \left\{ 1, \frac{W_{f\rightarrow i}}{W_{i\rightarrow f}} \exp \left( -\beta_\omega^* \delta E \right) \right\}
\]

(45)

for accepting or rejecting the final configuration \( X' \). Here, \( W_{i\rightarrow f} = W[X \rightarrow X'; \beta_\omega^i] \) and \( W_{f\rightarrow i} = W[X' \rightarrow X; \beta_\omega^f] \), with \( \beta_\omega^i = \beta_\omega(X) \) and \( \beta_\omega^f = \beta_\omega(X') \), represent the transition probabilities of the direct and the reverse process, respectively. The mathematical forms of these depend on the particularities of each non-local MC method.

\textsuperscript{4} The exponent \( \alpha = 1/2 \) follows from the size behavior of the energy dispersion \( \Delta E \sim \sqrt{N} \).

doi:10.1088/1742-5468/2010/02/P02002
3. Application examples

3.1. The model

For the sake of simplicity, let us consider in the present study the \( q \)-state Potts model \[13\]:

\[
H_q = \sum_{\langle ij \rangle} (1 - \delta_{\sigma_i \sigma_j})
\]

(46)

defined on a \( d \)-dimensional hypercubic lattice \( N = L_1 \times L_2 \times \cdots L_d \) with \( L_i = L \) and periodic boundary conditions. Here, the sum is only over pairs of nearest neighbor lattice sites and \( \sigma_i = [1, 2, \ldots, q] \) is the spin state at the \( i \)th site. This model system is just a generalization of the known Ising model:

\[
H_{\text{Ising}} = -\sum_{\langle ij \rangle} s_i s_j,
\]

(47)

with \( s_i = \pm 1 \), which appears as a particular case with \( q = 2 \) after considering a linear transformation among their respective energies per particle \( \varepsilon_{\text{Ising}} = -d + 2\varepsilon_{2\text{-Potts}} \) and inverse temperatures \( \beta_{\text{Ising}} = \frac{1}{2}\beta_{2\text{-Potts}} \).

As discussed elsewhere \[7\], this model system undergoes a first-order phase transition when \( q > 4 \) for \( d = 2 \) and \( q \geq 3 \) for \( d = 3 \), and hence, the microcanonical caloric curves \( \beta(\varepsilon) \) for these model realizations exhibit the back-bending behavior associated with the presence of macrostates with \( C < 0 \) such as the one sketched in figure 1. In addition to the consideration of a local MC method such as the Metropolis importance sampling \[10,11\], the canonical MC study of Potts models can be carried out by using other accelerating methods such as the Swendsen–Wang \[12,13\] or Wolff \[14\] cluster algorithms. However, none of these MC algorithms can account for the existence of an anomalous regime with \( C < 0 \), and they also even suffer from the existence of a super-critical slowing down as a direct consequence of the bimodality of the canonical energy distribution function. The existence of several canonical MC algorithms for such models enables us to perform a comparative study among their respective extended versions described below.

3.2. Monte Carlo methods

3.2.1. Metropolis importance sampling. The simplest and most general way to implement a thermal coupling with a bath at constant inverse temperature \( \beta_B \) is by using the Metropolis importance sampling \[10,11\]. In this method, a Metropolis move is carried out as follows:

1. A site \( i \) is chosen at random, and the initial spin state \( \sigma_i \) is changed (also at random) by considering any other one of its \( q \) admissible values.

2. This move, from an initial state with energy \( E \) and variation \( \delta E \), is accepted in accordance with the transition probability

\[
W (E \rightarrow E + \Delta E; \beta_B) = \min \{ \exp [\beta_B \delta E] , 1 \}.
\]

(48)

A MC step is produced after considering \( L^d \) moves regardless of whether they have been accepted or rejected.

The extension of this local algorithm with the consideration of an environment associated with an arbitrary probabilistic weight \( \omega(E) \) is achieved by using the transition
probability (6). Clearly, equation (6) fulfills the detailed balance condition (39). When
the system size $N$ is sufficiently large, such a transition probability is practically given
by equation (7), which is exactly the canonical transition probability (48) modified by
the inclusion of a variable inverse temperature, $\beta_B \rightarrow \beta_\omega(E)$. Remarkably, the transition
probability (6) can be exactly rewritten in the form (42) by using the value $\beta_\omega^\star = \beta_\omega(E_m)$
corresponding to $\theta = 0$ for the particular case of the linear ansatz (29). This is a useful
representation because both the initial and the final system configurations, $X$ and $X'$, are
a priori known for this local MC method.

3.2.2. Cluster algorithms. Among other important MC methods are the so-called cluster
algorithms, which are usually more efficient than any local MC method (such as
Metropolis). However, the success of these methods is not universal because the proper
cluster moves needed seem to be highly dependent on the system, and efficient cluster
MC methods have only been found for a small number of models [12]–[14], [22]–[27].

The idea of using non-local moves was first suggested by Swendsen and Wang [12,13]
for the case of the Ising model and its generalizations, the Potts models. Such cluster
algorithms are based on a mapping of this model system to a random cluster model of
percolation through the equation

$$Z(\beta_B) = \sum_{\{\sigma\}} \exp[-\beta_B H_\sigma] = \sum_{\{\sigma\}} q^{N_c} p^b (1-p)^{N-d-b}, \quad (49)$$

where $p = p(\beta_B) = 1 - e^{-\beta_B}$, $b$ is the number of bonds and $N_c$ the number of clusters. We
shall consider in the present study the Swendsen–Wang cluster algorithm, whose scheme
reads as follows:

1. Examine each nearest neighbor pair and create a bond with probability $p(\beta_B)$. That
   is, if the two nearest neighbor spins are the same, a bond is created between them
   with probability $p(\beta_B)$; if the spin values are different, there will be no bond.
2. Identify clusters as a set of sites connected by zero or more bonds (i.e., a connected
   component of a graph). Relabel each cluster with a fresh new value at random.

The extension of this cluster MC method by using the present ideas is achieved by
introducing a third step:

3. Redefine the inverse temperature of the bath $\beta_B = \beta_\omega(E_i)$ employed to obtain the next
   system configuration $X_{i+1}$ by using the energy $E_i = E(X_i)$ of the present configuration
   $X_i$.

While the bath inverse temperature $\beta_B$ is redefined after every local move in the
Metropolis importance sampling, such a redefinition only takes place in a cluster algorithm
such as the Swendsen–Wang one after every MC step because the cluster moves demand
a constancy of the bath temperature. The present method is a simple example of an
extended canonical MC algorithm that does not fulfill the detailed balance condition (39)
due to its using the approximated inverse temperature $\beta_\omega^\star = \beta_\omega(E_i)$ instead of the exact
value $\beta_\omega^\star = \beta_\omega(E_m)$ with $E_m = [E_i + E_{i+1}] / 2$. As is naturally expected, the violation
of the detailed balance condition introduces finite size effects in the calculation of the expectation values of the thermal fluctuations, which affects the calculation of the heat
Figure 2. Microcanonical caloric and curvature curves, $\beta(\varepsilon)$ (back-bending dependences) and $\kappa(\varepsilon)$ ($U$-like dependences), obtained from MC simulations by using the extended version of the Metropolis importance sampling algorithm for some realizations of 2D $q$-state Potts models.

capacity $C$ via equation (31). We have verified by means of preliminary calculations that the use of the transition inverse temperature $\beta^{\ast}(E_{m_{i}})$ instead of the instantaneous value $\beta_{\omega}(E_{i})$ in the MC estimation of the expectation values:

$$
\langle \beta_{\omega}(E)A(E) \rangle \sim \frac{1}{M} \sum_{i=1}^{M} \beta_{\omega}(E_{i})A(E_{i}) \longrightarrow \frac{1}{M} \sum_{i=1}^{M} \beta^{\ast}_{\omega}(E_{m_{i}})A(E_{i})
$$

(50)

significantly reduces the occurrence of such undesirable errors.

3.3. Results and discussion

Results of extensive MC calculations using the extended version of Metropolis importance sampling (hereafter referred to as extended MIS) are shown in figure 2. We are limited to considering here the 2D Potts models with $L = 25$ for several values of $q$, where each point of these curves has been obtained from a MC run with $10^6$ steps. While the microcanonical curvature curve $\kappa(\varepsilon)$ for the case $q = 4$ only slightly touches the horizontal line with $\kappa = 0$, the other cases with $q > 4$ clearly exhibit negative values $\kappa(\varepsilon) < 0$, that is, macrostates with negative heat capacities $C < 0$. This last observation evidences that the 2D Potts model with $q = 4$ undergoes a continuous phase transition, while the phase transition is discontinuous (of first order) for those cases with $q > 4$ [28].

As expected, the anomalous macrostates with $C < 0$ cannot be accessed by using the ordinary Metropolis importance sampling (hereafter referred to as canonical MIS). This limitation is explicitly shown in figure 3 for the particular case of the 2D seven-state Potts model. These results constitute a simple exemplification of the schematic behavior represented in figure 1. To allow a better understanding, we also show here the corresponding energy distribution functions and the dynamical evolutions of the average system energy $\langle \varepsilon \rangle$ (inset panel) obtained from MC runs by considering both canonical MIS and extended MIS algorithms with $\beta_{e} = 1.682$.

doi:10.1088/1742-5468/2010/02/P02002
Extending canonical Monte Carlo methods

Figure 3. Comparison between the caloric curves of the 2D seven-state Potts model with $L = 25$ obtained from MC simulations by using the canonical Metropolis importance sampling (canonical MIS) and its extended version (extended MIS). The distribution functions as well as the evolution of the average system energy $\langle \varepsilon \rangle$ shown in the inset panel correspond to MC runs using these methods with $\beta_e = 1.682$. See the text for further explanations.

The stationary macrostates at $\varepsilon_a$ and $\varepsilon_c$ derived from the intersection of the microcanonical curve $\beta(\varepsilon)$ and the bath inverse temperature $\beta_B = \beta_e = 1.682$ are thermodynamically stable within the canonical ensemble. Consequently, the canonical energy distribution function is bimodal and its peaks are related to the coexisting phases. As a consequence of such bimodality, the system energy exhibits eventual random transitions between the coexisting peaks, which lead to a slow equilibration of the corresponding average energy $\langle \varepsilon \rangle$ (inset panel).

The thermodynamic behavior radically changes when this system is put in a thermodynamic situation with non-vanishing correlated fluctuations $\langle \delta \beta_\omega \delta E \rangle \neq 0$, which is implemented here by considering a dependence $\beta_\omega(\varepsilon) = \beta_e + \lambda(\varepsilon - \varepsilon_c)$, where $\beta_e = 1.682$, $\lambda = \lambda_{opt}(\kappa_b) \neq 0$ and $\varepsilon_c = \varepsilon_b$, with $\kappa_b$ being the curvature at the stationary point $\varepsilon_b$ located within the anomalous region with $C < 0$. The canonically stable stationary points $\varepsilon_a$ and $\varepsilon_c$ become thermodynamically unstable and their corresponding peaks disappear from the energy distribution function. Conversely, the canonically unstable stationary point $\varepsilon_b$ now becomes thermodynamically stable, and its position determines the maximum of the only peak of the energy distribution function. Once the bimodal character of the energy distribution function is suppressed, the average system energy $\langle \varepsilon \rangle$ shows fast convergence towards its equilibrium value $\varepsilon_b$ (inset panel).

A comparative study of a 2D seven-state Potts model with $L = 25$ by using the Swendsen–Wang cluster algorithm is shown in figure 4. Although the canonical Swendsen–Wang method (hereafter referred to as canonical SW) is more efficient than the canonical MIS, it cannot provide a description of the existence of an anomalous region with $C < 0$ and suffers from a slow relaxation in this kind of situation, as is clearly shown in figure 4. Such limitations are circumvented by using its extended version (hereafter referred to as extended SW). Like the previously discussed example, the extended SW algorithm

doi:10.1088/1742-5468/2010/02/P02002
Extending canonical Monte Carlo methods

Figure 4. Comparison among the caloric curves of the 2D seven-state Potts model with $L = 25$ obtained from MC simulations using the canonical Swendsen–Wang cluster algorithm (canonical SW) and its extended version (extended SW), as well as the extended Metropolis importance sampling method (extended MIS). The energy distribution functions and the dynamical evolution of the average system energy $\langle \varepsilon \rangle$ shown in the inset panel correspond to MC runs using the above cluster algorithms with $\beta_e = 1.29$.

eliminates the bimodality of the energy distribution function for $\beta_e = 1.29$ and leads to a fast convergence of the average energy per particle during the simulation. Note also the remarkable agreement between the extended MIS and extended SW algorithms despite the latter not fulfilling the detailed balance condition (39). This is clear evidence that a suitable approximation of the inverse temperature $\beta^*_e$ of equation (42) is enough to provide a precise estimation of the microcanonical caloric curve $\beta(E)$ as long as the system size is sufficiently large.

The comparison among the above extended canonical MC algorithms and the Wang–Landau sampling method [18] is shown in figure 5. Although these results exhibit very good consistency, one can note small but appreciable discrepancies within the anomalous region with $C < 0$. These relative differences are naturally expected for two reasons. Firstly, a direct numerical differentiation of the entropy $S = \log \Omega$ (inset panel) obtained from the Wang–Landau method enhances the underlying statistical errors of its MC calculation, and therefore, the final result crucially depends on how one defines this mathematical operation for this discrete observable. On the other hand, the work in equation (25) is supported by the Gaussian character of the energy distribution function $\rho(\varepsilon)$, which arises as an asymptotic distribution when the system size $N$ is sufficiently large. Clearly, the distribution function undergoes small deviations from the Gaussian shape when $N$ is not as large. We shall show in a forthcoming paper that these size effects can be taken into account to improve the precision of some extended canonical MC algorithms.

Overall, it is important to remark that the present proposal is focused on the solution of the slow sampling problems observed in large-scale canonical MC simulations, that is, in systems undergoing a temperature driven discontinuous phase transition with sizes
Figure 5. Comparative study of the extended MIS and SW algorithms and the Wang–Landau sampling method. Despite the relatively good agreement, one can appreciate some small discrepancies, which could be mainly associated with size effects.

Figure 6. Microcanonical caloric curves of the ten-state Potts models for lattice dimensions $d = (2, 3, 4)$ with a fixed number of lattice sites $N = L^d = 4096$, which were obtained from MC simulations by using the extended SW algorithm with $2 \times 10^4$ MC steps for each calculated point. The plot is performed in terms of the re-scaled variables $\varepsilon^d = 2\varepsilon/d$ and $\beta^d = \beta d/2$. Inset: comparison between the extended SW and Wang–Landau methods for the 2D ten-state Potts model with $L = 64$.

sufficiently large to support the Gaussian approximation. In particular, the previous extended canonical MC algorithms can be useful for studying Potts models with many spin states and higher dimensions, such as the cases shown in figure 6, where we illustrate the caloric curves derived from MC simulations by using the extended SW algorithm for $q = 10$, $d = (2, 3, 4)$, and a fixed number of lattice sites $N = L^d = 4096$. The comparative study case with the Wang–Landau method shown in the inset panel allows us to verify that the agreement between these methods becomes more significant with increase of $N$. doi:10.1088/1742-5468/2010/02/P02002
To quantitatively characterize the efficiency of the previously discussed extended canonical MC methods, one should obtain the decorrelation time \( \tau \), that is, the minimum number of MC steps needed to generate effectively independent, identically distributed samples in the Markov chain. Its calculation is performed here by using the expression

\[
\tau = \lim_{M \to \infty} \tau_M = \lim_{M \to \infty} \frac{M \cdot \text{var}(\varepsilon_M)}{\text{var}(\varepsilon_1)},
\]

where \( \text{var}(\varepsilon_M) = \langle \varepsilon_M^2 \rangle - \langle \varepsilon_M \rangle^2 \) is the variance of \( \varepsilon_M \), which is defined as the arithmetic mean of the energy per particle \( \varepsilon \) over \( M \) samples (consecutive MC steps):

\[
\varepsilon_M = \frac{1}{M} \sum_{i=1}^{M} \varepsilon_i.
\]

This quantity is calculated for the particular MC runs considered in figures 3 and 4. This study allows us to verify that the use of a variable dependence \( \beta_\omega(\varepsilon) \) instead of a constant parameter \( \beta_B \) enables the reduction of the decorrelation time of the Metropolis importance sampling from \( \tau \simeq 450 \) to 80. The improvement is even more significant for the Swendsen–Wang cluster algorithm, which experiences a reduction of the decorrelation time from \( \tau \simeq 300 \) to 12.

Let us now analyze the behavior of the decorrelation time \( \tau \) with the increasing system size \( N = L^d \) at the inverse temperature \( \beta_c \) of the first-order phase transition. Such a study demands the performance of preliminary calculations of the quantity \( \beta_c \) due to its underlying dependence on the system size \( N \). Because of computational limitations, we decide to restrict our analysis to the case of the 2D ten-state Potts model with \( L = (8, 16, 32, 64) \), whose results are shown in figure 7. The entropy per particle \( s(\varepsilon) \) was obtained from the second-order approximation of the power expansion \( \Delta s = \beta \Delta \varepsilon - \frac{1}{2} \kappa \Delta \varepsilon^2 \), which allows a direct calculation of this thermodynamic function through the inverse temperature \( \beta \) and curvature \( \kappa \). These numerical results are shown in panel (b) of figure 7—or more exactly, the quantity \( s^*(\varepsilon) = s(\varepsilon) - \beta_c \varepsilon - s_0 \), with \( s_0 \) being a suitable constant, which allows us to appreciate better the ‘convex intruder’ of the entropy accounting for the existence of macrostates with \( C < 0 \). The mathematical form of this last anomaly enables the acquisition of the latent heat \( q_L = \varepsilon_3 - \varepsilon_1 \) and the entropy loss per particle \( \Delta s_{\text{surf}} \) associated with the existence of surface correlations [3], and \((\varepsilon_1, \varepsilon_2, \varepsilon_3)\) represent the three stationary points where the caloric curve \( \beta(\varepsilon) \) takes the value of the phase transition inverse temperature \( \beta_c \). It worth clarifying that the critical value \( \beta_c \) is simply the point of discontinuity of the first derivative of the Planck thermodynamic potential \( p(\beta) \) estimated from the microcanonical entropy \( s(\varepsilon) \) via the Legendre transformation:

\[
p(\beta) = \min_E [\beta \varepsilon - s(\varepsilon)].
\]

In fact, once one has obtained the microcanonical entropy \( S(E) \), one can calculate any thermo-statistical quantity by using the canonical distribution function:

\[
dp(E|\beta) = Ae^{-\beta E + S(E)} \, dE.
\]

Relevant physical observables and thermodynamic parameters derived from this type of analysis are summarized in table 1.
Figure 7. Study of the thermodynamic behavior of the 2D ten-state Potts model for different lattice sizes $L$: panel (a), microcanonical caloric curves $\beta(\varepsilon)$; panel (b), microcanonical entropy plotted in terms of the quantity $s^*(\varepsilon) \equiv s(\varepsilon) - \beta_c \varepsilon - s_0$ in order to reveal the convex intruder associated with the existence of surface correlations.

Table 1. Dependence of some physical observables and thermodynamic parameters of the 2D ten-state Potts model on the increasing lattice size $L$.

| $L$ | 8   | 16  | 32  | 64  |
|-----|-----|-----|-----|-----|
| $\beta_c$ | 1.415 | 1.422 | 1.424 | 1.426 |
| $\varepsilon_1$ | 0.319 | 0.319 | 0.321 | 0.329 |
| $\varepsilon_2$ | 0.767 | 0.755 | 0.737 | 0.719 |
| $\varepsilon_3$ | 1.165 | 1.114 | 1.074 | 1.049 |
| $q_L$ | 0.846 | 0.795 | 0.753 | 0.72 |
| $\Delta s_{\text{surf}} \times 10^3$ | 72   | 5.6  | 3.1  | 1.7  |

Results of extensive calculations of dependences of the decorrelation time $\tau$ versus the system size $N$ at the point of phase transition $\beta_c$ are shown in figure 8. It is clearly evident that the extended MC methods are always much more efficient than their respective canonical counterparts. In particular, the extended SW method reduces the exponential growth of the decorrelation time $\tau$ with $N$ to a weak power-law dependence $\tau \propto N^{\alpha}$. doi:10.1088/1742-5468/2010/02/P02002
Extending canonical Monte Carlo methods

Figure 8. Dependences of the decorrelation time $\tau$ versus the system size $N$ at the point of first-order phase transition $\beta_c$ for different MC methods considered in the study of the 2D ten-state Potts model. Inset panel: a more detailed plot for the decorrelation time of the extended SW method, which clearly shows a weak power-law dependence $\tau \propto N^\alpha$ with $\alpha = 0.2$.

with $\alpha = 0.2$. As expected, the extended MIS is less efficient than the extended SW. However, the dependence on $N$ of its corresponding decorrelation time $\tau$ does not differ in a significant way from the one associated with the cluster algorithm. Since the calculation of the decorrelation time $\tau$ of the canonical MC methods demands very large MC runs, we use the $1/M$ extrapolation in order to obtain some rough estimates of the non-equilibrated points.

According to Berg [29], the multicanonical method and its variant are able to reduce the exponential divergence of the decorrelation time $\tau$ to a power dependence with typical exponent $\alpha = 2-2.5$. Although such a power-law behavior accounts for a convergence that is less efficient than the one achieved by using the present proposal, it is important to remark that the multicanonical methods provide possibilities for an effective exploration of the entire energy region in a single MC run. Conversely, our methodology only explores a small energy region $(\varepsilon_c - \sigma, \varepsilon_c + \sigma)$ with a typical width $\sigma \sim 1/\sqrt{N}$, so a minimum of $n \propto \sqrt{N}$ MC runs are required to study the same energy range as multicanonical methods. Thus, the size dependence of the number of MC samples $M$ can be estimated by considering the number of MC runs $n$ and the decorrelation time $\tau$ as follows:

$$M \sim n\tau \sim N^{\alpha^*},$$

which leads to the effective exponent $\alpha^* \simeq 0.7$. Such an effective power-law growth still considers a convergence more efficient than the one associated with multicanonical methods, overall, when one also takes into account the fact that these re-weighting techniques demand a preliminary reconstruction of the probabilistic weight $\omega(E)$, which is a procedure that consumes a significant fraction of the available computational resources.

Let us finally reconsider the question of the optimal value of the coupling constant $\lambda$ that allows us to achieve the best efficiency by using an extended canonical MC algorithm. An example of such a study is shown in figure 9, where we illustrate the behavior of the
Figure 9. Dependence of the decorrelation time $\tau$, the total dispersion $\Delta^2_T$ and the efficiency factor $\eta = \tau \Delta^2_T$ on the coupling constant $\lambda$ obtained from MC simulations of the 2D ten-state Potts model with $L = 32$ by using the extended SW cluster algorithm. See the text for further explanations.

decorrelation time $\tau$ and the total dispersion $\Delta^2_T = (\Delta E)^2/N + N(\Delta \beta)^2$ for positive values of the coupling constant $\lambda$. The data set was obtained from MC simulations of the 2D ten-state Potts model with $L = 32$ by using the extended SW cluster algorithm with environment inverse temperature $\beta_\omega(\epsilon) = \beta_c + \lambda(\epsilon - \epsilon_b)$, whose parameters are $\beta_c$ and $\epsilon_2$, in order to access the canonically unstable stationary macrostate $\epsilon_2$ located within the anomalous region with $C < 0$ at the point of the first-order phase transition $\beta_c$.

As expected, the decorrelation time $\tau$ and the total dispersion $\Delta^2_T$ show large values when the coupling constant $\lambda$ is close to zero, as a consequence of the imposition of the external conditions associated with the canonical ensemble (1), where the thermodynamic instability of the macrostates with $C < 0$ and the bimodal character of the energy distribution function $\rho(\epsilon)$ are found. The increase of the coupling parameter $\lambda$ for values that obey the condition (33) produces a progressive reduction of the energy dispersion $\Delta E$ in equation (32). This latter behavior leads to a reduction of the decorrelation time $\tau$ due to the MC sampling performing a faster exploration for a smaller energy region. However, the decorrelation time $\eta$ starts to grow after reaching its minimum value at a certain $\lambda_\eta$. Apparently, the system configurations $X$ cannot be modified in an appreciable way by the MC sampling if the energy can only undergo very small changes.

For a general case of the environment inverse temperature (29) with $\lambda \neq 0$, both the system fluctuating behavior and the decorrelation time depend on the particular value of the coupling constant $\lambda$. In such cases, the efficiency of an extended canonical MC method can be evaluated through the minimal number of MC steps needed to achieve a certain precision in the calculation of a given thermo-statistical observable. The statistical error $\delta\epsilon$ associated with a set of $n$ independent outcomes $\{x_i, i = 1, \ldots, n\}$ can be estimated from the standard deviation $\sigma$ as $\delta\epsilon^2 = \sigma^2/n$. In MC calculations, independent samples are only obtained after considering a number of MC steps equal to the decorrelation time $\tau$. Thus, the number of MC steps $S$ needed to obtain an estimation of a given point $(\epsilon, \beta)$ of the caloric curve with a precision $\sqrt{\delta\epsilon^2 + \delta\beta^2} < \delta a$ can be evaluated in terms of the
Extending canonical Monte Carlo methods

The decorrelation time $\tau$, the total dispersion $\Delta_T^2$ and the system size $N$ as follows:

$$S = \frac{\tau \Delta_T^2}{N \delta a^2} \equiv \frac{\eta}{N \delta a^2}. \quad (56)$$

The quantity $\eta = \tau \Delta_T^2$ could be referred to as the efficiency factor. Clearly, a dynamic criterion for providing an optimal value of the coupling constant $\lambda$ for achieving the best efficiency by using such MC calculations is given by minimizing the efficiency factor $\eta$.

According to the data shown in figure 9, the value of the coupling constant $\lambda_0$ corresponding to the minimum of the efficiency factor $\eta$ is located within the interval between the position of the minimum $\lambda_c$ of the decorrelation time $\tau$ and the minimum $\lambda_s$ of the total dispersion $\Delta_T^2$, $\lambda \in [\lambda_s, \lambda_c]$. Remarkably, the efficiency factor $\eta$ does not change in a significant way within this last interval. By considering this last observation and the fact that the calculation of the correlation time $\tau$ demands extensive calculations, one can realize that the value of the coupling constant $\lambda_s$ corresponding to the minimum of the total dispersion $\Delta_T^2$, equation (35), provides a good value for performing a very efficient MC calculation of the calorific curve by using extended canonical algorithms. In general, it is always convenient to keep the value of the coupling constant $\lambda$ as small as possible. Indeed, the statistical error $\epsilon$ involved in the MC calculation of the standard deviation $\sigma_E \equiv (\Delta E^2)/N$ considered for obtaining the system curvature $\kappa = \beta^2 N/C$ from equation (31) leads to the existence of a statistical error $\delta \kappa \sim (\lambda + 1)(\lambda + \kappa)\epsilon$ which grows with increase of the coupling constant $\lambda$.

4. Final remarks

We have shown that conventional Monte Carlo methods based on the consideration of the Gibbs canonical ensemble (1) can be easily extended in order to capture the existence of an anomalous regime with negative heat capacities and avoid the occurrence of the super-critical slowing down. The key ingredient is replacing the use of a bath with an infinite heat capacity by imposing an environment with a finite heat capacity that obeys Thirring’s constraint (22). Such an equilibrium situation, characterized by the existence of non-vanishing correlations $\langle \delta \beta \delta E \rangle \neq 0$, is inspired by the generalized equilibrium fluctuation-dissipation relation (2), which allows one to introduce several improvements to the methodology of Gerling and Hüller on the basis of the consideration of the dynamical ensemble (26).

The way to introduce an environment with a variable inverse temperature $\beta_\omega$ depends on the specific features of each canonical Monte Carlo method, although such a question seems not to be a difficult problem in the case of classical algorithms. While it could be desirable for the implementation of this kind of methodology to obey the detailed balance condition (39), the application examples considered in section 3 show that one can still obtain a good MC estimation of the microcanonical calorific curve $\beta(E)$ without fulfilling the detailed balance as long as the system under analysis is sufficiently large.

Before concluding, it is worth mentioning that equation (2) constitutes a particular case of a more general equilibrium fluctuation-dissipation theorem, which accounts for the system fluctuation behavior in a thermodynamic situation characterized by the presence of several control parameters [30]. Roughly speaking, this theorem provides a general extension of some other well-known fluctuation relations such as the one involving the
Extending canonical Monte Carlo methods

isothermal magnetic susceptibility and magnetization fluctuations of a magnetic system, \( \chi_T = \beta \langle \delta M^2 \rangle \), or the isothermal compressibility and volume fluctuations of a fluid system, \( V_K T = \beta \langle \delta V^2 \rangle \), which are compatible with the existence of anomalous response functions, e.g., negative isothermal susceptibilities \( \chi_T < 0 \) and negative isothermal compressibilities \( K_T < 0 \). Clearly, this general theorem suggests a direct extension of the present methodology in order to enhance MC methods based on the so-called Boltzmann–Gibbs distributions:

\[
dp_{BG}(E, X) = \frac{1}{Z(\beta, Y)} \exp \left[ -\beta (E + YX) \right] \, dE \, dX
\]

(57)

to account for the existence of macrostates with anomalous response functions.

Acknowledgments

It is a pleasure to acknowledge partial financial support by FONDECYT 3080003. L Velazquez also thanks for partial financial support the project PNCB-16/2004 of the Cuban National Program of Basic Sciences.

References

[1] Padmanabhan T, 1990 Phys. Rep. 188 285
[2] Lynden-Bell D, 1999 Physica A 263 293
[3] Gross D H E, 2001 Microcanonical Thermodynamics: Phase Transitions in Small Systems (Lectures Notes in Physics vol 66) (Singapore: World Scientific).
[4] Gross D H E and Madjet M E, 1997 Z. Phys. B 104 521
[5] Moretto L G, Ghetti R, Phair L, Tso K and Wozniak G J, 1999 Phys. Rep. 287 250
[6] D’Agostino M, Gulminelli F, Chomaz P, Bruno M, Cannata F, Bougault R, Gramegna F, Iori I, Le Neindre N, Margaglioti G, Moroni A and Vannini G, 2000 Phys. Lett. B 473 219
[7] Landau P D and Binder K, 2000 A Guide to Monte Carlo Simulations in Statistical Physics (Cambridge: Cambridge University Press)
[8] Velazquez L and Curilef S, 2009 J. Phys. A: Math. Theor. 42 095006
[9] Velazquez L and Curilef S, 2009 J. Stat. Mech. P03027
[10] Metropolis N, Rosenbluth A W, Rosenbluth M N, Teller A H and Teller E, 1953 J. Chem. Phys. 21 1087
[11] Hastings W K, 1970 Biometrika 57 97
[12] Swendsen R H and Wang J-S, 1987 Phys. Rev. Lett. 58 86
[13] Wang J-S, Swendsen R H and Kotecký R, 1989 Phys. Rev. Lett. 63 109
[14] Wolff U, 1989 Phys. Rev. Lett. 62 361
[15] Thirring W, 1980 Quantum Mechanics of Large Systems (Berlin: Springer) chapter 2.3
[16] Viana Lopes J et al, 2006 Phys. Rev. E 74 046702
[17] Berg B and Neuhaus T, 1991 Phys. Lett. B 267 249
[18] Berg B and Neuhaus T, 1992 Phys. Rev. Lett. 68 9
[19] Wang F and Landau D P, 2001 Phys. Rev. Lett. 86 2050
[20] Wang F and Landau D P, 2001 Phys. Rev. E 64 056101
[21] Gerling A and Hüller R W, 1993 Z. Phys. B 90 207
[22] Challa M S S and Hetherington J H, 1988 Computer Simulation Studies in Condensed Matter Physics I ed D P Landau, K K Mon and H-B Schüttler (Heidelberg: Springer)
[23] Hetherington J H, 1987 J. Low Temp. Phys. 66 145
[24] Edwards R G and Sokal A D, 1988 Phys. Rev. D 38 2009
[25] Niedermayer F, 1988 Phys. Rev. Lett. 61 2026
[26] Evertz H G, Hasenbusch M, Marcu M, Pinn K and Solomon S, 1991 Phys. Lett. B 254 185
[27] Hasenbusch M, Marcu M and Pinn K, 1994 Physica A 211 255
[28] Dress C and Krauth W, 1995 J. Phys. A: Math. Gen. 28 L597
[29] Liu J W and Luijten E, 2004 Phys. Rev. Lett. 92 035504

doi:10.1088/1742-5468/2010/02/P02002
Extending canonical Monte Carlo methods

[28] Wu F Y, 1982 Rev. Mod. Phys. 54 235
[29] Berg B A, 1997 Proc. Int. Conf. on Multiscale Phenomena and Their Simulations (Bielefeld, Oct. 1996) ed F Karsch, B Monien and H Satz (Singapore: World Scientific)
[30] Velazquez L and Curilef S, Equilibrium fluctuation-dissipation relations: a generalization compatible with macrostates with anomalous response functions, 2009 arXiv:0910.2870