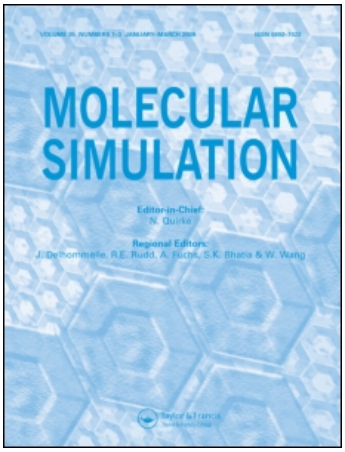

Molecular Simulation
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713644482

A New Equilibration Algorithm for Monte Carlo Simulations

E. Yurtsever[a]
[a] Technische Hochschule Darmstadt, Physikalische Chemie I, Darmstadt, F.R. Germany

To cite this Article Yurtsever, E.(1989) 'A New Equilibration Algorithm for Monte Carlo Simulations', Molecular Simulation, 2: 3, 189 — 199
To link to this Article: DOI: 10.1080/08927028908031367
URL: http://dx.doi.org/10.1080/08927028908031367

Molecular Simulation, 1989, Vol. 2, pp. 189–199
Reprints available directly from the publisher
Photocopying permitted by license only

A NEW EQUILIBRATION ALGORITHM FOR MONTE CARLO SIMULATIONS

E. YURTSEVER

Technische Hochschule Darmstadt, Physikalische Chemie I, Petersen str. 20, D-6100 Darmstadt, F.R. Germany*

(*Received July, 1988*)

An algorithm for speeding up equilibration processes in Monte Carlo simulations is proposed. By employing different maximum-allowed-displacements for each particle and updating them according to the previous acceptance ratios, it is shown that considerably faster convergence rates to equilibrium states are obtained.

KEY WORDS: Monte Carlo, equilibration, Lennard-Jones, maximum allowed displacement

INTRODUCTION

Molecular simulation is a very powerful tool for the theoretical analysis of condensed matter. These "numerical experiments," even though limited by their finite-size, finite-time averaging processes, provide robust methods of obtaining macroscopic information from fundamental interactions between particles. They allow quantitative tests of interaction potentials derived from other theoretical means such as ab-initio quantum calculations.

Molecular Dynamics techniques are based on time averaging of the clasical trajectories of interacting particles whereas Monte Carlo algorithms rely on ensemble averages to extract the necessary information. "Although in its actual realization the Monte Carlo method can be considered to also perform a "time" averaging (1)." In both approaches the initial configurations have to be chosen in an arbitrary fashion if no ad hoc knowledge of the system is available. A random configuration may be a suitable choice for simulating seeking out structural information since it precludes bias. In order to obtain reliable averages, the system is brought close to equilibrium; normally a slow process. In this work we outline a Monte Carlo algorithm which speeds up the equilibration and which can also be used to recognize the "equilibrium" state of the system.

THEORY

The Monte Carlo method is based on random walks in configurational space such that the configurations sampled are distributed according to the Boltzmann distribution (2).

*Address after March 1989: Orta Doğu Teknik Üniversitesi, Kimya Bölümü, Ankara, Turkey

$$P(R) = Z^{-1} \cdot \exp[-\beta V(R)] \tag{1}$$

where $P(R)$ and $V(R)$ are the probability distribution and potential energy functions for the configuration $R = (r_1, r_2, \ldots, r_n)$, r_i is the position vector of the i-th particle. z is configurational partition function and $\beta = (kT)^{-1}$. Following the notation of Rao *et al.* (3) we proceed to define the transition probability distribution $W(R'|R)$ of reaching R' from R as:

$$W(R'|R) = A(R'|R) \cdot T(R'|R) \tag{2}$$

with $T(R'|R)$ a conditional probability function of choosing R' (which generates the random walk in the configurational space) and $A(R'|R)$ the acceptance probability of the specified move. A and T should satisfy two general conditions in order that all possible R' configurations are accessible from R and there exists an equal number of systems which make the transitions from R to R' and from R' to R. Namely:

$$\int dR' \cdot W(R'|R) = \int dR' \cdot T(R'|R) = 1 \tag{3-a}$$

$$W(R'|R) \cdot \exp[-\beta \cdot V(r)] = W(R|R') \cdot \exp[-\beta \cdot V(R')] \tag{3-b}$$

$A(R'|R)$ is usually defined as:

$$q(R'|R) = \frac{T(R|R') \cdot P(R')}{T(R'|R) \cdot P(R)} \tag{4-a}$$

$$A(R'|R) = \text{Min}[1, q(R'|R)] \tag{4-b}$$

The differences between the standard Metropolis Monte Carlo (*MMC*) and the modified algorithms are in the choice of the sampling function $T(R'|R)$. In *MMC* this function is chosen to be uniform in a preselected domain:

$$T(R'|R) = \begin{cases} \text{constant} & R' \in D(R) \\ 0 & R' \notin D(R) \end{cases} \tag{5}$$

The normalization integral of $T(R'|R)$ is therefore a constant and (4-b) is reduced to:

$$A(R'|R) = \text{Min}\left[1, \frac{P(R')}{P(R)}\right] \tag{6}$$

A more efficient sampling procedure which is based on the forces acting on particles (force bias FBMC) has also been proposed (4) and used to study dilute solutions (5). In this technique $T(R'|R)$ is defined as:

$$T(R'|R) = \begin{cases} Q^{-1} \cdot \exp\{\lambda\beta \cdot F(R) \cdot (R' - R)\} & R' \in D(R) \\ 0 & R' \notin D(R) \end{cases} \tag{7}$$

where Q is the normalization, λ a variable parameter and $F(R)$ is the vector of the forces on all particles. Now the acceptance probability $A(R'|R)$ of equation (4-b) can no longer be simplified to the form of equation 6. FBMC is a powerful technique. However it requires the computation of the forces in each step in addition to the potential energy as well as the more complicated functional forms of $T(R'|R)$ and $A(R'|R)$.

SCALED-MAXIMUM-ALLOWED DISPLACEMENT METHOD

A new configuration on a Markov chain corresponds to an N-particle move from the previous one. In practice this is achieved by supplying $m \cdot N$ one-particle moves where m is a small integer. Then the domain $D(R)$ which is used in equations 5 and 7 is the maximum allowed displacement (MAD) of a given particle. MMC and FBMC assume equal MAD's for each particle and they are kept constant during the simulation. The role of the magnitude of MAD on the convergence rate is very clear. A MAD which is too large will result in the rejection of many moves; on the other hand too small a MAD will result in a very slow convergence. It is also obvious that all the particles should not have the same MAD. This may be a more pronounced effect for clusters and the dilute solutions, in which the different regions of the sample may have different physical characteristics. One attempt to incorporate variable MAD into an MC algorithm is due to Goldman (6) who suggested modifying the MAD of each particle by comparing its potential energy to the average energy (ESDMC). His method was later modified to include detailed energy balance as given in equation 3-b (7).

Here we describe an algorithm which is similar to ESDMC in that it uses differently scaled MADs for each particle. The scaling is based on the acceptance ratio of the preceeding cycle during the walk on the Markov chain. It can be applied to both MMC and FBMC and is shown to diminish the equilibration times considerably.

For a system far from equilibrium one has to use large a MAD to reach configurations which lie closer to the equilibrium. However, then the rejection rate would also be high, since a large number of unstable structures would be tested. This is analogous to the well known problem of numerical optimization algorithms. With large step sizes on one has difficulties to converge and with small steps it converges and with small steps it converges very slowly and possibly to some local minimum. In order to optimize the magnitude of MAD we proceed in the following way. If the acceptance ratio is larger than a predetermined value we allow the particles to step out more by increasing the MAD. Conversely, if the acceptance ratio is smaller than the mentioned threshold the domain acessible to an individual particle is contracted. The scaling of each MAD is restricted so that they cannot become larger than 1/8-th of the box in a periodic system (along each coordinate only half of the length is allowed since minimum distance conventions are used). On the other hand, accidentally low acceptance ratios can be obtained when only a small number of trial moves for an individual particle are tested. Consequently the MAD for such a particle may decrease too quickly, thus slowing the convergence process. A recipe to correct this unwanted reductions is to enforce an maximum reduction of MADs at every MC cycle. The scaled MAD (SMAD) is summarized as:

$$\Delta^i = \Delta^i_x \cdot \Delta^i_y \cdot \Delta^i_z \tag{8-a}$$

$$\Delta^{i+1}_x = \Delta^i_x \cdot c \text{ with } \Delta^{i+1}_x \leqslant X/2 \tag{8-b}$$

$$c = \text{Max}\left[b, \left[\frac{H/T}{a}\right]^{1/3}\right] \tag{8-c}$$

where Δ^i_x is the x component of MAD in the i-th cycle, X is the length of the periodic box along the x-axis, H/T is the acceptance ratio during a cycle of $m \cdot N$ moves, a and b are variable constants. Here all three components of MAD are scaled uniformly but

the method can be extended so that they are scaled individually provided that each move is composed of a single particle move along a single coordinate.

RESULTS AND DISCUSSION

SMAD is tested on two different Lennard-Jones systems. The first one (I) is a cluster composed of 108 Argon particles at $T = 30\,K$ in a cubic box of dimensions $64 \times 64 \times 64\,Å$. The corresponding reduced particle number density is $\varrho^* = 0.016264$ (defined as $N/V^* \sigma^3$ where σ is the Lennard-Jones diameter). The other relevant parameters are given in Table I. This model is almost equivalent to a geometry optimization problem as the low density and temperature would surely bring the system close to a minimum energy configuration. The initial configuration is chosen in a random fashion and 1500 Monte Carlo cycles are carried out both for MMC and FBMC. Each cycle is $2 \times 108 = 216$ trial moves and at the end of cycles, configurational energies and hit/try ratios for each particle are computed. To study the effect of the magnitude of MAD on the convergence properties, a series of simulations is carried out employing various initial Δ^0 values. In our simulations one-particle moves have three random cartesian components. The acceptance or rejection of a move solely depends on the final position of the particle, not the individual cartesian components. Therefore it is not possible to analyze the fate of the trial move in terms of cartesian displacements. An alternative approach of defining each trial move along a single coordinate would probably be inefficient as a large number of energy and the random number calculations are required. Therefore Δ_x^0, Δ_y^0 and Δ_z^0 are taken to be equal and scaled identically for SMAD simulations. For the constant MAD, $\Delta_x^0 = 0.15, 0.25, 0.50, 1.0, 2.0$ and $4.0\,Å$ are selected and the average energy per particle (averaged over 10 cycles) as a function of the number of MC cycles is given in Figure 1, both for MMC and FBMC. Because of the low density and the random initial distribution, one would expect the large step size simulations to converge faster. Figure 1 confirms this expectation and it is clear that by employing small MAD, it may not even be possible to come close to an equilibrium state in a reasonable amount of time. For this problem FBMC does not give substantially better results than MMC but it is known that the force-bias techniques perform better close to equilibrium. The convergence rates are almost directly proportional to the size of MAD and only at significantly high step sizes equilibration seems to take place at all.

Table I Computational parameters of Lennard-Jones Argon systems.

	SYSTEM I	*System II*
Number of particles	108	256
Temperature	30	86.4956
Length of the box	64.0	22.99342
Reduced number density	0.01626	0.85
m^*	2	1

Lennard-Jones parameters:
$\varepsilon = 0.2380426$ kcal/mol
$\sigma = 3.405\,Å$
γ parameter of the FBMC is 0.5.
For SMAD: a = 0.75, b = 0.95
The number of trial moves in a Monte Carlo cycle is m number of particles.

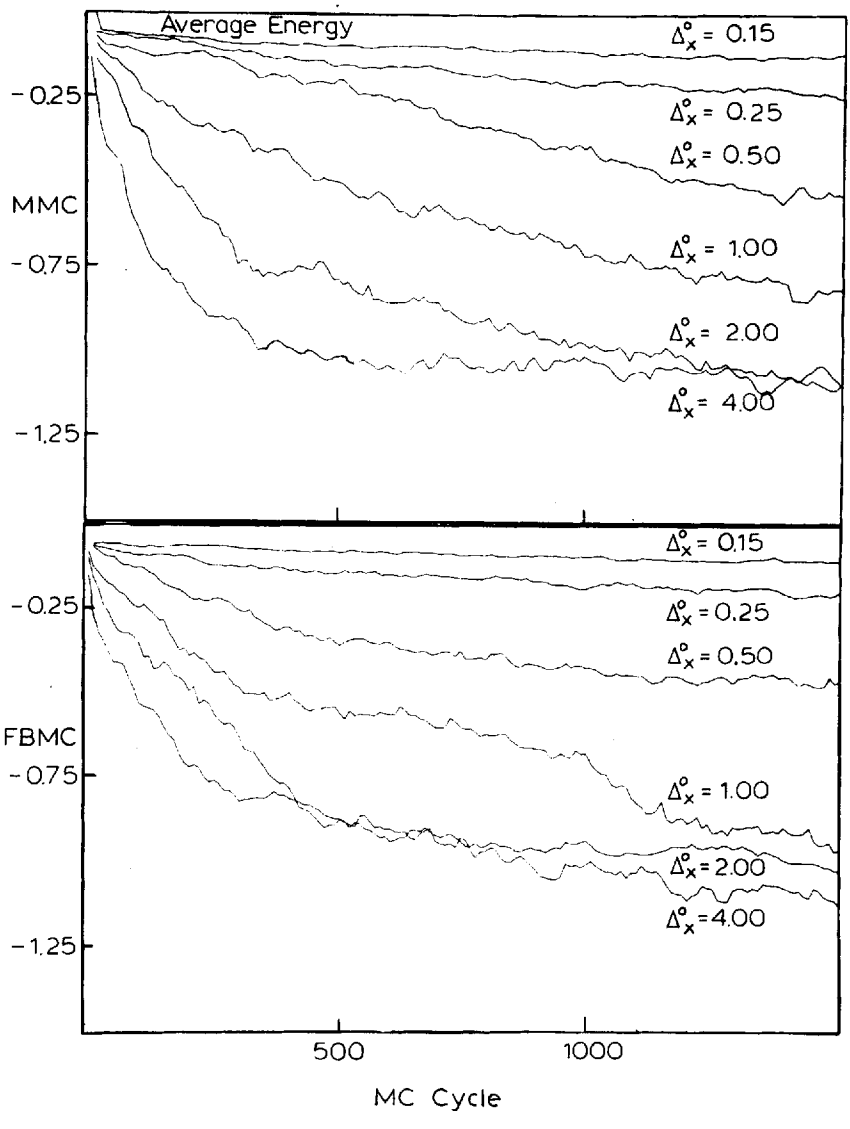

Figure 1 Average energy of a particle in LJ cluster.

However the price to pay is the high rejection rate of the trial moves. In Figure 2 the average HIT/TRY ratios are plotted again as functions of the number of MC cycles. In those cases where the equilibrium state is attainable, the rejection rates as high as 95% are observed. The proposed method of SMAD scaling finds a compromise between high rejection and low convergence rates.

The parameters of Equation 8c are chosen as $a = 0.75$ and $b = 0.95$. All particles with HIT/TRY ratios of more than 75% within a cycle will have a larger domain of

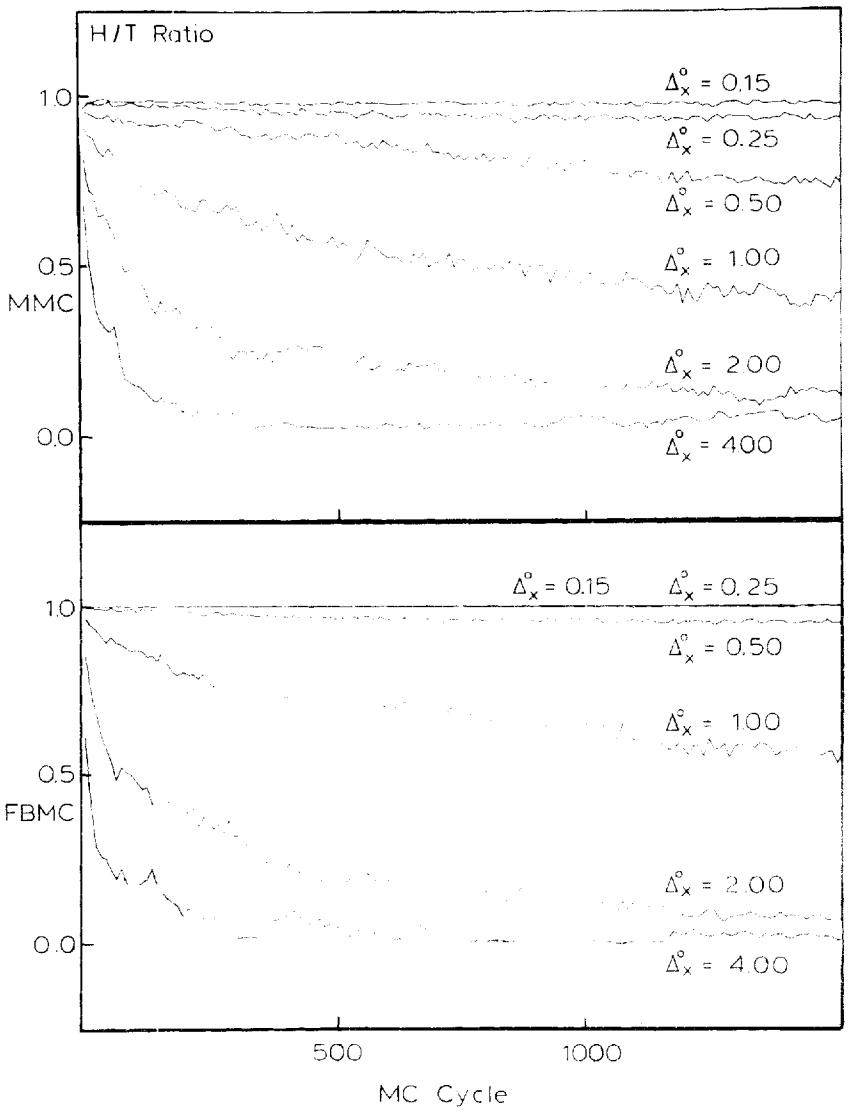

Figure 2 Average HIT/TRY ratio of a particle in LJ cluster.

motion during the next cycle as this high ratio implies a larger energy gradient. In contrast if the acceptance ratio is less than 75% then smaller moves are perscribed, however this contraction of the domain cannot be more than 0.95*0.95*0.95 of the previous one. Figures 3 and 4 show average energy and HIT/TRY ratios for SMAD, respectively. Now the effect of the initial choice of the step size diminishes as the simulation proceeds. The spread between average energies is in the order of 20% which is comparable to fluctuations of an equilibrium state for this small size. Also the advantage of the force-bias technique becomes more apparent. The average

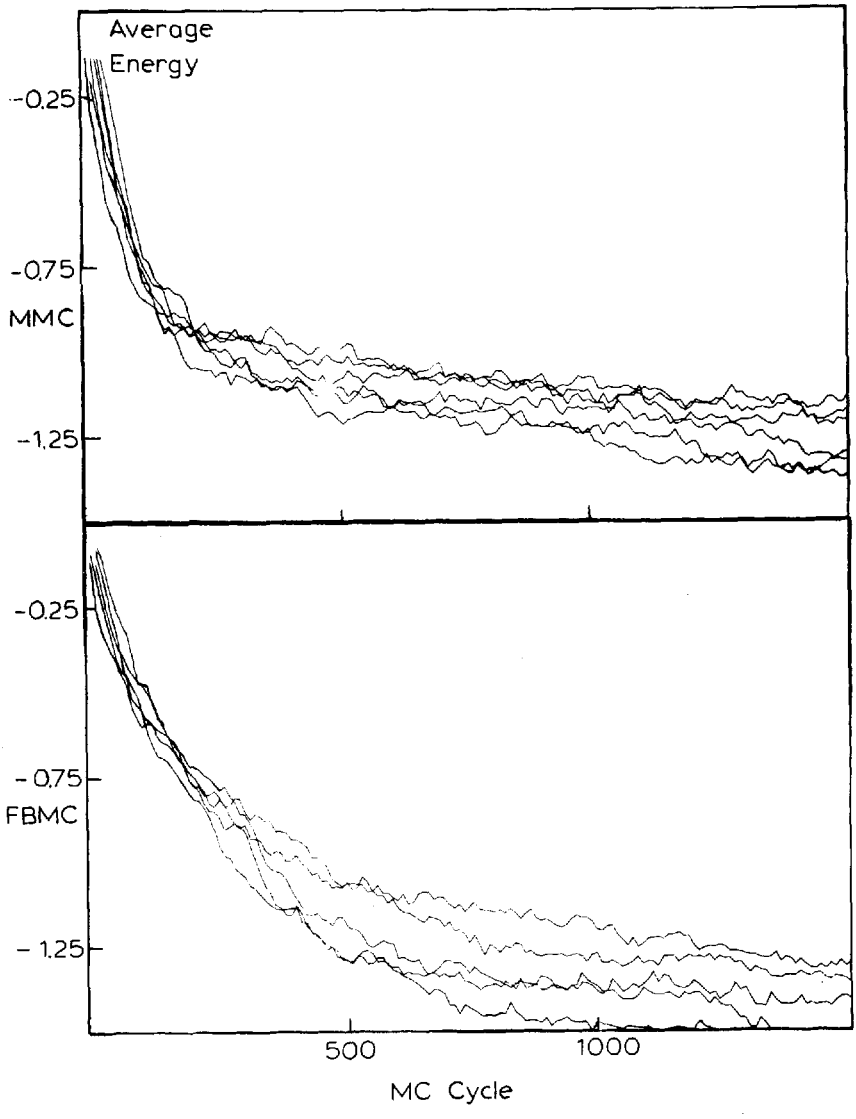

Figure 3 Average energy of a particle in LJ cluster with scaled MAD.

HIT/TRY ratio quickly goes to 25% for both MMC and FBMC. Of course these numerical results can be manipulated by modifying the parameters a and b. From a qualitative analysis of the results it is clear that by employing SMAD the step-size optimization process is efficiently automatized. One can always start with a large MAD to find some of the deep minima in the configurational space and then let the scaling take over for the rest. It also seems that the size of MAD vs MC cycle plots give a more descriptive picture for equilibration as they tend to form a flat plateau which signals convergence. From Figure 4 one sees that after 600 MC cycles the

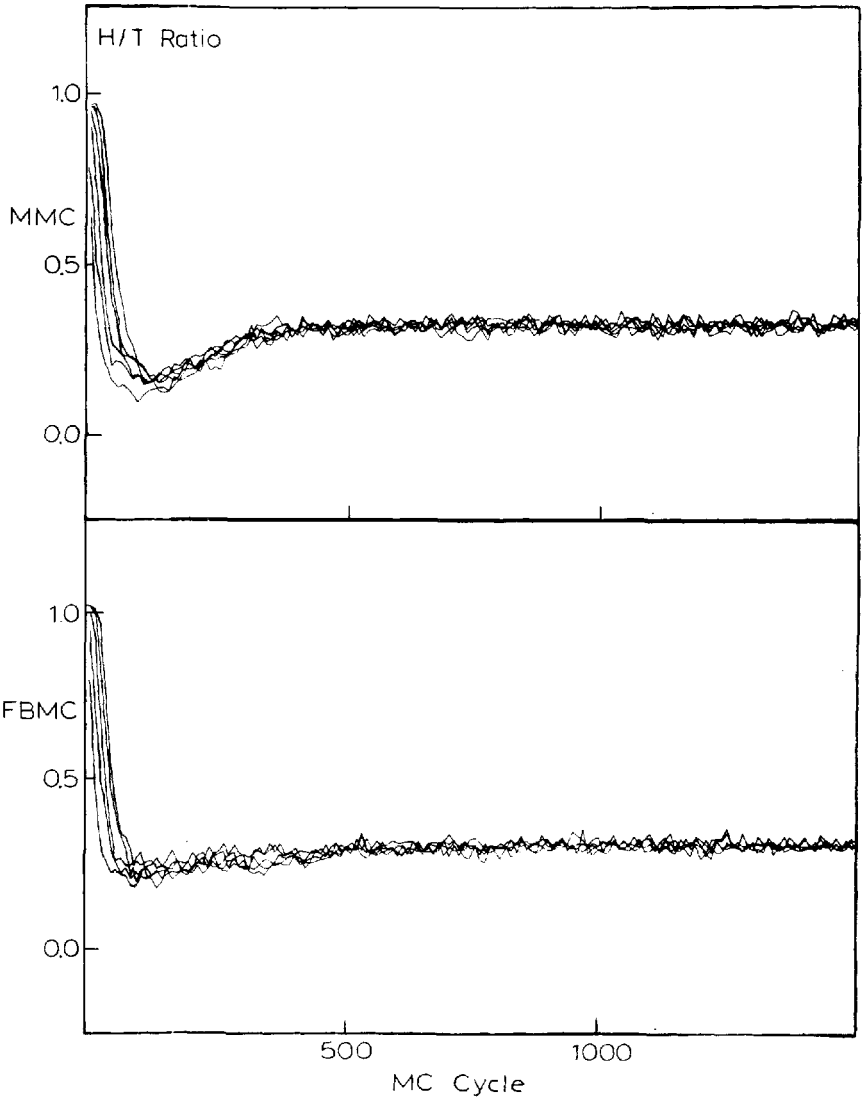

Figure 4 Average HIT/TRY ratio of a particle in LJ cluster with scaled MAD.

average acceptance ratio is a flat function and the computations of the physical properties can be started from there on.

The second test case is an Argon fluid of 256 particles at T = 86.5 K. Since the motion of the particles is restricted by the high density there are a large number of energetically close configurations available and faster convergence rates are obtained. In Figure 5 average energy plots are given for MMC, FBMC and their SMAD counterparts for the initial step size of $\Delta_x^0 = 1.0$ Å. The constant MMC gives the "worst" results and FBMC has a definite advantage over MMC. However upon

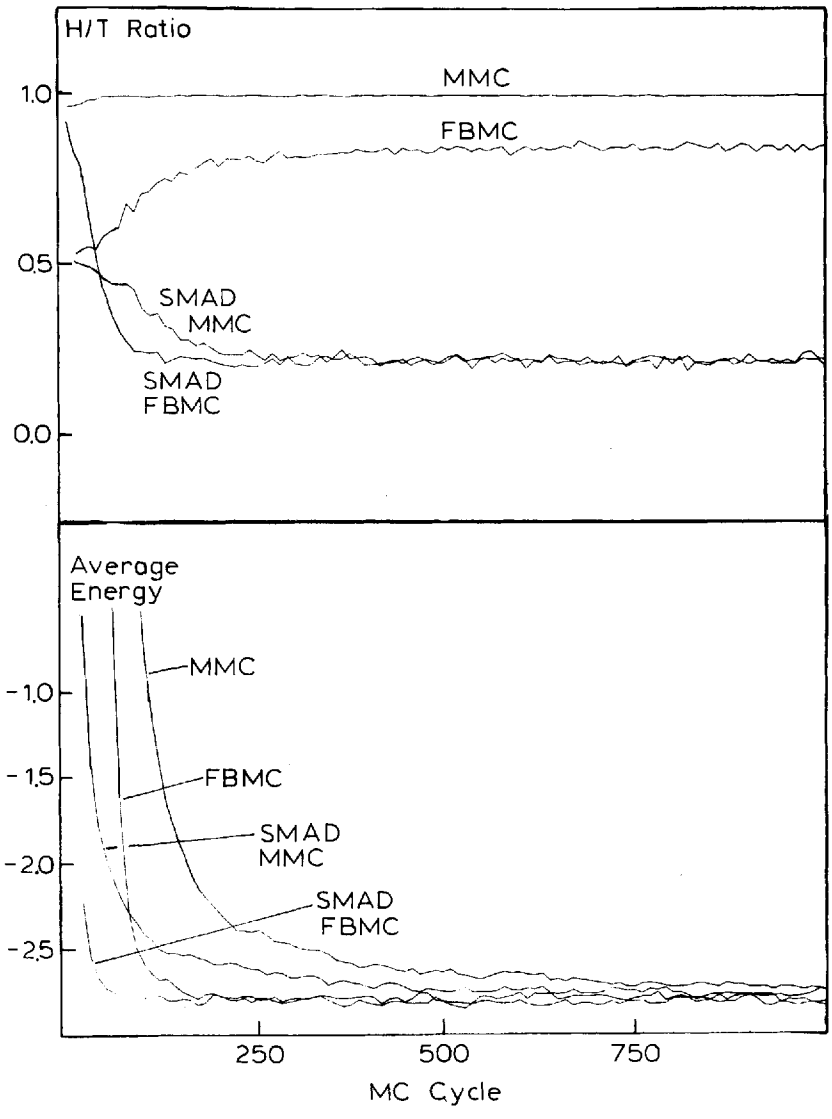

Figure 5 Comparison of average energies and HIT/TRY ratios of the standard and scaled MAD methods for dense LJ fluid.

applying SMAD this difference is less significant and better rates of convergences are obtained. The quantitative measures of the improvements over the constant MAD methods would surely depend on the system chosen and the parameters. However the computational effort to carry out the scaling is minimal, consisting of a single integer addition for accepted moves and a small loop to modify MADs at the end of every cycle. The generalization of SMAD to nonspherical molecules where translational and rotational moves are allowed should be relatively simple keeping in mind that acceptance ratios of both these types of moves may have different significance.

Finally we should like to show that the detailed energy balance is not really disturbed during the scalings. The conditional probability function of choosing a new configuration and the acceptance probability of this configuration has to be balanced so that the old configuration should be equally accessible from the new one. In SMAD the sampling function is a constant namely 1/MAD as in MMC. As long as MAD remains constant the acceptance probability is simply the ratio of Boltzmann distributions. Therefore for all the moves within a cycle the detailed energy balance is satisfied rigorously. When MAD of a particle is changed at the end of the cycle, there are two possibilities. The new MAD may be so small that the old configuration is no longer within the defined domain. Then there is no possibility of reaching the old configuration again. This effect can be alleviated by selecting b of Equation 8c around 0.85 to 0.95 so that no large reductions of the domain are allowed. The other case is that the new MAD is of sufficient size so that the old configuration is still in the allowed domain. In this case the probability of going back is only slightly altered. For both of these cases there seems to be no analytical corrections to be made. Modifying the transition and acceptance probabilities requires an *a priori* knowledge of the complete configurational space. However our observations point out that the changes in MADs are very small near the equilibrium and we can safely assume that the detailed enery balance is numerically satisfied for the regime in which ensemble averages are computed.

A comparsion between the results of ESDMC (6, 7) and this method shows several trends. First of all the convergence rates depend on the variable parameters A of Equation 1 of Reference 6 and a and b of Equation 8c. Computations with several sets of parameters and initial Δ_x^0 values indicate that both of the methods go to equilibrium with approximately same speeds provided that reasonable parameters are chosen. It is clear that SMAD is less sensitive to the selection of parameters since it updates MADs linearly where in ESDMC exponential changes are introduced. The overall acceptance rates are somewhat higher in case of ESDMC. One significant though expected difference is in the updating procedure. ESDMC performs the best when new MADs are obtained from the initial values but in SMAD the previous set has to be used. The main reason in this difference is again the use of exponential and linear transformations respectively. For the systems studied the best results are obtained with SMAD and small initial Δ_x^0 however this observation is not conclusive enough to state that SMAD is faster than ESDMC. The choice of the scaling would probably depend on the problem to be studied.

We also note that with the automatic optimization of MADs Monte Carlo methods which employ many-particle moves may be practical. Chapman and Quirke (8) have pointed out that contrary to common belief, such moves are permissible for reasonable choices of MAD. Since SMAD should be able to find the "reasonable" range, it is feasible to write faster algorithms which are based on many-particle moves with automatic scalings.

In conclusion we like to state that a scaling of the maximum allowed displacement of each particle according to its acceptance ratio seems to be a very economical and efficient way of reaching the equilibrium conditions.

The author would like to thank J. Brickmann and P. Bopp of THD for various discussions and valuable suggestions. This project is supported by the BMFT grant 03 8678 0.

References

[1] K. Binder, Monte Carlo Methods in Statistical Physics, Springer Verlag, Berlin 1, (1979).
[2] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller and E.Teller, "Equation of state calculation by fast computing machines," *J. Chem. Phys.*, **21,** 1087 (1953).
[3] M.Rao, C. Pangali and B.J. Berne, "On the force bias Monte Carlo simulation of water: methodolgy, optimization and comparsion with molecular dynamics," *Mol. Phys.*, **37,** 1773 (1979).
[4] C. Pangali, M. Rao and B.J. Berne, "On a novel Monte Carlo scheme for simulating water and aqueous solutions," *Chem. Phys. Lett.*, **55,** 413 (1978).
[5] P.K. Mehrotra, M. Mezei and D.L. Beveridge," Convergence accelaration in Monte Carlo computer simulation on water and aqueous solutions," *J. Chem. Phys.*, **78,** 3156 (1983).
[6] S. Goldman, "A simple new way to help speed up Monte Carlo conavergence rates: Energy-scaled displacement Monte Carlo," *J. Chem. Phys.*, **79,** 3938 (1983).
[7] M. Mezei, K.A. Bencsath, S. Goldman and S. Singh, " The detailed balance energ-scaled displacement Monte Carlo algorithm," *Mol. Sim.*, **1,** 87 (1987).
[8] W. Chapman and N. Quirke, "Metropolis Monte Carlo simulation of fluids with multiparticle moves," *Physica*, **131B,** 34 (1985).