Smart Darting Monte Carlo

Ioan Andricioaei[a] and John E. Straub[b]
Department of Chemistry, Boston University, Boston, Massachusetts 02215

Arthur F. Voter
Theoretical Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87545

(Received 18 September 2000; accepted 2 February 2001)

The ''Smart Walking'' Monte Carlo algorithm is examined. In general, due to a bias imposed by the interbasin trial move, the algorithm does not satisfy detailed balance. While it has been shown that it can provide good estimates of equilibrium averages for certain potentials, for other potentials the estimates are poor. A modified version of the algorithm, Smart Darting Monte Carlo, which obeys the detailed balance condition, is proposed. Calculations on a one-dimensional model potential, on a Lennard-Jones cluster and on the alanine dipeptide demonstrate the accuracy and promise of the method for deeply quenched systems. © *2001 American Institute of Physics.*
[DOI: 10.1063/1.1358861]

I. BACKGROUND

In recent years, a variety of algorithms have been developed to address the problem of ''broken ergodicity'' in the simulation of complex many body systems.[1] One of the more promising classes of methods is based on the ''simulated tempering'' Monte Carlo method;[2] the class can be thought to include multicanonical,[3] parallel tempering,[4] J-walking,[5] and q-jumping[6–8] Monte Carlo. Simulations based on such methods employ two or more Monte Carlo walkers at a series of higher temperatures (or on a modified potential surface) where the sampling efficiency is greater than that at the temperature of physical interest. By an exchange of walkers (in a way that satisfies detailed balance), it is possible to improve the sampling for the computation of thermostatistical averages at the temperature (or on the potential surface) of interest. One shortcoming of such methods is that the distribution of walkers must sample energetic distributions with a significant overlap in probability so that the exchange of walkers between temperatures (or potential surfaces) occurs with reasonable probability. This problem can be acute in the simulation of deeply quenched systems.

Consider, for example, a peptide or protein at low temperature. A number of conformational ''basins'' may contribute to the thermostatistical average properties such as the equilibrium distribution of solvent molecules in vacancies of the protein. It is highly desirable to be able to compute the relative free energies of the ''quenched states'' of the system without simulating a manifold of states connecting in energy the quenched state distributions.

An alternative approach, discussed in this work, involves the direct exchange of walkers using ''displacement'' moves constructed from a set of low energy configurations of the system. To be more precise, consider the partition function for the ensemble of states for N distinguishable particles,

$$Q(\beta)=\left(\prod_{k=1}^{N}\Lambda_k^d\right)^{-1}\int d\mathbf{r}e^{-\beta U(\mathbf{r})}=\left(\prod_{k=1}^{N}\Lambda_k^d\right)^{-1}Z(\beta), \tag{1}$$

where $\Lambda_k=(h^2\beta/2\pi m_k)^{1/2}$ is the thermal de Broglie wavelength for a particle of mass m_k, $U(\mathbf{r})$ is the potential energy, and $Z(\beta)$ is the configuration integral. In the inherent structure picture of statistical mechanics, proposed by Stillinger and Weber,[9] the dN dimensional configuration space is decomposed as a set of basins of attraction. Any point in configurational space (excluding maxima, saddle points, and ridges) will be mapped by a steepest descent to a minimum on the potential surface. Labeling the basins i and calling $\mathcal{B}_i$ the regions of configuration space which form basins of attraction draining to the ith minimum located at $\mathbf{R}_i$ of energy U_i, the configuration integral may be written

$$Z(\beta)=\sum_i Z_i(\beta)=\sum_i\int_{\mathcal{B}_i}d\mathbf{r}e^{-\beta U(\mathbf{r})}. \tag{2}$$

The potential energy in the ith basin can be written $U(\mathbf{r})=U_i+\Delta_i U(\mathbf{r})$ leading to

$$Z_i(\beta)=e^{-\beta U_i}\int_{\mathcal{B}_i}d\mathbf{r}e^{-\beta\Delta_i U(\mathbf{r})}. \tag{3}$$

These are exact expressions for the configuration integrals. Alternatively, we can write the partition function as

$$Q(\beta)=\sum_i \exp[-\beta F_i], \tag{4}$$

where F_i is the Helmholtz free energy of the ith basin. A method that couples local sampling of the basins $\mathcal{B}_i$ together with a convenient means of transportation between all the basins would constitute a good approach towards the sampling of the available configuration space. That is the object of this study.

[a]Present address: Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts 02138.
[b]Author to whom correspondence should be addressed. Phone: (617) 353-6816; electronic mail: straub@bu.edu

II. SMART WALKING MONTE CARLO

In a Monte Carlo simulation, we hope to generate the equilibrium Boltzmann distribution of states

$$\rho(\mathbf{r}) = \frac{1}{Z(\beta)} \exp[-\beta U(\mathbf{r})]. \tag{5}$$

However, for many systems of interest, high energetic barriers separate basins on the potential surface. This can make it difficult for the Monte Carlo walk to move between basins and sample the equilibrium distribution of states. This problem is referred to as ''broken ergodicity.''[1]

An effective approximate algorithm known as ''Smart Walking'' Monte Carlo has been proposed to address this problem.[10] This S-walking algorithm is a variation on the J-walking Monte Carlo of Frantz, Freeman, and Doll.[5] In the S-walking algorithm, there are two types of moves. One move is of the normal Metropolis MC type. Trial moves occur with a probability $T(\mathbf{r}\rightarrow\mathbf{r}') = \text{constant}$ within a local region and zero otherwise. Since the trial move distribution is symmetric, $T(\mathbf{r}\rightarrow\mathbf{r}') = T(\mathbf{r}'\rightarrow\mathbf{r})$, the acceptance probability that generates the Boltzmann distribution is

$$p = \min[1, q(\mathbf{r}\rightarrow\mathbf{r}')], \tag{6}$$

where

$$q(\mathbf{r}\rightarrow\mathbf{r}') = \frac{T(\mathbf{r}'\rightarrow\mathbf{r})\rho(\mathbf{r}')}{T(\mathbf{r}\rightarrow\mathbf{r}')\rho(\mathbf{r})} = \exp[-\beta(U(\mathbf{r}') - U(\mathbf{r}))]. \tag{7}$$

Moves generated with this acceptance probability are designed to give effective *local* sampling within basins of the potential energy surface. From here on we will use the term ''walker'' to describe the point in configuration space that is moving during the simulation.

A second type of move used in S-walking is a ''jump'' move designed to transport the walker between basins separated by significant energetic barriers, thereby facilitating *global* sampling. The trial move distribution for the ''jump'' moves is taken to be the set $\{\mathbf{R}_i\}$ of minima of the potential energy basins, obtained by quenching the Boltzmann distributed phase points at the elevated temperature T_J. Each minimum may be found multiple times with the relative number of occurrences being proportional to the partition function for that basin for the case of adequate sampling. That is,

$$T(\mathbf{r}\rightarrow\mathbf{r}') = \sum_i w_i(\beta_J)\,\delta(\mathbf{r}' - \mathbf{R}_i), \tag{8}$$

where $\mathbf{R}_i$ is the position of the minimum of the ith basin and

$$w_i(\beta_J) = \frac{Z_i(\beta_J)}{Z(\beta_J)} \tag{9}$$

is the probability of visiting the ith basin at temperature T_J. The sum of the weights is unity.

In S-walking, the trial ''jump'' moves are accepted with probability

$$p_J = \min[1, q_J(\mathbf{r}\rightarrow\mathbf{r}_i)], \tag{10}$$

where

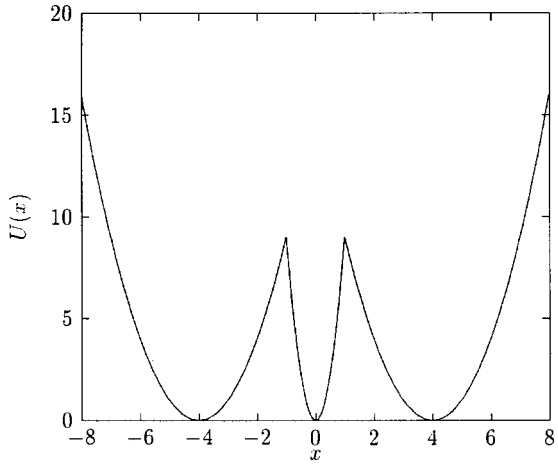

FIG. 1. The three-parabola potential. At high temperatures, the central well is only rarely visited.

$$q_J(\mathbf{r}\rightarrow\mathbf{R}_i) = \exp[-\beta(U(\mathbf{R}_i) - U(\mathbf{r}))]. \tag{11}$$

However, since the trial move distribution is a function of the jumping temperature, and this bias is not corrected in the S-walk acceptance probability, there is no guarantee that the equilibrium Boltzmann distribution is sampled by the S-walk.

We would like to be able to generate the S-walk acceptance probability for the jump moves from the condition of microscopic reversibility,

$$q(\mathbf{r}\rightarrow\mathbf{R}_i) = \frac{T(\mathbf{R}_i\rightarrow\mathbf{r})\rho(\mathbf{R}_i)}{T(\mathbf{r}\rightarrow\mathbf{R}_i)\rho(\mathbf{r})}. \tag{12}$$

We want to balance the probability of moving $\mathbf{r}\rightarrow\mathbf{R}_i$ with that of moving $\mathbf{R}_i\rightarrow\mathbf{r}$. However, using a jump move it is only possible to move to a *minimum* of a basin. Therefore, the trial move $\mathbf{R}_i\rightarrow\mathbf{r}$ is not allowed unless $\mathbf{r}$ is a basin minimum. It follows that in general microscopic reversibility is not satisfied.

III. ONE-DIMENSIONAL EXAMPLE

We can demonstrate this point by considering a potential composed of three parabolas,

$$U(x) = \begin{cases} k_2(x+\xi)^2 & \text{if} \quad x < -x^* \\ k_1 x^2 & \text{if} \quad -x^* \leqslant x \leqslant x^* \\ k_2(x-\xi)^2 & \text{if} \quad x > x^*, \end{cases} \tag{13}$$

as shown in Fig. 1.

Let us assume good local sampling in the wells and that broken ergodicity ensures that the only means of transportation between wells is by a jump to a minimum from the distribution generated at the higher temperature T_J. In order for the average of a physical property to be correct, one requires that the total ''time'' spent in the each basin be proportional to $Z_i(\beta)$. However, in a general S-walk simulation, if the simulation is of adequate length, the period actually spent in each well is proportional to $Z_i(\beta_J)$.

At very high temperatures T_J, the probability of the random walker is much greater to the left and to the right of the

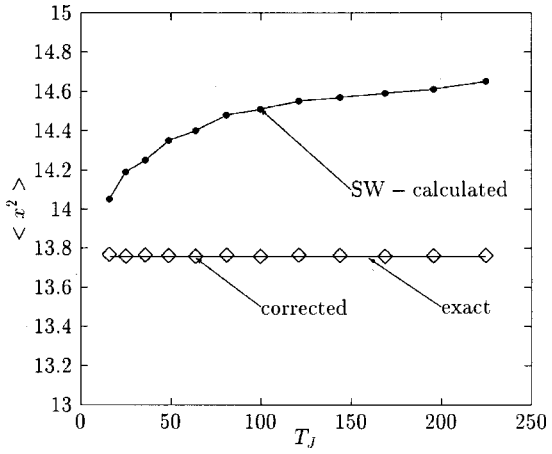

FIG. 2. The dependence of the average mean square position at a thermal energy of $T=0.1$ (in reduced energy units) as a function of the jumping temperature T_J. At high jumping temperatures, the S-walk is biased towards the outermost wells. The contribution from the central well is underestimated and the average exceeds the exact value. Correcting the S-walking MC method eliminates the bias and produces results that are in good agreement with the exact average.

two energy barriers. The central parabola is very seldom visited. This corrupts the S-walk average since the trial move distribution introduces a bias, dependent on T_J, that is not corrected by the acceptance probability.

A numerical study of this potential employing the S-walk algorithm yielded an average for x^2 different than the exact average. The parameters were chosen such that the finite energy barriers are equal to 9 energy units ($\xi=4, k_1=9, k_2=1$). For $T_J=100$ and $T=0.1$ the result is too large by 20%. For $T_J=10\,000$ and $T=0.1$ the result is too large by 60%.

This can have catastrophic repercussions, especially if the ''physics'' happens exactly in this middle region. The sampling in the peripheral regions will be wasted.

In principle, one can correct the average calculated by the S-walking algorithm by including the correct state-by-state configuration integrals into the acceptance criterion in such a way as to undo the incorrect biasing of each state, so that Eq. (8) becomes

$$T(\mathbf{r}\rightarrow\mathbf{r}')=\sum_i Z_i(\beta)\,\delta(\mathbf{r}'-\mathbf{R}_i). \tag{14}$$

In the simple one-dimensional example of the three parabola potential, this is easily accomplished because one can analytically compute the necessary correction ratios. The corrected average of the squared position for several values of the upper temperature T_J is shown in Fig. 2, for the case of the three parabola model. However, in general, one will not know partition functions for each state at the temperature of interest, T.

IV. SMART DARTING MONTE CARLO

We have seen that the main drawback of the Smart Walking algorithm is the breakdown of microscopic reversibility. If, in the evolution of the Markov chain, the system goes from one microscopic state to a minimum energy state, the reverse move is not possible (unless the original state is itself at a minimum). In this section we present a new method based on a combination of (1) the Smart Walking scheme and (2) the use of displacement vectors connecting minimum energy configurations, based on an idea employed previously.[11,12] The resulting algorithm is shown to obey detailed balance. The gist of the method is to perform, among good local sampling moves, jumps not exactly to the local minimum of another basin, but from and to suitably chosen small neighborhoods around such minima. As shown below, this can be done in such a way that detailed balance is achieved.

The Smart Darting method can be formulated as follows. In the process of some search method (high temperature MC, q-jumping MC,[6] etc.), pick configurations from which to perform steepest descents on the potential energy surface, generating a set $\{\mathbf{R}_i\}_{i=1,2,\ldots,M}$ of M distinct local minimum configurations. From this set of minima construct $M(M-1)$ displacement vectors, or ''darts,'' of the type

$$\mathbf{D}_{ij}=\mathbf{R}_j-\mathbf{R}_i, i\neq j, i,j=1,2,\ldots,M. \tag{15}$$

Also, choose a real number ϵ. Then define an ϵ-sphere around each element of the set $\{\mathbf{R}_i\}$,

$$S_\epsilon(\mathbf{R}_i)=\{\mathbf{r}\,|\,\|\mathbf{r}-\mathbf{R}_i\|<\epsilon\}. \tag{16}$$

For efficient sampling, the value of ϵ should be chosen small in the following sense. The difference in the potential energy of any configuration within any $S_\epsilon(\mathbf{R}_i)$ and the configuration $\mathbf{R}_i$ of the local minimum of that ϵ-sphere should be much less than the thermal energy of a degree of freedom. Loosely speaking, ϵ should be smaller than the size of the smallest catchment basin. The strict requirement on ϵ is that it be small enough that no two spheres overlap (or the sampling procedure presented here requires modification). During the habitual Monte Carlo sampling done locally within a basin, check, with probability P, whether the current configuration $\mathbf{r}$ is in an ϵ-sphere and do one of the following two things:

(1) If it is, say, $\mathbf{r}\in S_\epsilon(\mathbf{R}_k)$, then randomly pick another local minimum, say, the lth one, and jump to the $S_\epsilon(\mathbf{R}_l)$ sphere by the translation,

$$\mathbf{r}\rightarrow\mathbf{r}+\mathbf{D}_{kl}. \tag{17}$$

Accept or reject the move according to the Boltzmann criterion.

(2) If $\mathbf{r}$ is outside any ϵ-sphere, then *count again* the current configuration $\mathbf{r}$ (i.e., reject an implicitly attempted jump along $\mathbf{D}_{kl}$ because it would land outside the ϵ-sphere for minimum l). The remainder of the simulation steps, on the average a fraction $1-P$ of them, are local MC steps drawn from a uniform distribution and accepted or rejected according to the Boltzmann criterion.

For this algorithm, one can prove that detailed balance is guaranteed and the exact average of any configuration-dependent property over the accessible space is obtained. Two key issues determine the detailed balance. The first is the fact that the trial probability to pick the displacement vector $\mathbf{D}_{kl}$ to go from the kth ϵ-sphere to the lth one equals the trial probability to pick the displacement vector $\mathbf{D}_{lk}$ for the reverse step. The second issue is that the trial probability

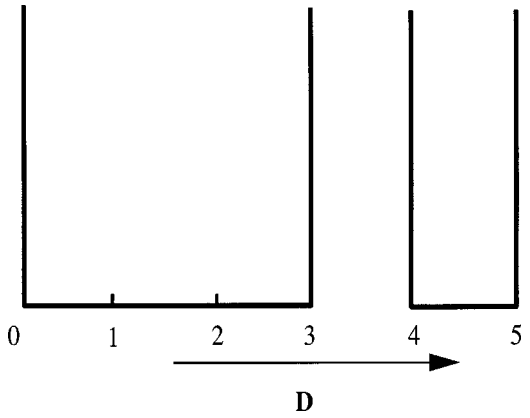

FIG. 3. The square double well potential.

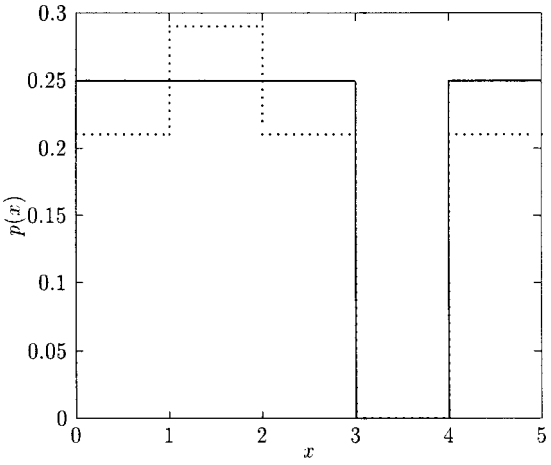

FIG. 4. Probability density function for the square double well potential. The thick line is the correct density, while the dotted line results from omitting the double counting when the walker is outside any ϵ-sphere.

for a local MC step that moves the walker from a point inside an ϵ-sphere to a point outside that sphere is the same as for the reverse move, i.e., $(1-P)$ times what it would be in a walk restricted to local moves. To give insight in the details of the scheme, consider a simple example of a one-dimensional potential made of square wells.

Say a random walk is performed in the double square well potential,

$$U(x)=\begin{cases}0 & \text{if} \quad x\in[0,3]\cup[4,5]\\ \infty & \text{otherwise}\end{cases} \tag{18}$$

shown in Fig. 3. We assume a small bias that makes the centers of the two segments the local minima. Choose $2\epsilon=5-4=2-1$. There are two displacement vectors, **D** (shown in Fig. 3), and $-\mathbf{D}$. They are used for interwell transitions, only between $[1,2]$ and $[4,5]$. The rules of the walk are:

(1) If $x\in[0,1)$, the walker moves locally with probability $1-P$; the rest of the time, i.e., with probability P, it does not move and the position is counted again;
(2) If $x\in[1,2]$, the walker moves locally with probability $1-P$, or is transported to the other well with probability P;
(3) If $x\in(2,3]$, see (1);
(4) If $x\in[4,5]$, see (2).

It is easy to see, by writing down the master equation, that $1-P$ is the correct weighting for the local moves out of the regions complementary to the ϵ-spheres. Only then is the detailed balance obeyed. If this is not accounted for, then the erroneous distribution plotted in Fig. 4 is obtained.

That this is so can be seen when taking into consideration two points, $x\in[1,2]$ and $x'\in[0,1]$ (see Fig. 3) within Δ of each other. Here Δ is the maximum displacement of a local Monte Carlo move. The probability to move from x to x' is $W(x\to x')=(1-P)/\Delta$. Since, in the reverse move, x is accessible from x' only by a local move, $W(x'\to x)$ would equal $1/\Delta$ and detailed balance would fail to hold. Thus we see that the walker should move out of x' not always, but with probability $1-P$, and, for reasons of conservation of probability, has to stay in x' with probability P. This is why we count again, with probability P, points out of ϵ-spheres. It is easy to extend the proof to nonflat, one-dimensional potentials and further to multiple dimensions.

Three-parabola model results are shown in Fig. 5, where we represent the average of the position squared at different temperatures, compared with the exact result obtained by analytical integration. For this example, the S-walking MC results stand outside the bounds of the plot.

It is worth discussing the design of the Smart Darting Monte Carlo approach from the point of view of computational efficiency. Conceptually, the following procedure would be simpler than the one discussed so far, in that it eliminates the ϵ-spheres and the associated rejection of null moves. First, construct the set of $M(M-1)$ darts representing both forward and reverse moves between all pairs of minima. At each step, with probability P, choose one of these $M(M-1)$ darts at random, without regard to the current position of the walker. Accept or reject this jump with the usual Metropolis probability. This gives a valid MC walk with unbiased trial move probabilities;[11] we have simply ex-

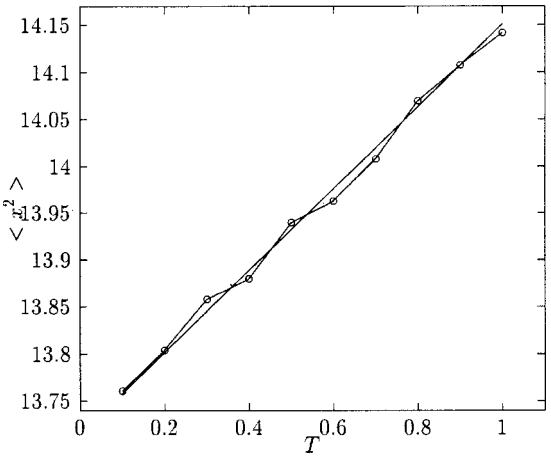

FIG. 5. The dependence of the average mean square position, in the potential of the three-parabola model, as a function of the temperature T, generated using Smart Darting Monte Carlo (line with circles) overlapped with the exact theoretical dependence (continuous line).

tended the normal list of possible moves to include the set of dart moves. It is not an efficient procedure, however, for two reasons. Selecting randomly from all the darts will give a dart that originates in the current state only a fraction $(1/M)$ of the time. As M becomes large, this is a serious problem, as most dart jumps will be rejected. Second, even when a ''proper'' dart (one that originates in the present basin) is chosen, the probability that the jump will be accepted decays with increasing distance from the basin minimum. The creation of the ϵ-sphere solves both of these problems. Because all ϵ-spheres are the same size and the darts connect the centers, every point in one sphere maps (via darts) onto an equivalently located point in every other sphere. Thus, the trial probabilities for the forward and reverse jumps are automatically matched for jumps originating within spheres, and are zero for points outside any sphere. The second problem is addressed by limiting the size of ϵ, as discussed above. So, in the procedure proposed above, the $1/M$ inefficiency is eliminated, but the procedure is slightly more complicated. It requires determining if the walker is presently within a sphere and double counting of positions outside the spheres P fraction of the time. Moreover, if the walker enters a new basin purely by local moves (as might happen at high temperatures), this must be recognized, because the walker is now entering the proximity of a different ϵ-sphere. For large M, efficient implementation of this monitoring may require neighbor-listing of the minima or some other form of bookkeeping, to prevent frequent searches over all M minima.

A method similar in spirit with the one we presented is the Jumping Between Wells method,[13] which uses the same paradigm of generating good thermal averages once important low-energy conformations are known. While it has been successful for the calculations of free energy differences for small organic molecules, the JBW method lacks the concept of the ϵ-sphere and jumping is allowed from any configuration. As a result, it is expected to perform poorly as the complexity of the potential energy surface increases, because jumping from points far from low energy minima will be rejected with a large probability, as discussed above.

V. APPLICATION TO LENNARD-JONES CLUSTERS

We have applied the method of Smart Darting Monte Carlo to the calculation of thermodynamic averages of clusters of atoms. In this section we describe the implementation of the algorithm and the results obtained.

We model the interaction between pairs of atoms in the cluster by the Lennard-Jones potential,

$$V(r)=\frac{1}{r^{12}}-\frac{2}{r^{6}}, \tag{19}$$

where the energy unit is the well depth and the length unit is the separation between two atoms when at equilibrium. To prevent the ''vaporization'' of atoms during simulations at high temperatures, we added a confining potential of the type (r/r_0),[20] where r_0 is larger than the size of the minimum configuration of the cluster. This confinement has negligible effect on the energetics of the ''condensed'' configurations of the clusters. To eliminate the six degrees of freedom corresponding to the rotation and translation of the cluster, one atom was fixed at the origin of the coordinate system, a second atom was constrained to move on a line containing the first atom, and a third atom was constrained to move in a plane containing this line.

The number of minima increases rapidly with the number of atoms in the cluster. Hoare and Pal[14] have cataloged the configurations of small clusters. An 8-atom cluster possesses 8 minima, which we have found and stored for use by the algorithm. Connecting these minimum energy configurations thus creates 56 darts. The reduced temperature was 0.05.

To measure the extent to which the phase space is sampled we have computed the ergodic measure, whose dependence on the simulation time gives the rate of self-averaging in an equilibrium calculation.[15] Self-averaging is a necessary, but not sufficient, condition for the ergodic hypothesis to be satisfied. The rate of self-averaging for a given property is expected to be proportional to the rate of phase space sampling. We have chosen to use the potential energy as the property whose average is calculated for two independent trajectories α and β. We define the ''move average'' over the Monte Carlo trajectory of the potential energy U for the jth particle along the α trajectory after n moves as

$$u_j^{\alpha}(n)=\frac{1}{n}\sum_{k}^{n} U_j(\mathbf{r}_k^N). \tag{20}$$

The ergodic measure is then defined as the sum over N particles,

$$d_U(n)=\frac{1}{N}\sum_{j} [u_j^{\alpha}(n)-u_j^{\beta}(n)]^2. \tag{21}$$

For an ergodic system, if $n\to\infty$, then $d_U(n)\to 0$. For large n we expect the form of the convergence to be diffusive,[16]

$$\frac{d_U(0)}{d_U(n)}=D_U n, \tag{22}$$

where D_U is a rate for self-averaging of U over the two independent trajectories. We associate rapid and effective sampling of phase space with a large value of D_U. The choice of the potential energy metric is arbitrary. However, it has been shown to be a good measure of the extent of phase space sampling in a variety of systems.[17]

The results are shown in Fig. 6, where the energy metric is displayed as a function of the number of Monte Carlo steps; the plot is the result of averaging over 10 runs with the same two initial basins but different random seeds, and thus different evolution in configurational space. The darting probability is $P=0.1$. In Fig. 7, the continuous line represents a typical ''time'' series of the Euclidean distance Δs of the current configuration from the minimum of the current basin in which the last dart landed. The tallest peaks are due to ''darting.'' They are followed by a rapid decrease of the distance to the local minimum.

To address the problem of the dependence of the result on the number of minima we have also calculated the average internal energy of the cluster not only by constructing all the darts between all the eight minimum energy configura-

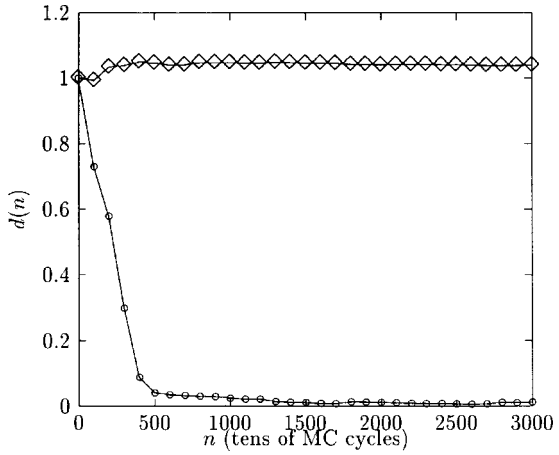

FIG. 6. The ergodic measure for the Smart Darting MC (shown with circles) and for the standard MC (shown with diamonds).

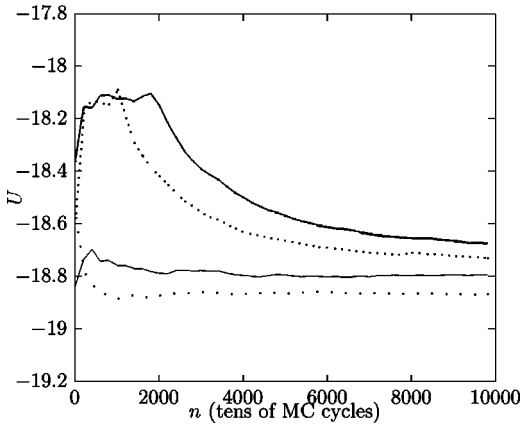

FIG. 8. The average potential energy for the 8-atom cluster in the following cases, from top to bottom, respectively, using the Smart Darting method when only 2 of the minima considered (top, thick line), when only 5 of the minima are considered (dotted line), with all 8 minima considered (thin line), and using the Smart Walking method with all 8 minima considered (bottom, rare dotted line).

tions, but also by leaving out a number of the minima, chosen at random. Figure 8 presents the results when two, five, and all eight minima are used in the construction of the darting vectors. Even if the exclusion of some of the minima still gives the average around the correct result, one expects that in general applications of this method a sufficient number of the thermodynamically significant conformational states are at hand. We also plot in this figure the average energy for the case of the Smart Walking algorithm[10] using all 8 minima. The slightly different value of this thermodynamic average is due to the breakdown of detailed balance.

VI. APPLICATION TO ALANINE DIPEPTIDE

We have performed tests on the alanine dipeptide, N-acetylalanyl-N-methylamide (AcAlaNHMe), a model system used extensively in theoretical studies of the conformational equilibrium in proteins. The peptide is on one hand small enough to allow thermodynamical[18] and quantum mechanical[19] calculations, and on the other hand has backbone configurations which are prototypical for the polypeptide backbone of proteins.

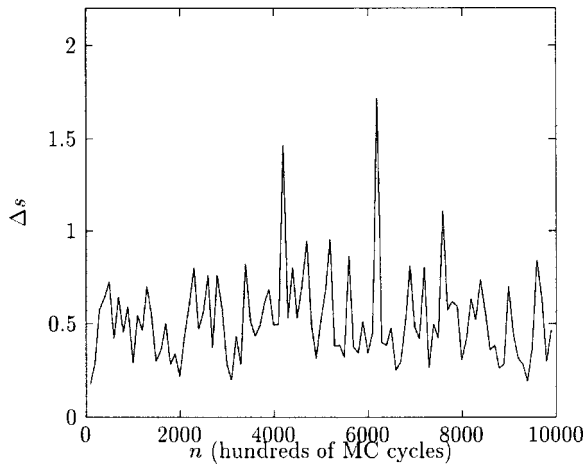

FIG. 7. The distance of the current position from the position of the minimum of the last basin at low temperature, when the assumption that interbasin moves occur only by darting is true.

We have used the CHARMM force field with the version 19 parameter set in vacuum. The conformational distribution of the alanine dipeptide in vacuum is dominated by two configurations, $C_{7\mathrm{eq}}$ and $C_{7\mathrm{aq}}$, and we have used these two conformations to create the (two) darts. In order to describe the conformational sampling efficiency, we have used the nonbonded force metric, which has been originally used for proteins[17] and has been proven to be a good measure of ergodicity for such systems. The nonbonded force metric is similar to the energy metric used for the cluster, except that the self-averaging quantity is now the force coming from the nonbonded interactions of the peptide. That is, we calculate, for two trajectories α and β, the move average

$$\mathbf{F}_j^{\alpha}(n)=\frac{1}{n}\sum_{k}^{n}\mathbf{F}_j(\mathbf{r}_k^N). \tag{23}$$

The nonbonded force metric is defined as the average over the number of particles of the mean square difference of the two move averages,

$$d_{\mathbf{F}}(n)=\frac{1}{N}\sum_{j}[\mathbf{F}_j^{\alpha}(n)-\mathbf{F}_j^{\beta}(n)]^2. \tag{24}$$

For an ergodic system, we again expect $d_{\mathbf{F}}(n)\to 0$ as $n\to\infty$.

Figure 9 shows the force metric as a function of the Monte Carlo sweeps for regular MC vs Smart Darting MC. The temperature was $T=200$ K, the darting probability $P_J=0.1$ and we have used a value of $\epsilon=1.5$ Å. At this low temperature, the peptide faces a rugged energy landscape with barriers that cannot be overcome using standard Monte Carlo in the time scale of our simulations.

VII. DISCUSSION

Due to its importance for condensed phase simulation, a host of methods have been devised to calculate equilibrium averages by an enhanced configurational space search of systems suffering from broken ergodicity. There exists a num-

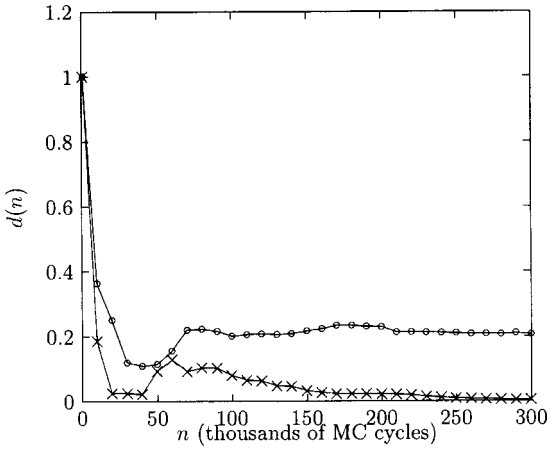

FIG. 9. The nonbonded force ergodic measure computed during simulations of the AcAlaNHMe dipeptide using the Smart Darting MC (shown with crosses) and for the standard MC (shown with circles).

ber of methods which efficiently identify collections of local minima, ranging from simple random generation of configurations followed by local minimization,[20] to more elaborate propulsions over the saddles of the potential energy surface.[21] The method presented here is an implementation of an exact (under reasonable assumptions) algorithm that calculates thermodynamic averages over the available configurational space, once its minima have been found.

The Smart Darting MC algorithm is effective in overcoming broken ergodicity and providing accurate equilibrium averages. The local minimum configurations that form the center of the ϵ-spheres, can, in principle, be generated in any number of ways including sampling from non-Gibbs–Boltzmann distributions. Care must be taken, however, to choose ϵ small enough that trial jumps have a significant probability of being accepted.

To achieve this efficiency, the size of ϵ can be determined in a number of ways. One way would be to perform a standard MC run at the temperature of interest and measure the standard deviation of the position of the walker. Alternatively, one could choose a random direction, plot the potential energy dependence along that direction, and choose ϵ as an average of the length around minima along which the variation of the potential energy is much less than the thermal energy. The only strict requirement on ϵ is that the spheres do not overlap, or the sampling rules must be modified.

The advantage of this method, relative to other techniques such as simulated tempering, parallel tempering, J-walking or q-jumping Monte Carlo is that one is relieved of the burden of connecting states with overlapping distributions created by a manifold of temperatures or deformed potential energy surfaces. As such, the algorithm is particularly well suited for the simulation of deeply quenched systems such as peptides, proteins, or spin systems at low temperature.

ACKNOWLEDGMENTS

J.E.S. recognizes the generous support of the National Science Foundation (CHE-9975494) and the Petroleum Research Fund of the American Chemical Society (34348-AC6). Work at Los Alamos National Laboratory was supported by the United States Department of Energy (DOE), Office of Basic Energy Sciences, under DOE contract No. W-7405-ENG-36.

[1] B. J. Berne and J. E. Straub, Curr. Opin. Struct. Biol. **7**, 181 (1997).
[2] E. Marinari and G. Parisi, Europhys. Lett. **19**, 451 (1992).
[3] B. A. Berg and T. Neuhaus, Phys. Lett. B **267**, 249 (1991).
[4] C. Geyer and E. Thompson, J. Am. Stat. Assoc. **90**, 909 (1995).
[5] D. D. Frantz, D. L. Freeman, and J. D. Doll, J. Chem. Phys. **93**, 2769 (1990).
[6] I. Andricioaei and J. E. Straub, J. Chem. Phys. **107**, 9117 (1997).
[7] Y. Pak and S. M. Wang, J. Chem. Phys. **111**, 4359 (1999).
[8] Y. Pak and S. Wang, J. Phys. Chem. B **104**, 354 (2000).
[9] F. H. Stillinger and T. A. Weber, Phys. Rev. A **25**, 978 (1982).
[10] R. Zhou and B. J. Berne, J. Chem. Phys. **107**, 9185 (1997).
[11] A. F. Voter and J. Doll, J. Chem. Phys. **80**, 5814 (1984).
[12] A. F. Voter, J. Chem. Phys. **82**, 1890 (1985).
[13] H. Senederowitz, F. Guarnieri, and W. Still, J. Am. Chem. Soc. **117**, 8211 (1995).
[14] M. R. Hoare and P. Pal, Adv. Chem. Phys. **20**, 161 (1971).
[15] D. Mountain and D. Thirumalai, J. Phys. Chem. **93**, 6975 (1989).
[16] D. Thirumalai and R. D. Mountain, Physica A **210**, 453 (1994).
[17] J. E. Straub and D. Thirumalai, Proc. Natl. Acad. Sci. U.S.A. **90**, 809 (1993).
[18] D. Tobias and C. L. III, J. Phys. Chem. **96**, 3864 (1992).
[19] T. Head-Gordon *et al.*, J. Am. Chem. Soc. **113**, 5989 (1991).
[20] Z. Li and H. A. Scheraga, Proc. Natl. Acad. Sci. U.S.A. **84**, 6611 (1987).
[21] G. T. Barkema and N. Mousseau, Phys. Rev. Lett. **77**, 4358 (1996).