JOURNAL OF COMPUTATIONAL PHYSICS 62, 464–471 (1986)

Some Practical Considerations in the Use of Metropolis and Force-Biased Monte Carlo Simulations on Liquids

SHRI SINGH* AND SAUL GOLDMAN†

Department of Chemistry, The Guelph–Waterloo Centre for Graduate Work in Chemistry, Guelph Campus, Guelph, Ontario, N1G 2W1, Canada

Received August 16, 1984; revised December 6, 1984

A series of four force-biased and two Metropolis Monte Carlo calculations were run on a dense Lennard–Jones fluid just above its melting point. Our purpose was to test some recent theoretical predictions [1] as to the relative speed and stochastic characteristics of these algorithms, and to learn what the most propitious maximum step-size was, in each case. We found that the principal predictions in [1] were upheld for this system, that the Metropolis algorithm was speeded up by using the larger of the two maximum step-sizes tried, that the force-biased algorithm with the full bias was speeded up by using the smaller of the two step-sizes tried, and that the force-biased algorithm, with the bias partially on, was relatively insensitive to the choice of this parameter.

INTRODUCTION

In a previous article [1], a theoretical study was carried out whose purpose was to try to characterize and understand the convergence characteristics that are intrinsic to the Metropolis [2] and the force-biased [3–5] Monte Carlo algorithms (hereafter MMC and FBMC, respectively) when these are applied to liquids. The approach taken was to approximate the real situation by model transition matrices based on Gauss-like histogram distribution functions. Since these transition matrices were defined over a computationally manageable number of states ($O(10^2)$), it was possible to obtain values for their subdominant eigenvalues, which control the convergence rates, and to evaluate the stochastic characteristics of the algorithms. It was found, both from the subdominant eigenvalues and from the stochastic characteristics of the Markov chains generated, that the FBMC algorithm with the biasing strength partially turned on (i.e., with $0 < \lambda < 1$) was best. It was also found that the FBMC algorithm with $\lambda = 1$, i.e., with the biasing fully turned on, tended to procedure an over-occupation of the low-energy wing of the distribution function, at the expense of the high-energy wing, in the early stages

* On leave of absence from: Department of Physics, Banaras Hindu University, Varanasi–221005, India.

† Author to whom correspondence should be addressed.

0021-9991/86 $3.00

of the walk. Of course, in the limit of an infinite number of steps all the algorithms converge to the same limiting distribution function. This skewing behavior was shown to be due to both limited state-to-state accessibility at each step, and to distortion of the bias due to first-order truncation of the energy expansion. As the one-step accessibility was reduced, or as the distortion was increased, the skewing with any $\lambda > 0$ algorithm increased, as did the magnitude of the subdominant eigenvalue, indicating slower overall convergence.

The purpose of this article is to test the principal results and predictions found in [1] against what happens in actual simulations on liquids. Our motivation to do this is twofold. First, the approximations that were required to do the calculations in [1], of course, made tentative the connection of [1] to actual simulations, so we simply wanted to see how well or otherwise the theoretical results [1] fared against reality. Second, a question arose implicitly in [1] that could not there be resolved. The question was, how does the maximum move size, in both MMC and FBMC, affect the convergence rate and stochastic characteristics of the algorithm? For MMC a large maximum step size seemed to be indicated since this would increase the number of states accessible at each step, and this increase was found to have a salutory effect on the convergence rate (see Figs. 6–8 of [1]). But the approximate way in which accessibility was modelled in [1] rendered this conclusion tentative. The question, applied to FBMC, was even less answerable. Here, increasing the maximum step-size increases both accessibility (which is beneficial) and distortion of the biasing function due to first-order truncation (which is detrimental). Since it was impossible to know which effect would dominate, the important practical question of how to choose the maximum step size in FBMC calculations remained unanswered. Thus part of the purpose of this article was to shed some light on this issue.

Calculations

We chose to study the details by which the different algorithms, each run with different maximum step sizes (Δ) brought about melting in a dense Lennard–Jones solid at a temperature just above its melting point. A low temperature was needed since the differences between the algorithms diminish with increasing temperature. The melting transition was chosen since we wanted to monitor the evolution of the limiting potential energy distribution function, with all energy levels except one initially unoccupied. Thus our starting configuration was always an FCC lattice wherein all the particles have the same initial potential energy. Our temperature was $T^* = 0.719$ and density ρ^* was 0.85, since at $\rho^* = 0.85$ a Lennard–Jones solid melts at $T^* \simeq 0.704$ [6]. Here $T^* = kT/\varepsilon$, $\rho^* = \rho\sigma^3$, the Lennard–Jones potential is $u(r) = 4\varepsilon[(\sigma/r)^{12} - (\sigma/r)^6]$, k is the Boltzmann constant and ρ is the number density.

To test the predictions in [1], we monitored (in addition to the usual functions) the evolution of the histogram distribution function shown in Fig. 1, from the

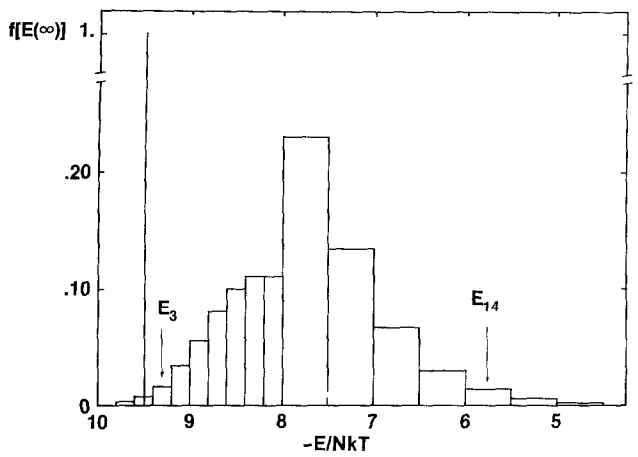

FIG. 1. Sixteen selected energy regions (E_1–E_{16}) from left to right that comprise the normalized limiting histogram distribution function for the Lennard–Jones liquid at $\rho^* = 0.85$, $T^* = 0.719$, $N = 108$. $f[E(\infty)]$ is the fraction of particles within the energy range indicated by the appropriate rectangle. The regions E_3 and E_{14} are used in the text to illustrate the depopulation and population of the low- and high-energy wings of the distribution function, respectively. The solid line at $E/NkT = -9.4905$ represents the potential energy for the initial FCC configuration. Further details are given in Table I.

starting FCC configuration. We also stored system configurations periodically, to check for the onset of melting.

All the calculations were done with 108 particles, the usual periodic boundary conditions, spherical cut off at 2.5136σ and the test particle picked at random. The FBMC calculations were done as in [4, 5]. The calculations were done on an FPS-164 Array Processor that carries 15 significant figures in its arithmetic operations.

RESULTS

We want to compare the different ways by which the limiting histogram distribution function shown in Fig. 1 is filled in by the different algorithms. In particular, to test the results of [1] we focus on the development of the wings of the distribution function. Thus we describe the rate of population of the high energy wing, for which we use region E_{14} (Fig. 1) and the corresponding depopulation of the low-energy wing, for which we use region E_3 (Fig. 1). Our situation of having region E_3 over-occupied and region E_{14} under-occupied in the early stages of the walk is a consequence of starting from an FCC configuration for which the potential energy per particle is considerable lower than the final ensemble average energy. (Here $E(\text{initial})/NkT = -9.49$ versus $\langle E/NkT \rangle = -7.93$, both without a tail correction).

These results are given in Figs. 2 and 3, where we plot the fraction of particles in

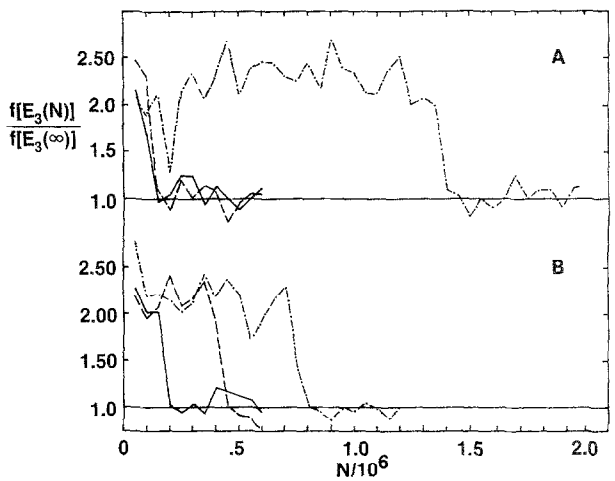

FIG. 2. These graphs trace the progress by which region E_3 of Fig. 1 is depopulated in the course of each of the six walks that were generated, starting in each from an FCC configuration. (—) FBMC $\lambda = \frac{1}{2}$, (– –) FBMC $\lambda = 1$, (– · –) MMC. (A) is for $\Delta = 0.10\sigma$; (B) is for $\Delta = 0.15\sigma$, both here and in Figs. 3 and 4.

the low- and high-energy regions, respectively, over the course of the walk. Each fraction is normalized with respect to its infinite time value (Table I), so that achievement of the equilibrium distribution function is indicated by the plotted function oscillating about unity. It is clear from these graphs that

(a) the MMC algorithm is the slowest to reach equilibration, but its performance improves with increasing maximum step-size,

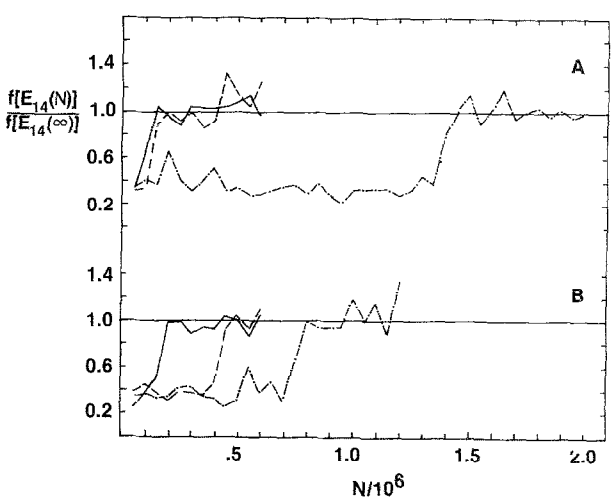

FIG. 3. These graphs trace the progress by which region E_{14} of Fig. 1 is populated, starting from the FCC configuration. Legend and distinction between (A) and (B) as in Fig. 2.

(b) the FBMC $\lambda = \frac{1}{2}$ algorithm is the fastest to come to equilibrium; its performance is affected slightly and negatively by use of the larger maximum step-size,

(c) the FBMC $\lambda = 1$ algorithm is almost as fast as the $\lambda = \frac{1}{2}$ algorithm for the smaller maximum step-size but is significantly slower for the larger maximum step-size.

The above results for equilibrating the wings of the energy distribution function are completely consistent with those for the onset of melting shown in Fig. 4. Here we plot the evolution of the function [6],

$$S = (S_x + S_y + S_z)/3, \tag{1}$$

where

$$S_x = \sum_{i=1}^{N} \cos \frac{4\pi X_i}{a} \tag{2}$$

with similar expressions for S_y and S_z. In Eq. (2) X_i is the absolute X coordinate of particle "i," and "a" is the side of the unit cell in the initial FCC lattice. S_x equals N

TABLE I

Asymptotic Values for the Fractions of Particles Found in the Energy Ranges E_1–E_{16} of Fig. 1

LABEL	$-E/NkT$	$f\{E(\infty)\}^a$
E_1	9.8–9.6	0.0023
E_2	9.6–9.4	0.0069
E_3	9.4–9.2	0.0168
E_4	9.2–9.0	0.0334
E_5	9.0–8.8	0.0559
E_6	8.8–8.6	0.0799
E_7	8.6–8.4	0.0997
E_8	8.4–8.2	0.1105
E_9	8.2–8.0	0.1103
E_{10}	8.0–7.5	0.2282
E_{11}	7.5–7.0	0.1347
E_{12}	7.0–6.5	0.0668
E_{13}	6.5–6.0	0.0306
E_{14}	6.0–5.5	0.0133
E_{15}	5.5–5.0	0.0058
E_{16}	5.0–4.5	0.0025

[a]These were obtained by averaging over all six Monte Carlo walks reported in the text, after the onset of melting in each. The total number of trial moves on which these entries are based was 1.805×10^6. The standard deviation for regions E_3–E_{14} was in the range 0.001–0.002; for the regions E_1, E_2, E_{15}, E_{16} it was ~ 0.0005.

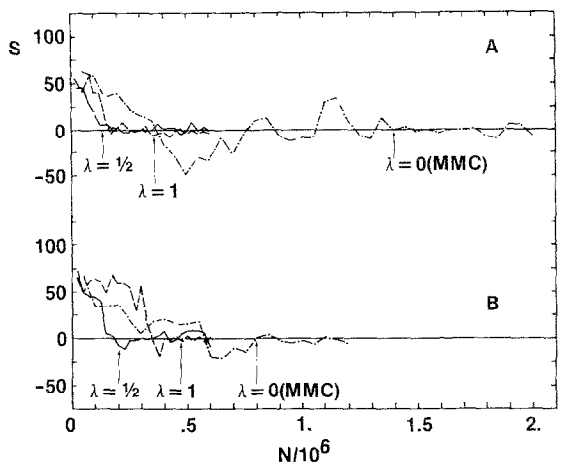

FIG. 4. Development of the function S (Eq. (1)) with the different algorithms. Legend and distinction between (A and B) as in Fig. 2. The arrows, determined from the results in Figs. 2 and 3, are seen here to coincide with the onset of the liquid (see text).

exactly in an appropriately oriented FCC lattice, it is of the order of N in any (appropriately oriented) lattice, and it oscillates around zero with an amplitude of order $N^{1/2}$ for a liquid [6]. A comparison of Fig. 4 with Figs. 2 and 3 reveals that in every walk the onset of melting (deduced by S oscillating around zero within a range of ± 10) coincided essentially exactly with the point at which the wings of the energy distribution reached their equilibrium values. This was taken to occur when

TABLE II

Ensemble Averages for the Internal Energy and Cumulative Constant Volume Molar Heat Capacities of the Lennard–Jones Liquid

Algorithm[a]	$-\langle E/NkT\rangle$[b]	Cv(cal/mol K)[b]	No. trial moves $\times 10^{-3}$[c]
MMC $\Delta = 0.10$	7.93	5.3	600
MMC $\Delta = 0.15$	7.91	5.1	450
FBMC $\lambda = 0.5$, $\Delta = 0.10$	7.94	5.3	500
FBMC $\lambda = 0.5$, $\Delta = 0.15$	7.94	5.1	425
FBMC $\lambda = 1.$, $\Delta = 0.10$	7.93	5.7	475
FBMC $\lambda = 1.$, $\Delta = 0.15$	7.91	5.0	200

[a] See text for meaning of symbols.

[b] These are for the liquid only, with no tail correction. For comparison, molecular dynamics values [8, 9] are $\langle E/NkT\rangle = -7.96 \pm 0.14$, $Cv = 5.3$ cal/mol K. This Cv value was estimated from the temperature-dependence of the configurational energy, obtained with the energy equation and fitted molecular dynamics radial distribution functions [9]. The Cv's reported in [7] for these state conditions are now known to be too high, partly because of round-off error (due to single precision arithmetic [7]) and partly because of incomplete equilibration after liquefaction in [7].

[c] This is the number of single-particle trial moves made subsequent to melting, in each walk.

the plotted functions in Figs. 2 and 3 started to oscillate around a value of 1. The arrows in Fig. 4 refer to these points in Figs. 2 and 3.

The ensemble averages for the energy, heat capacity, and radial distribution functions also showed shifts with the formation of the liquid, but these are not shown. Their response to the onset of the liquid was not nearly as sharp as those of the functions shown in Figs. 2 and 3. The ensemble averages for the energies and heat capacities for the liquid are summarized in Table II. The heat capacities were obtained as described previously [7].

It is worth noting that the rate of filling of much of the central part of the equilibrium energy distribution function eg regions E_8 to E_{10} in Fig. 1 showed no noticeable trend either with algorithm or with maximum step-size. This central region reached its equilibrium value rapidly and at roughly the same rate in all six walks. Thus, for the system at hand at least, the fastest algorithms were characterized by their ability to rapidly populate and depopulate the wings rather than the central part of the energy distribution function.

Discussion

It is gratifying to be able to report that those predictions in reference [1] that were testable, were borne out by the foregoing results. In particular, as predicted, we found that

(a) the FBMC $\lambda = \frac{1}{2}$ algorithm was faster to provide the equilibrium distribution function than either the MMC or the FBMC $\lambda = 1$ algorithms,

(b) the slowness of the FBMC $\lambda = 1$, relative to the $\lambda = \frac{1}{2}$ algorithm, was connected with the slowness of the $\lambda = 1$ algorithm to, respectively, depopulate and populate the low- and high-energy wings of the energy distribution function in the walk toward equilibrium,

(c) the relative order by which the high-energy wing is populated, starting from a state in the low-energy wing was: FBMC $\lambda = \frac{1}{2} >$ FBMC $\lambda = 1 >$ MMC. This was precisely the order found in our model calculations (Fig. 9b, [1]).

We are also now able to address the question mentioned in the Introduction regarding the most propitious choice of the maximum step-size in the different algorithms. For the system studied here, the MMC algorithm was significantly speeded up in going from a maximum step-size of 0.1σ for which the mean acceptance rate was ~ 0.42, to a maximum step-size of 0.15σ and a mean acceptance rate of ~ 0.25. So for MMC, within these ranges and for this system at least, it pays to increase the one-step accessibility, despite the resultant lower acceptance rate. Just the reverse turns out to be true for the FBMC $\lambda = 1$ algorithm. Here the speed of equilibration was made significantly faster by reducing the maximum step-size from 0.15σ, where the mean acceptance rate is ~ 0.35, to 0.10σ, where the mean acceptance rate is ~ 0.70. As discussed below, this result is noteworthy in two respects.

First, it is partially at variance with an earlier suggestion [4], that the maximum step-size for *any* algorithm be selected so as to maximize the root-mean-squared (rms) displacement per move. For liquids at relatively high densities this latter criterion leads to a choice of larger rather than smaller maximum step-sizes. So while the root-mean-squared criterion is in accord with our MMC results, it clearly conflicts with our (high density ie $\rho^* = 0.85$) FBMC $\lambda = 1$ results. We point out, however, that at lower densities (e.g., $\rho^* = 0.77$) the rms displacement criterion, does not lead to large maximum displacements for $\lambda = 1$ [5].

Second, the result clears up, at least for this system, which of the two competing effects on the convergence rate: an increase in accessibility or a decrease in the distortion of the biasing function is the more important. We now know it is reduction of the distortion that is here more important when $\lambda = 1$. This also turned out be true for the $\lambda = \frac{1}{2}$ algorithm, but for $\lambda = \frac{1}{2}$ the effect of using different maximum step-sizes was much less than with $\lambda = 1$ (cf., e.g., the FBMC $\lambda = \frac{1}{2}$ versus $\lambda = 1$ tracings in Figs. 2A versus 2B, 3A versus 3B, and 4A versus 4B). So a final conclusion that can be drawn is that the choice of the maximum move size with the FBMC $\lambda = \frac{1}{2}$ algorithm is relatively unimportant, at least for choices that give reasonable acceptance rates.

ACKNOWLEDGMENT

We are grateful to the Natural Sciences and Engineering Research Council of Canada for financial support.

REFERENCES

1. S. GOLDMAN, *J. Comput. Phys.* (1984), in press.
2. N. METROPOLIS, A. W. ROSENBLUTH, M. N. ROSENBLUTH, A. H. TELLER, AND E. TELLER, *J. Chem. Phys.* **21** (1953), 1087.
3. C. PANGALI, M. RAO, AND B. J. BERNE, *Chem. Phys. Lett.* **55** (1978), 413.
4. M. RAO, C. PANGALI, AND B. J. BERNE, *Mol. Phys.* **37** (1979), 1773.
5. M. RAO AND B. J. BERNE, *J. Chem. Phys.* **71** (1979), 129.
6. L. VERLET, *Phys. Rev.* **159** (1967), 98.
7. S. GOLDMAN, *J. Chem. Phys.* **79** (1983), 3938.
8. J. J. NICOLAS, K. E. GUBBINS, W. B. STREETT, AND D. J. TILDESLEY, *Mol. Phys.* **37** (1979), 1429.
9. S. GOLDMAN, *J. Phys. Chem.* **83** (1979), 3033.