

# Temperature and density extrapolations in canonical ensemble Monte Carlo simulations

A. L. Ferreira and M. A. Barroso

*Universidade de Aveiro, Departamento de Física, 3810-193 Aveiro, Portugal*

(Received 14 June 1999)

We show how to use the multiple histogram method to combine canonical ensemble Monte Carlo simulations made at different temperatures and densities. The method can be applied to study systems of particles with arbitrary interaction potential and to compute the thermodynamic properties over a range of temperatures and densities. The calculation of the Helmholtz free energy relative to some thermodynamic reference state enables us to study phase coexistence properties. We test the method on the Lennard-Jones fluids for which many results are available.

PACS number(s): 05.10.Ln, 05.20.Jj, 64.70.Fx

## I. INTRODUCTION

Histogram and multiple-histogram methods have been proposed as an optimized way of analyzing Monte Carlo data [1–3]. These methods can be included in the more general class of reweighting methods [4]. The idea is to combine a given set of standard Monte Carlo simulations to get improved estimates of observables in a given parameter region. A related idea is to sample a suitably chosen probability distribution rather than a given statistical mechanics ensemble. The sampling distribution is such that the configuration space visited is typical of the interval of thermodynamic parameters of interest thus allowing the reconstruction of the appropriate statistical mechanics ensemble. The methods of umbrella sampling [5–7], multicanonical [8,9] and expanded ensemble methods [10,11] can be seen as belonging to this class.

In this paper we show how to combine NVT Monte Carlo simulations made at different temperatures and volumes. Our method is a generalization to volume extrapolations of the multiple histogram method. We further show that the method can be applied not only to systems of particles that interact through interaction potentials that have a simple scaling with particle distance but also to those with arbitrary distance dependence. Furthermore, as we are able to calculate relative free energies as a function of volume and temperature, the method can be applied to study phase coexistence properties [6,7].

Several simulation methods have been proposed to study phase coexistence properties. In the Gibbs Monte Carlo ensemble [12,13] two simulation boxes equilibrate by exchanging particles and volume and the system separates into two phases, each one located in one of the boxes. Grand canonical ensemble simulations with multiple histogramming have been used to study critical properties and finite-size scaling in fluid systems [14–16]. However, due to the low probability of particle exchanges or insertions at high densities these methods cannot be used to study dense phases. For such systems special methods have been suggested that rely on the calculation of the absolute free energies of the two phases [17]. Recently an alternative method based on Gibbs-Duhem integration was proposed and the concept of pseudoensembles was introduced [18–20]. For all of these methods the use of the multiple histogram technique can be a valuable

auxiliary tool [17,20]. Our method can also be applied to the study of solid fluid coexistence [21]. An introduction to the application of simulation techniques to the study of phase coexistence properties can be found in the book of Frenkel and Smit [22].

## II. METHOD

Consider a system of  $N$  interacting particles contained in a box of volume  $V_0$ . For simplicity we consider a pairwise additive interaction potential and  $\{\vec{r}_i\}$  denotes a given configuration of particle coordinates. The total potential energy of the system in a configuration is given by  $E(\{\vec{r}_i\}) = \sum_{\langle i,j \rangle} u(|\vec{r}_i - \vec{r}_j|)$ , where the sum runs over all pairs of particles. Uniformly expanding the system from the volume  $V_0$  to the volume  $V$  changes the configuration from  $\{\vec{r}_i\}$  to  $\{\vec{r}'_i\}$ , such that  $\vec{r}'_i = (V/V_0)^{1/3} \vec{r}_i$ . The energy of the system of volume  $V$  in the new configuration is given by  $E(\{\vec{r}'_i\}) = \sum_{\langle i,j \rangle} u[(V/V_0)^{1/3} |\vec{r}_i - \vec{r}_j|]$ .

We will show that it is always possible to find a set of  $n_c$  variables,  $C_n(\{\vec{r}_i\})$ , with  $0 \leq n \leq n_c - 1$ , that depend on the particle coordinates. These variables can be seen as coordinates of a column vector,  $\vec{C} = (C_0, C_1, \dots, C_{n_c-1})$ . Their choice is arbitrary provided two properties are fulfilled. First, it should be possible to write the potential energy in terms of these variables. Second, there is a known linear relation between the value of the variables in the expanded system of volume  $V$  and their value for the system of volume  $V_0$ :

$$\vec{C}(\{\vec{r}'_i\}) = \mathbf{M} \cdot \vec{C}(\{\vec{r}_i\}), \quad (1)$$

$\mathbf{M}$  being a square matrix with coefficients that depend only on  $V$  and  $V_0$ .

For example, for the Lennard-Jones potential,  $u(r) = 4\epsilon[(\sigma/r)^{12} - (\sigma/r)^6]$ , the vector  $\vec{C}$  can be chosen with only two components,  $C_0$  and  $C_1$

$$C_0(\{\vec{r}_i\}) = \sum_{\langle i,j \rangle} \left( \frac{\sigma}{r_{ij}} \right)^{12}, \quad (2a)$$

$$C_1(\{\vec{r}_i\}) = \sum_{(i,j)} \left( \frac{\sigma}{r_{ij}} \right)^6, \quad (2b)$$

satisfying the two properties mentioned above.

For an arbitrary potential it may not be possible to find variables  $C_n$  with a given volume scaling as in the Lennard-Jones case. However, there is always a method based on volume expansions that we describe next. We define the coefficient  $C_n$  from the volume derivatives of  $E[(V/V_0)^{1/3}\{\vec{r}_i\}]$ :

$$C_n(\{\vec{r}_i\}) = \left( \frac{\partial^n E[(V/V_0)^{1/3}\{\vec{r}_i\}]}{\partial V^n} \right)_{V_0}. \quad (3)$$

The two properties are fulfilled since the energy of a given configuration is  $E(\{\vec{r}_i\}) = C_0(\{\vec{r}_i\})$  and the series expansion

$$C_n(\{\vec{r}_i\}) = \sum_{l=n}^{\infty} \frac{C_l(\{\vec{r}_i\})}{(l-n)!} (V-V_0)^{l-n}, \quad (4)$$

provide the linear relation (1). However the vector  $\vec{C}$  has an infinite number of components. In practical numerical work the above expansion needs to be stopped at a sufficiently high-order. As it will be seen below the approximation introduced can be controlled either by increasing the order of the approximation or by combining simulations at closer densities.

We denote the density of states with variables  $\vec{C}(\{\vec{r}_i'\})$  in some neighborhood of  $\vec{c}(V)$  for a system of volume  $V$ , by  $\Omega[\vec{c}(V), V]$ . This quantity can be obtained from a phase space integration,

$$\Omega[\vec{c}(V), V] = \int_V d\vec{r}'_1 \dots d\vec{r}'_N \delta[\vec{C}(\{\vec{r}'_i\}) - \vec{c}(V)]. \quad (5)$$

Changing the variables of integration,  $\vec{r}'_i = (V/V_0)^{1/3}\vec{r}_i$  and noting relation (1) between the vector  $\vec{C}(\{\vec{r}'_i\})$  and  $C(\{\vec{r}_i\})$  we can write

$$\begin{aligned} \Omega[\vec{c}(V), V] &= \left( \frac{V}{V_0} \right)^N \int_{V_0} d\vec{r}_1 \dots d\vec{r}_N \\ &\times \delta(\mathbf{M}[\vec{C}(\{\vec{r}_i\}) - \vec{c}(V_0)]), \end{aligned} \quad (6)$$

where

$$\vec{c}(V) = \mathbf{M} \cdot \vec{c}(V_0). \quad (7)$$

Using the property of the Dirac  $\delta$  function,  $\delta(\mathbf{M}[\vec{C}(\{\vec{r}_i\}) - \vec{c}(V)]) = \delta(\vec{C}(\{\vec{r}_i\}) - \vec{c}(V))/|\mathbf{M}|$ , where  $|\mathbf{M}|$  is the determinant of the matrix  $\mathbf{M}$ , we see that the densities of states at different volumes are related by

$$\Omega[\vec{c}(V), V] d\vec{c}(V) = \left( \frac{V}{V_0} \right)^N \Omega[\vec{c}(V_0), V_0] d\vec{c}(V_0), \quad (8)$$

with  $\vec{c}(V)$  and  $\vec{c}(V_0)$  related by Eq. (7).

Suppose that we perform several Monte Carlo simulations at inverse temperature  $\beta_i$  and volume  $V_i$ ,  $1 \leq i \leq R$ . Each simulation provides an estimate of the density of states

$$\Omega[\vec{c}(V_i), V_i] d\vec{c}(V_i) \approx \exp(\beta_i \{E[\vec{c}(V_i)] - f_i\}) \frac{h_i[\vec{c}(V_i)]}{M_i}, \quad (9)$$

where  $E[\vec{c}(V_i)]$  is the potential energy of the system,  $h_i[\vec{c}(V_i)]$  is the histogram of the variables  $C_n$  measured in the simulation  $i$ ,  $\delta\vec{c}(V_i)$  is the histogram bin size,  $M_i$  is the number of measures and  $f_i$  is the Helmholtz free energy at inverse temperature  $\beta_i$  and volume  $V_i$ . The values of  $f_i$  are not known by now but will be self-consistently determined later.

We use (8) to relate the density of states at volume  $V$  to the density of states estimated at the simulation volume by the above equation,

$$\begin{aligned} \Omega[\vec{c}(V), V] d\vec{c}(V) &\approx \left( \frac{V}{V_i} \right)^N \\ &\times \exp(\beta_i \{E[\vec{c}(V_i)] - f_i\}) \frac{h_i[\vec{c}(V_i)]}{M_i}. \end{aligned} \quad (10)$$

The estimates of the density of states given by each of the  $R$  simulations are now combined [2],

$$\begin{aligned} \Omega[\vec{c}(V), V] d\vec{c}(V) &= \sum_{i=1,R} p_i \left( \frac{V}{V_i} \right)^N \\ &\times \exp(\beta_i \{E[\vec{c}(V_i)] - f_i\}) \frac{h_i[\vec{c}(V_i)]}{M_i}, \end{aligned} \quad (11)$$

assigning to each of them a weight  $p_i$ . The normalized ( $\sum_{i=1,R} p_i = 1$ ) weights are obtained from the condition of minimization of the statistical uncertainty on the density of states,

$$\delta^2 \Omega[\vec{c}(V), V] = \overline{\Omega^2[\vec{c}(V), V]} - \overline{\Omega[\vec{c}(V), V]}^2. \quad (12)$$

The number of measures in each bin of the histogram is a random variable. Neglecting the correlations between the measures and using the independence of the simulations we have [23],

$$\overline{h_i[\vec{c}(V_i)] h_j[\vec{c}(V_j)]} - \overline{h_i[\vec{c}(V_i)]} \overline{h_j[\vec{c}(V_j)]} \approx \overline{h_i[\vec{c}(V_i)]} \delta_{i,j}. \quad (13)$$

The result for the weights is

$$\begin{aligned} p_i^{-1} &= \exp(\beta_i \{E[\vec{c}(V_i)] - f_i\}) \sum_{l=1}^R \left( \frac{M_l}{M_i} \right) \left( \frac{V_l}{V_i} \right)^N \\ &\times \exp(-\beta_l \{E[\vec{c}(V_l)] - f_l\}). \end{aligned} \quad (14)$$

The partition function at inverse temperature  $\beta$  and volume  $V$  is thus

$$Z(\beta, V) = \sum_{c(V)} \sum_{i=1}^R \frac{h_i[\vec{c}(V_i)] \exp(-\beta E[\vec{c}(V)])}{\sum_{l=1}^R M_l (V_l/V)^N \exp(-\beta_l \{E[\vec{c}(V_l)] - f_l\})}, \quad (15)$$

and the canonical average of any function  $f[\vec{c}(V)]$  is

$$\langle f \rangle = \frac{1}{Z(\beta, V)} \times \sum_{c(V)} \sum_{i=1}^R \frac{f[\vec{c}(V)] h_i[\vec{c}(V_i)] \exp(-\beta E[\vec{c}(V)])}{\sum_{l=1}^R M_l (V_l/V)^N \exp(-\beta_l \{E[\vec{c}(V_l)] - f_l\})}, \quad (16)$$

where  $\sum_{c(V)}$  is a sum over bins in the multidimensional  $\vec{c}$  space.

For the expansion (4), the system pressure,  $P(\beta, V)$ , is obtained directly from  $C_1$ :

$$P(\beta, V) = \frac{N}{\beta V} - \langle C_1 \rangle. \quad (17)$$

It is clear that in the actual calculations there is no need to compute the histograms [23]. Denoting by  $\vec{c}^{i,j}$  the measure  $j$  ( $1 \leq j \leq M_i$ ) in the simulation number  $i$  we have

$$\langle f \rangle = \frac{1}{Z(\beta, V)} \times \sum_{i=1}^R \sum_{j=1}^{M_i} \frac{f[\vec{c}^{i,j}(V)] \exp(-\beta E[\vec{c}^{i,j}(V)])}{\sum_{l=1}^R M_l (V_l/V)^N \exp(-\beta_l \{E[\vec{c}^{i,j}(V_l)] - f_l\})}, \quad (18)$$

where

$$Z(\beta, V) = \sum_{i=1}^R \sum_{j=1}^{M_i} \frac{\exp(-\beta E[\vec{c}^{i,j}(V)])}{\sum_{l=1}^R M_l (V_l/V)^N \exp(-\beta_l \{E[\vec{c}^{i,j}(V_l)] - f_l\})}. \quad (19)$$

One should remark that in the above expression the values  $\vec{c}^{i,j}(V_l)$ , with  $l=i$  are measured while the corresponding coefficients with  $l \neq i$  as well as  $\vec{c}^{i,j}(V)$  are computed from the measured values using Eq. (7). The free energies  $f_i$  are self-consistently obtained from the conditions  $f_i = -\beta_i^{-1} \ln Z(\beta_i, V_i)$  and by setting  $f_1 = 0$ . Thus, we are able to compute free energies relative to some thermodynamic state  $\beta_1, V_1$ .

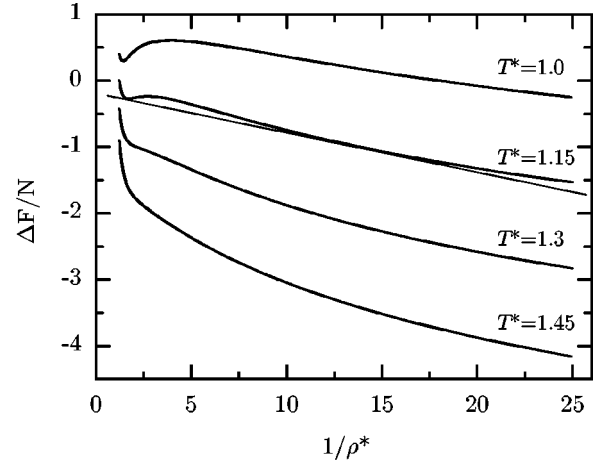


FIG. 1. Relative Helmholtz free energy as a function of volume per particle at four different temperatures  $T^* = 1.0, 1.15, 1.3$ , and  $1.45$ . The results obtained from the two simulation sets at temperatures  $T_1^* = 1.15$  and  $T_2^* = 1.3$  are plotted and they are not distinguishable. We also show the double tangent straight line for  $T^* = 1.15$ .

Application of our method to rigid molecular systems [24] is also possible. The expansion coefficients,  $C_n$  are computed by scaling the center of mass coordinates of the molecules while keeping the intramolecular coordinates fixed.

### III. APPLICATION TO THE LENNARD-JONES FLUID

We measure values of  $\vec{C}$  every 10 MCS/N and the simulation lengths were  $10^5$  MCS/N. The value of the cut-off radius was always equal to half the side of the simulation box. Standard long range corrections were added to the measured values at the end of the simulation.

We first considered the choice (2) for the vector  $\vec{C}$ . Two sets of simulations, with 108 particles, were done at two reduced temperatures,  $T_1^* = 1.15$ , and  $T_2^* = 1.3$ . For the first temperature we made 40 simulations at equally spaced reduced densities:  $0.02 \leq \rho^* \leq 0.8$ . For the second temperature we made simulations at densities  $\rho_i^* = 0.02 \times 1.1^{i-1}$ ,  $1 \leq i \leq 40$ .

Every pair of simulations close in density were combined using the proposed method to obtain results for densities in between the two simulations and for a given range of temperatures (above and below the simulation temperature). From the free energy values obtained we built the volume and temperature dependence of the free energy. In Fig. 1 we show the free energy as a function of volume per particle at four different temperatures  $T^* = 1.0, 1.15, 1.3$ , and  $1.45$ . In this figure we also compare the extrapolations obtained from each of the two sets of simulations made at  $T_1^* = 1.15$  and  $T_2^* = 1.3$ . The curves from these two simulations are nearly coincident and they are not distinguishable in the figure. For  $T^* = 1.15$  we also show the common tangent straight line at the liquid and gas coexisting phases. The double tangent construction allows us to find the volumes of the coexisting phases at each temperature. A new set of simulations with 256 particles at  $T^* = 1.3$  and densities  $\rho_i^* = 0.1 \times 1.047^{i-1}$ ,  $1 \leq i \leq 40$  was also done.

In Fig. 2 we show the phase diagram computed from each set of simulations. In order to ascertain the usefulness of volume expansion method (4) we also made simulations

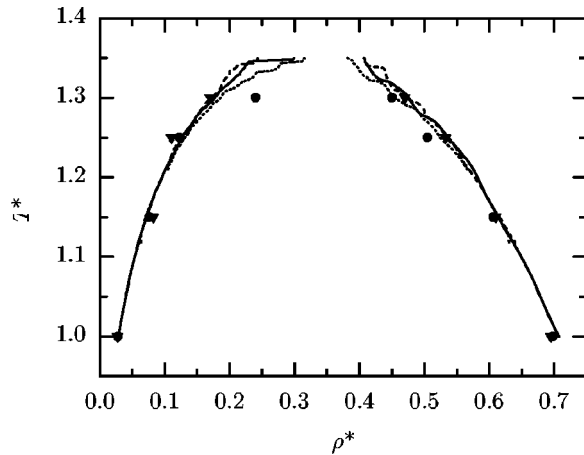


FIG. 2. Liquid-gas phase diagram of the three dimensional Lennard-Jones model. Solid line: Simulation temperature  $T^* = 1.3$  and 108 particles; dashed line:  $T^* = 1.15$  and 108 particles; dotted line:  $T^* = 1.3$  and 256 particles; circles and triangles are Gibbs Ensemble results from reference [12] with 500 particles and 300 particles, respectively.

where the variables,  $C_n$ ,  $0 \leq n \leq 5$ , were measured. This set of simulations was done for a system of 108 particles at  $T^* = 1.3$  and at the same densities chosen before. In Fig. 3 we show the convergence of the results for the relative free energy as a function of volume for a temperature  $T^* = 1.15$  as we include an increasing number of coefficients in the expansion (4). The curves obtained with 4, 5, and 6 coefficients are nearly coincident and agree with results obtained from the choice based on Eq. (2).

#### IV. CONCLUSIONS

We have proposed a method which allows simultaneous extrapolations in volume and temperature based on the multiple histogram method. An arbitrary number of Monte Carlo simulations made in the canonical ensemble can be combined providing improved estimates of thermodynamic prop-

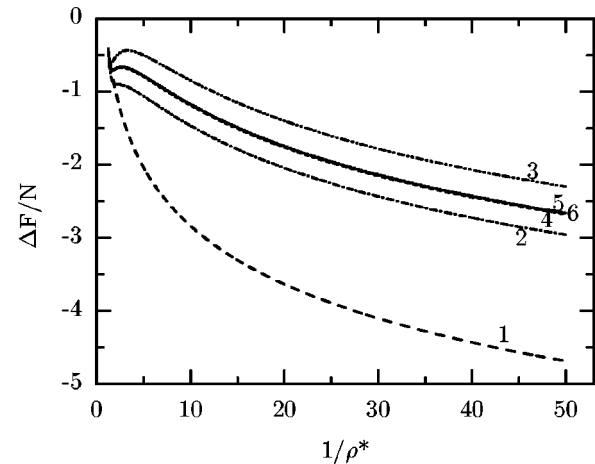


FIG. 3. Convergence of the free energy with an increasing number of coefficients in the expansion based in Eq. (4). Results with 1 to 6 coefficients correspond, respectively, to long-dashed, dot dashed, dot-dot-dashed, short-dashed, dotted, and solid curves. The curves obtained with 4, 5, and 6 coefficients are nearly coincident.

erties. We show test bed results on the three-dimensional Lennard-Jones system which confirm that the method works well and that the volume expansion scheme based on equation (4) can be used with a good control of the approximations involved. Calculation of relative Helmholtz free energy coupled with the double tangent construction allows an efficient determination of the phase diagram. The method is general and can be applied to interaction potentials that do not have a simple scaling with system volume.

#### ACKNOWLEDGMENTS

We thank Professor E. J. S. Lage, Professor S. K. Mendiratta, Professor T. P. Gasche, and Professor J. M. Pacheco for a critical reading of the manuscript. A. L. Ferreira gratefully acknowledges J. M. Pacheco and J. P. Ramalho for fruitful discussions. This work was partially supported by Project Nos. PRAXIS/2/2.1/299/94 and PRAXIS/2/2.2/FIS/302/94.

- [1] A. M. Ferrenberg and R. H. Swendsen, Phys. Rev. Lett. **61**, 2635 (1988).
- [2] A. M. Ferrenberg and R. H. Swendsen, Phys. Rev. Lett. **63**, 1195 (1989).
- [3] R. H. Swendsen, Physica A **194**, 53 (1993).
- [4] E. Marinari, cond-mat/9612010, 1996 (unpublished).
- [5] G. M. Torrie and J. P. Valleau, J. Chem. Phys. **66**, 1402 (1977).
- [6] J. P. Valleau, J. Comput. Phys. **96**, 193 (1991).
- [7] N. V. Brilliantov and J. P. Valleau, J. Comput. Phys. **108**, 1115 (1998).
- [8] B. A. Berg and T. Neuhaus, Phys. Rev. Lett. **68**, 9 (1992).
- [9] W. Janke, in *Computer Simulations in Condensed Matter Physics VII*, edited by D. P. Landau, K. K. Mon, and H. B. Schuttler (Springer-Verlag, Heidelberg, 1994).
- [10] A. P. Lyubartsev *et al.*, J. Chem. Phys. **96**, 1776 (1992).
- [11] F. A. Escobedo and J. de Pablo, Mol. Phys. **87**, 347 (1996).
- [12] A. Z. Panagiotopoulos, Mol. Phys. **61**, 813-826 (1987).
- [13] A. Z. Panagiotopoulos, Molecular Simulations **9**, 1 (1992).
- [14] J. M. Caillol, J. Chem. Phys. **109**, 4885 (1998).
- [15] N. B. Wilding, Phys. Rev. E **52**, 602 (1995).
- [16] A. D. Bruce and N. B. Wilding, Phys. Rev. Lett. **68**, 193 (1992).
- [17] E. J. Meijer *et al.*, J. Chem. Phys. **92**, 7570 (1990).
- [18] D. A. Kofke, J. Chem. Phys. **98**, 4149 (1993).
- [19] M. Mehta and D. A. Kofke, Mol. Phys. **86**, 139 (1995).
- [20] F. A. Escobedo, J. Chem. Phys. **108**, 8761 (1998).
- [21] A. L. Ferreira, J. M. Pacheco, and J. P. Ramalho (unpublished).
- [22] D. Frenkel and B. Smit, *Understanding Molecular Simulation* (Academic Press, San Diego, 1996).
- [23] A. Ferrenberg, D. P. Landau, and R. H. Swendsen, Phys. Rev. E **51**, 5092 (1995).
- [24] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Clarendon Press, Oxford, 1987).