# Efficient Monte Carlo methods for the computer simulation of biological molecules

Djamal Bouzida*
*Department of Physics, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213*

Shankar Kumar
*Department of Biological Sciences, University of Pittsburgh, Pittsburgh, Pennsylvania 15260*

Robert H. Swendsen
*Department of Physics, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213*
(Received 29 August 1991)

We present an alternative approach to efficient Monte Carlo simulations of biological molecules. By relaxing the usual restriction to Markov processes, we are able to optimize performance while dealing directly with the inhomogeneity and anisotropy inherent in these systems. This approach allows us to sample configurational space more efficiently than with either standard Monte Carlo or molecular-dynamics methods.

PACS number(s): 87.15.−v, 02.50.+s, 31.15.+q

## I. INTRODUCTION

In recent years, considerable effort and computational resources have been devoted to simulating biological molecules. These molecules, which include polymers, proteins, and nucleic acids, are of great importance in physics, chemistry, biology, and medicine [1,2]. The goal of these computer simulations is to provide insights into the structure-function relationships in biomolecular interactions and energetics [3] by supplying detailed information about the conformations and internal motions of biologically important molecules.

Due to the size and complexity of the task, these simulations require enormous amounts of supercomputer time. This has led us to investigate alternative methods for improving efficiency with the goal of reducing the burden on supercomputing resources, broadening the scope of applications, and increasing the reliability of the results.

Currently, most computer simulations of thermodynamic systems use either molecular-dynamics (MD) or Monte Carlo (MC) methods. Both methods involve the generation of molecular conformations to represent thermal fluctuations. Equilibrium properties are found by computing appropriate averages over the resulting set of conformations. The interactions between atoms in the molecule are represented by an effective Hamiltonian that has been constructed empirically from a variety of experimental information [4–8].

Molecular dynamics is a deterministic method that computes classical trajectories by iterating a discretized representation of Newton's equations. The total energy of the system is conserved so that the generated configurations trace out a microcanonical ensemble (if the system is ergodic). An advantage of the MD method is the possibility of following an explicit classical trajectory of the system.

Since it is often more convenient to analyze data obtained at constant temperature (and/or pressure), various modifications of the MD method have been developed. One possibility is to introduce terms representing noise and dissipation into the equations, which leads to a canonical ensemble (constant temperature) [9–11]. Another possibility is the rescaling of atomic velocities to impose a fixed average temperature [11,12]. Nosé has introduced particularly interesting modified equations of motion that couple the system to a fictitious external degree of freedom [13,14].

The main adjustable parameter that enters a molecular-dynamics simulation is the size of the time step. Increasing the time step moves the system through phase space more rapidly, but it can introduce errors and can even affect the stability of the algorithm. For either Gear [15] or Verlet [16] algorithms, the maximum time step for reasonable accuracy must be less than about $\frac{1}{15}$ of the shortest period of vibration. In the case of most biological molecules, this corresponds to roughly 1 fs, which is indeed commonly used in such work. It might also be noted in passing that since the highest frequencies are associated with the smallest masses, one way to increase the efficiency of the MD method without distorting the equilibrium properties would be to set all masses equal [17].

A Monte Carlo simulation is a stochastic Markov process that generates a sequence of configurations representing a canonical ensemble. Trial moves are generated from a random distribution and are either accepted or rejected with a probability given by the Boltzmann factor. The Markov property of the MC process means that the probability of transition to a new state depends only on the present state. A fundamental theorem of Markov processes states that if the transition probabilities satisfy detailed balance, and if any configuration can be reached from any other configuration in a finite number of steps with nonzero probability (ergodicity), then the simulation can correctly reproduce the equilibrium behavior [18].

 

There are also mixtures of MC and MD methods known as hybrid MC [19] and hybrid MD methods [20]. In these methods, larger time steps are used with a global acceptance step to ensure thermal equilibrium.

In developing alternative methods, we must keep in mind the nature of the system and the types of questions that are asked. Some of these questions concern the dynamics of short-time behavior (i.e., a few picoseconds), and the only current option is the MD method. However, most questions concern equilibrium properties. The equilibrium configuration is of major importance in determining biological function, and free-energy calculations are needed to predict and understand biochemical reactions. Many calculations of relaxation times actually require quasiequilibrium determination of free-energy barriers.

Early work by Northrup and McCammon [21] indicated that the MD method was more efficient than the MC method, even for equilibrium properties. This conclusion was based on their observation that the MC method required much more computer time to carry out the updating process. However, the two methods actually require about the same amount of computer time per sweep. When this is corrected for, their data, based on the rms deviations per sweep, indicate that the standard MC method is more than the MD method. We have chosen to base our approach on the MC method partly for this reason, but primarily because of the flexibility for introducing additional new moves in MC simulations to improve the efficiency.

Both Monte Carlo and molecular-dynamics methods were originally developed to simulate fluids. In setting up a simulation, both methods require preliminary calculations to equilibrate the system and to determine the appropriate simulation parameters (time step in the MD method or maximum jump size in the MC method) and to ensure stability in the case of the MD method. When simulating fluids, optimization of parameters is done on a global basis, which is appropriate because these systems are homogeneous and isotropic. However, the local structure of molecules is inhomogeneous and highly anisotropic [1,2,21]. Thus it is necessary to optimize the parameters locally in order to increase the efficiency of the simulation. Furthermore, both the inhomogeneity and anisotropy change with time.

We have developed alternative MC methods to address these problems. The essential feature of our approach is the optimization of the Monte Carlo parameters from data collected during the simulation. Because this feature allows us to optimize local moves, we are able to deal explicitly with the inhomogeneity and the local anisotropy of biological molecules. Our approach also allows for the optimization of a wide variety of global moves without long preliminary studies. This turns out to be extremely important in designing efficient algorithms.

Since our methods use information gathered during the simulation, they are no longer strictly Markovian, although they are *almost* or *piecewise* Markov processes [22]. Detailed balance is satisfied and we have been able to demonstrate that our algorithms reproduce the correct equilibrium behavior while substantially improving speed and efficiency. In this paper, we present a description of our approach and illustrate its advantages by applying it to a small molecule.

## II. ALMOST MARKOV SIMULATION METHODS

In this section, we derive two simulation methods from an analysis of a simple harmonic oscillator (SHO). The acceptance-ratio method (ARM) is an optimization technique that treats the inhomogeneity of biomolecules efficiently. The second method, the dynamically optimized Monte Carlo (DOMC) method, treats both the inhomogeneity and the anisotropy. These methods are then shown to retain their efficiency for the anharmonic potentials found in biological molecules.

The potential energy of a $d=1$ SHO is given by $V(x)=\frac{1}{2}kx^2$. The efficiency of a MC simulation is determined by the choice of maximum step size $\delta$ from which the MC trial moves are generated. If an optimum $\delta$ is known for some given $k$ and $\beta$, then the optimum $\delta$ for any other values of $k$ and $\beta$ is also known through the scaling relation

$$\tfrac{1}{2}k\beta\delta^2=F^2 , \tag{1}$$

where $\beta=1/k_BT$. Thus, if we can determine the optimal scale factor $F$ for any SHO, we have solved the problem for all such models.

An important quantity for understanding the efficiency of MC simulations is the average acceptance ratio $\langle P\rangle$, which is defined as the ratio of accepted moves to trial moves during a simulation. As shown in Fig. 1 (which

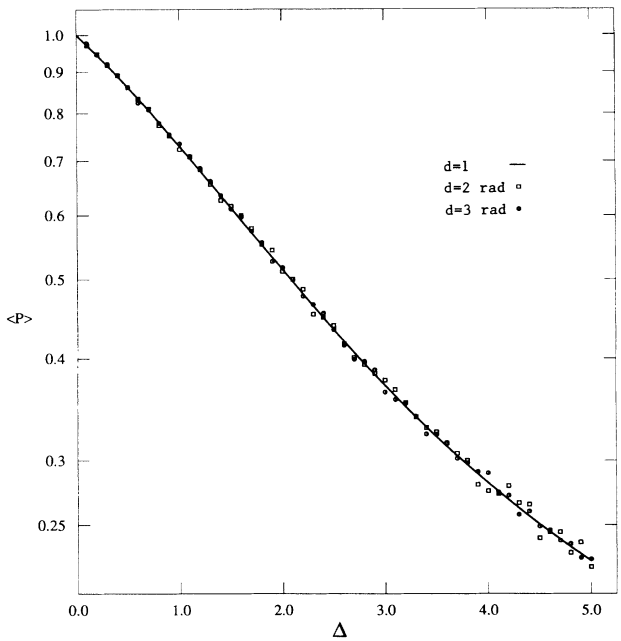


FIG. 1. Semilog plot for the average acceptance ratio $\langle P\rangle$ as a function of the maximum step size for $d=1$, 2, and 3 SHO with $\beta=1$. The one-dimensional curve is exact, while in higher dimensions, each point is obtained by averaging over 10 000 MC steps from a simulation where the trial jumps are generated uniformly in radius.

also shows data for higher dimensions discussed below), the acceptance ratio decreases monotonically as a function of step size $\delta$. This is expected because small trial step sizes corresponding to small energy changes will produce high acceptance ratios, while large moves have a high probability of being rejected due to large energy differences. The acceptance ratio decreases approximately exponentially as a function of $\delta$ for the range of values shown in Fig. 1. For larger $\delta$, the acceptance ratio varies inversely as $\delta$.

To determine the optimal value of $F$, we first performed a series of MC studies of a $d=1$ SHO, $V(x)=x^2$ with $\beta=1$. In addition to the acceptance ratio, we monitored the autocorrelation time $\tau$ and two measures of the displacements per MC step, $\langle(\Delta x)^2\rangle^{1/2}$ and $\langle|\Delta x|\rangle$, where $\Delta x$ represents the displacement during a MC move and the angular brackets indicate the usual thermal average.

The simplest measures of efficiency are the rms [21] and average absolute displacements. These quantities should go to zero for small acceptance ratio $\langle P\rangle$ since most moves are rejected, and for $\langle P\rangle$ near 1.0 since each trial move is small. Therefore, we expect a maximum in each curve, as seen in Fig. 2. The maximum rms displacement $\langle(\Delta x)^2\rangle^{1/2}$ occurs at $\langle P\rangle=0.42$ ($\delta=F=2.62$), while $\langle|\Delta x|\rangle$ has a maximum at the larger value of $\langle P\rangle=0.56$ ($\delta=F=1.76$). The optimal value of $F$ clearly depends on what is being calculated. However, even if $\langle P\rangle$ differs from the optimal value by as much as $\pm 0.15$, the displacements are only reduced by 10%. This leaves a fairly large region around $\langle P\rangle=0.5$ ($\delta=F=2$) for which both quantities are nearly optimized.

The integrated correlation time $\tau$ was determined from the normalized time-dependent energy-energy correlation function

$$f(t)=\frac{\langle E(t_0)E(t_0+t)\rangle-\langle E\rangle^2}{\langle E^2\rangle-\langle E\rangle^2} \tag{2}$$

by the usual expression

$$\tau=\sum_{i=1} f(t_i)\ , \tag{3}$$

where the sum is cut off when the fluctuations drive the correlation function negative [23]. The statistical error is proportional to $\sqrt{1+2\tau}$.

Figure 3 shows a plot of the correlation time $\tau$ as a function of the acceptance ratio $\langle P\rangle$. This plot shows a minimum in the correlation time with $\tau_{\min}\simeq 1.4$ MC steps corresponding to a 50% acceptance ratio. This provides a justification for the common practice of tuning the step size to accept about one-half of the trial moves. In fact, using $\langle P\rangle=0.5$ ($\delta=F=2$) to minimize $\tau$ also gives rms and average absolute displacements within 2.5% and 1.3% of their respective maxima.

In higher dimensions, the parameters that enter the optimization scheme are found to depend on how the jumps are generated. Two possibilities are to generate them uniformly in either the radius or the volume of a sphere (uniformly in radius or area of a circle for $d=2$). Interestingly, when the jumps are generated uniformly in radius, the acceptance ratio, the rms displacement, and the absolute displacement are nearly independent of the dimension as

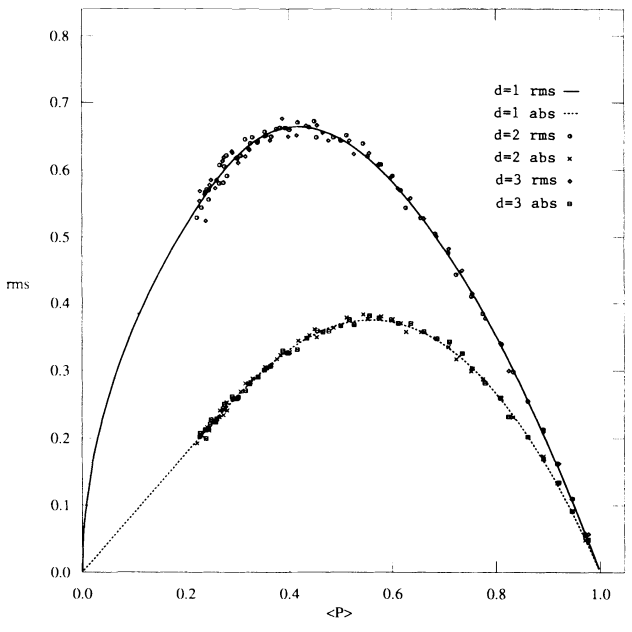


FIG. 2. The rms and absolute displacements vs acceptance ratio for $d=1$, 2, and 3 SHO with $\beta=1$. The one-dimensional curve is exact, while in higher dimensions, each point is obtained by averaging over 100 000 MC steps from a simulation where the trial jumps are generated uniformly in radius.

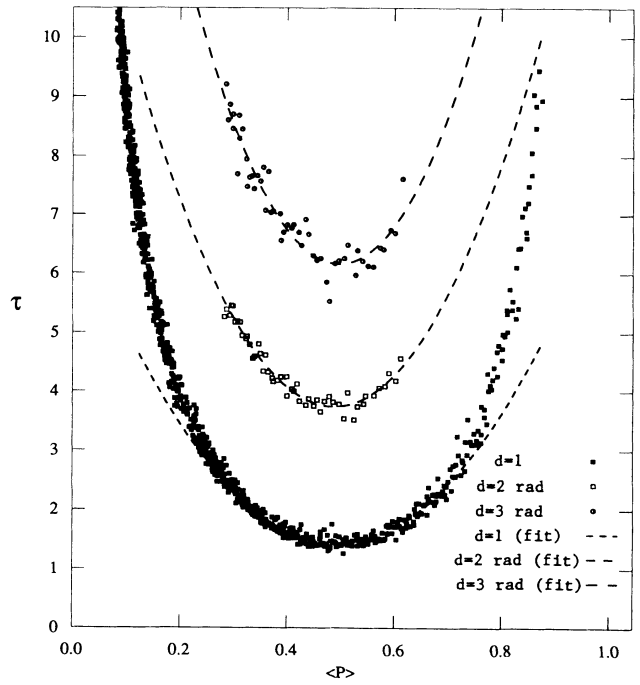


FIG. 3. Plot of the energy-energy autocorrelation time $\tau$ as a function of acceptance ratio $\langle P\rangle$ for $d=1$, 2, and 3 SHO with $\beta=1$. The trial jumps are generated uniformly in radius, and the dashed lines represent parabolic fits to the curves.

shown in Figs. 1 and 2. The correlation time still has a minimum at $\langle P \rangle = 0.5$, but its value increases for higher dimensions as shown in Fig. 3. The minimum correlation time is 3.8 MC steps in two dimensions, and 6.1 MC steps in three dimensions. Thus the optimal value of $F$ when the jumps are generated uniformly in radius is still $F=2$ corresponding to a $\langle P \rangle = 0.5$.

When the trial jumps are generated uniformly in volume, the behavior is somewhat different. As shown in Fig. 4, the acceptance ratio decreases more rapidly as a function of $\delta$ than in the previous case, but is still a nearly exponential function. The maximum rms and absolute displacements are higher than in the previous case but occur at lower acceptance ratios as shown in Fig. 5. The minimum correlation time occurs at lower acceptance ratios for higher dimensions as shown in Fig. 6. The minimum correlation time is 2.8 MC steps corresponding to $\langle P \rangle = 0.42$ in two dimensions, and it is 4.4 MC steps occurring at $\langle P \rangle = 0.39$ in three dimensions. The value of $F$ corresponding to the optimal correlation time decreases only slightly in higher dimensions, being about 1.9 for $d=2$, and about 1.8 for $d=3$ SHO.

Although most workers fix the step size for the duration of the simulation, some efforts have been made to update it as new information is generated. Allen and Tildesley suggested raising or lowering the global step size by 5% depending on whether the measured acceptance ratio is above or below 50% [24]. Corana *et al.* introduced variations in maximum step sizes to maintain the acceptance ratio at 50% in simulated annealing runs for minimizing functions of continuous variables [25]. Since they were not concerned with equilibrium properties, they did not discuss the non-Markovian nature of their procedure.

Our first optimization procedure is an acceptance-ratio method to carry out equilibrium simulations with different dynamically determined step sizes for each particle. We have used the approximately exponential dependence of $\langle P \rangle$ on $\delta$

$$\langle P \rangle = \exp(-\delta/\delta_0) , \tag{4}$$

where $\delta_0$ is some constant. Let $\langle P_i \rangle$ be the ideal or the

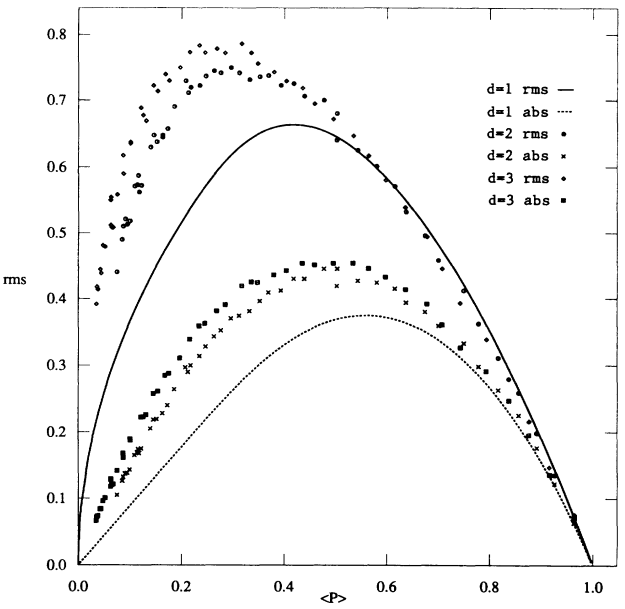


FIG. 5. Same as in Fig. 2, except the trial jumps are generated uniformly in volume (uniformly in area for $d=2$).

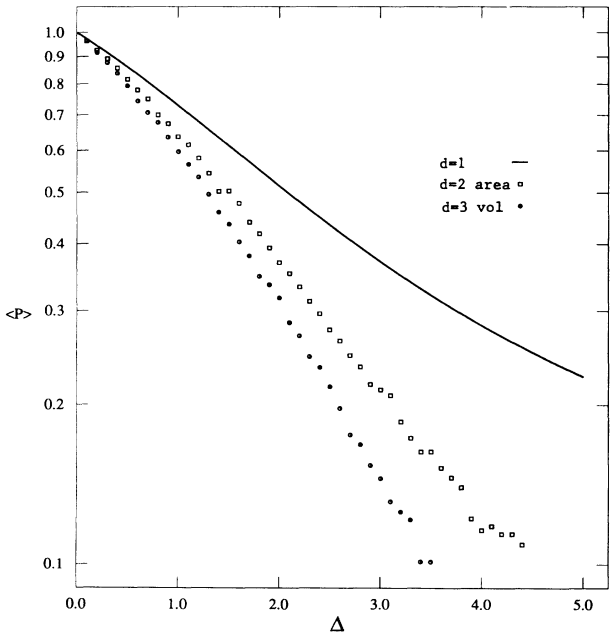


FIG. 4. Same as in Fig. 1, except the trial jumps are generated uniformly in volume (uniformly in area for $d=2$).

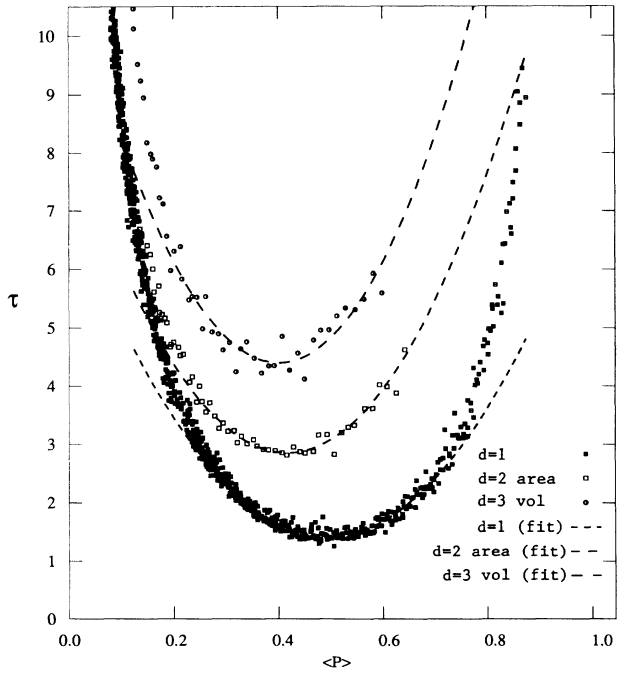


FIG. 6. Same as in Fig. 3, except the trial jumps are generated uniformly in volume (uniformly in area for $d=2$).

desired acceptance probability corresponding to an ideal maximum step size $\delta_i$. Then, clearly

$$\delta_i = \delta \frac{\ln\langle P_i \rangle}{\ln\langle P \rangle} . \tag{5}$$

A simulation is set up as a sequence of cycles. During a given cycle characterized by a maximum step size $\delta_{\text{old}}$, the acceptance ratio $\langle P_{\text{old}} \rangle$ for each particle is computed. An iteration procedure using Eq. (5) is set to update new values of $\delta$. However, this equation must be protected against overflow problems that occur whenever $P$ is either 0 or 1. Therefore, we have modified it to read

$$\delta_{\text{new}} = \delta_{\text{old}} \frac{\ln(a\langle P_i \rangle + b)}{\ln(a\langle P \rangle + b)} , \tag{6}$$

where $a$ and $b$ are real parameters chosen such that $\delta_{\text{old}}$ is either multiplied or divided by a convenient scale factor (about 5 or 10) whenever $P$ is 0 or 1 [25].

The ARM is a robust optimization technique and is especially useful at high temperatures for simulated annealing experiments. An advantage of the method is that each atom is treated separately, thus dealing with the inhomogeneity of the system efficiently. A weakness is that the accuracy of the optimization is limited by the discrete estimates of the acceptance ratio from the finite length of each simulation cycle. Another weakness is that the ARM does not deal with the local anisotropy of macromolecules, although it can be applied effectively to rotations of part of the molecule or other global moves. However, these problems can be dealt with using a different method we call the dynamically optimized Monte Carlo method.

We first discuss the one-dimensional DOMC equations for a $d=1$ SHO. Square brackets denote a direct average over all attempted moves regardless of whether they are accepted or not. It is easy to show that

$$[\Delta E] = \tfrac{1}{2} k [(\Delta x)^2] , \tag{7}$$

where $\Delta E$ represents the energy change corresponding to the jump size $\Delta x$. Combining this with Eq. (1) and eliminating $k$, we obtain the DOMC estimate for the optimum value of $\delta$ from the simulation data

$$\delta = F \left[ \frac{[(\Delta x)^2]}{\beta [\Delta E]} \right]^{1/2} , \tag{8}$$

or, more generally

$$\delta = F \left[ \frac{[(\Delta x)^{2n+2}]}{\beta [\Delta E (\Delta x)^{2n}]} \right]^{1/2} , \quad n = 0,1,2,\ldots . \tag{9}$$

This is the fundamental DOMC equation for $d=1$ systems. As discussed above, the scale factor $F$ for the SHO is about 2 for optimum efficiency. The extension of the DOMC method to anisotropic systems in two and three dimensions is described in the Appendix. Trial moves are made within an ellipsoid (an ellipse for $d=2$) that reflects the local anisotropy. The advantage of choosing moves from an ellipsoid was suggested by Northrup and McCammon [21] in 1980, although they had not developed a method for implementing it.

During each simulation cycle, the averages $[(\Delta x)^2]$ and $[\Delta E]$ are computed locally for each variable, and a new value of the maximum step size is obtained for the next cycle [26]. This procedure is done only once every cycle, and takes a negligible amount of computer time—about 3% for the adenosine simulations discussed below, and even less for larger molecules.

Because both the ARM and the DOMC method use information from past configurations in determining the transition probabilities, they are not strictly Markovian, which raises the possibility that systematic errors might be generated. To test for systematic errors, we have used the DOMC method to calculate the energy of a SHO using very short cycles. In the extreme case of only 1 MC step/cycle, we do find a large systematic error of 42%. However, with even 2 MC steps/cycle the error drops to 6%, and for 3 MC steps/cycle it is about 1%. No systematic error was measurable for 4 or more MC steps/cycle. To obtain small statistical errors in the estimates of $\delta$, at least 10 MC steps/cycle are needed for one-dimensional moves, and 50 to 100 MC steps/cycle for three-dimensional moves. Consequently, we conclude that the systematic errors are negligible for practical applications.

We found DOMC to be extremely effective for a wide range of anharmonic systems including both symmetric and asymmetric double-well potentials. The results were qualitatively similar to those for the SHO, although the optimal values of $F$ tended to be higher—about 3 or 4. It was also interesting to note that the average displacements showed a broader maximum when plotted against $F$ than when plotted against $\delta$, which implies that the precise value of $F$ is even less critical for strongly anharmonic systems. Comparisons between the results of DOMC and direct numerical integrations again confirmed the absence of any measurable systematic error for a cycle length of more than 4 MC steps. Furthermore, DOMC easily achieves optimal efficiency for two- and three-dimensional models, with anisotropies of 1000 in the ratio of the coupling constants. Tests on two- and three-dimensional potentials with anharmonicity and even double minima have demonstrated that the DOMC equations remain remarkably efficient.

In practice, it is possible for the straight averages used in Eq. (9) to take either very large or negative values. This may arise, for example, when a trial move puts an atom very close to another. Generally, this is rather rare at normal or low temperatures. However, to account for such cases, the stability of the program is protected by providing a branch to an alternate updating of the step size based on the ARM for the particular cycle that has run into a problem. The program reverts to the DOMC equations on the following cycle.

Although the ARM and the DOMC method have been derived and discussed in terms of single-particle moves, this is not a real restriction. In fact, collective moves that are important in biomolecules may be easily optimized using this approach. A particularly important class of such moves involves global rotations of a group of atoms with respect to the rest of the molecule [27–29]. These

moves are described by angles of rotation, and the only modification needed is to restrict the maximum trial move to the range $[-\pi,\pi]$. The DOMC method can be especially useful when the rotations are highly correlated. Efficient MC moves involving groups of two or three rotations can be optimized with the equations derived in the Appendix.

## III. APPLICATIONS

As a simple example, we have applied these simulation methods to adenosine, one of the building blocks of DNA, using the united-atom force field developed by Kollman and co-workers [8]. The potental energy is given by

$$E=\sum_{\text{bonds}} K_b(b-b_{\text{eq}})^2+\sum_{\text{angles}} K_\theta(\theta-\theta_{\text{eq}})^2 + \sum_{\text{dihedral}} \frac{V_\varphi}{2}[1+\cos(n\varphi-\gamma)] + \sum_{\text{nonbonds}} \left[\frac{A_{ij}}{R_{ij}^{12}}-\frac{B_{ij}}{R_{ij}^{6}}\right]+\sum_{\text{H bonds}} \left[\frac{C_{ij}}{R_{ij}^{12}}-\frac{D_{ij}}{R_{ij}^{10}}\right] + \sum_{\substack{i,j\\ i<j}} \frac{q_i q_j}{\epsilon R_{ij}} . \tag{10}$$

The first two terms represent the bond-length and bond-angle strain energies. The dihedral energy is represented by a cosine function, where $\varphi$ is the dihedral angle, $n$ denotes the symmetry of the torsional barrier, $V_\varphi$ is the force constant, and $\gamma$ is the phase angle. The last three terms describe the "nonbonded" interactions between pairs of atoms that do not belong to the same bond or angle. A hydrogen bond is represented by a 10-12 potential, and a Lennard-Jones potential is used for the other nonbonded pair interactions. The last term is intended to represent both the direct electrostatic interaction and the screening effects of solvent. This term is not well determined for typical separations of a few angstroms, but we have followed the common practice of taking the dielectric "constant" to be $\epsilon(R_{ij})=R_{ij}$ [30].

In performing the simulations, trial MC moves included nine global rotations along with single-particle jumps. The possibility of including such global moves to accelerate the simulation is a great advantage of the Monte Carlo approach.

One measure of the DOMC efficiency is the energy-energy correlation time, which is 8.6 sweeps, indicating a rapid sampling of phase space. However, this quantity is not easily compared with the results of other methods for reasons that will become clear below. Therefore, we have used the time dependence of the approach of the average rms displacement $\langle(\Delta\mathbf{r})^2\rangle^{1/2}$ to its equilibrium value as a measure of efficiency [21]. This quantity is defined as

$$\langle(\Delta\mathbf{r})^2\rangle^{1/2}=\left\langle\frac{1}{N}\sum_{i=1}^{N}\{\mathbf{r}_i-\langle\mathbf{r}_i\rangle\}^2\right\rangle^{1/2}, \tag{11}$$

where $N$ is the number of atoms and $\mathbf{r}_i$ and $\langle\mathbf{r}_i\rangle$ are the position and the average position of atom $i$, respectively. It was computed by superposing all the stored structures to the initial equilibrated structure. The superposition was done following the procedure introduced by Kabsch [31] with a modification that assures that the matching of structures is done through a rotation and not an inversion [32].

We have performed conventional MC, MD, and DOMC simulations at room temperature (298 K) using the same amount of computer time to account for the fact that the full DOMC simulation took a factor of 3 more computer time per sweep due to the global rotations. The DOMC simulation consisted of 1800 cycles with 100 sweeps each. Every single-particle move and rotation was performed once each sweep. Snapshots of the molecule are taken every 10 sweeps. The computed average position of each atom $\langle\mathbf{r}_i\rangle$ after superposition is updated after each snapshot. The rms is found to converge to about 1 Å.

In the MD simulation, a Verlet leap-frog algorithm was used with a time step of $10^{-15}$ s, starting from the same equilibrated initial configuration as the DOMC simulation. We used the average potential energy from the DOMC simulation to initialize the velocities of the atoms. Velocities were initially generated from a Maxwellian distribution and a global adjustment was made to set the total linear and angular momenta to zero. The velocities were then rescaled to make the initial kinetic energy equal to the difference between the previously calculated total energy and the configuration's poten-

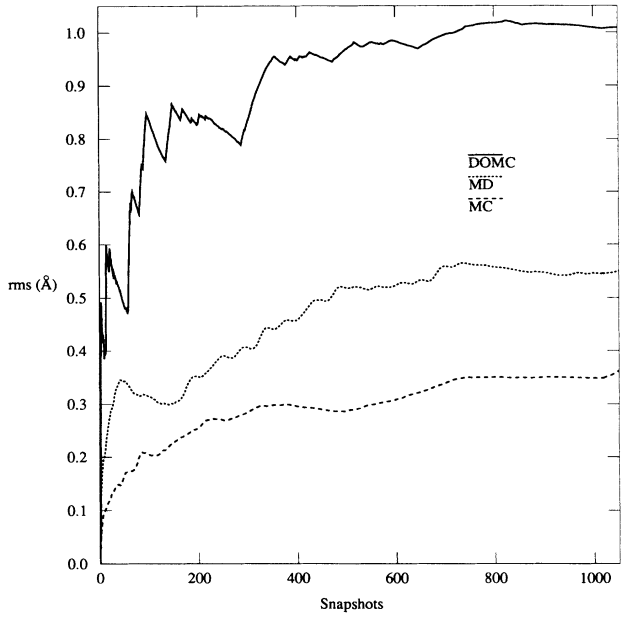


FIG. 7. The rms displacements for simulations of adenosine obtained from MC, MD, and DOMC simulations corresponding to the same amount of computer time. On a DEC 3100 workstation, the CPU times are 0.12 s/sweep for the DOMC simulation, and 0.04 s/sweep for the MC and MD simulations. For the DOMC simulation, snapshots are taken every 10 sweeps, while snapshots are taken every 30 sweeps for the MC and MD simulations. In this figure, the rms obtained during the first 1000 snapshots is shown.

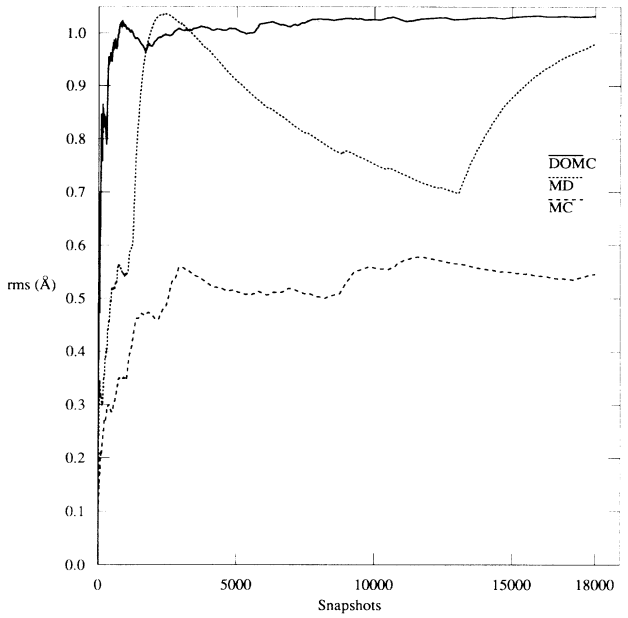


FIG. 8. Same as in Fig. 7, except that the length of the runs are longer. The DOMC simulation is 180 000 sweeps, while the MC and MD simulations are 540 000 sweeps long.

tial energy.

The standard MC simulation was done with a global optimizaton. Single atom moves were generated uniformly in volume from spherical neighborhoods with a radius of 0.08 Å, corresponding to the maximum rms displacement for this method. The overall global acceptance ratio was 31%.

Plots of rms displacements obtained for each of the three methods starting from the same well-equilibrated structure corresponding to the same amount of computer time are shown in Fig. 7 and 8. Figure 7 shows the "short-time" behavior of the rms fluctuations, while a lengthier run is shown in Fig. 8. From the short run, the rms fluctuations obtained from the DOMC simulation exceed the MD fluctuations, and approach the asymptotic value rapidly.

The plot of the DOMC rms fluctuations also shows some kinks (or zigzags). These represent energy-barrier crossings when the molecule changes conformation states. These kinks are also seen in the MD data in Fig. 8; however, they occur on much longer (and unpredictable) time scales.

The rms fluctuations obtained from MD depend strongly on the initial conditions—both on the initial configuration and the initial random Maxwellian velocities. They do not usually rise as fast as shown in Fig. 7. Our work indicates that for any initial configuration, the DOMC simulation converges much faster than either standard MC or MD simulation.

## IV. CONCLUSION

In this paper, we have introduced two methods for carrying out optimized Monte Carlo simulations of thermodynamic systems with strong inhomogeneity and local anisotropy. This approach is particularly intended for simulations of macromolecules, although we expect it to be useful in other situations. These methods make essential use of information gathered during the course of the simulation, which requires a slight relaxation of the usual restriction of Monte Carlo simulations to Markov processes. Our calculations have shown that under normal conditions, all systematic errors introduced by the non-Markovian nature of the simulation are negligible.

An important advantage of the current approach is the automatic optimization of any kind of MC move that would be useful in accelerating the convergence of the simulations. Large-scale collective motions can be emphasized, simulations can be carried out in either internal coordinate space or Cartesian space, or a mixture of both. This has far-reaching implications, especially in the calculation of free-energy differences by free-energy perturbation, multistage sampling, or umbrella sampling techniques, where lack of proper convergence can make the simulations very long and time consuming [33].

By providing for each incorporation of NOE and/or crystallographic constraints, the ARM and the DOMC method can be used for efficient structure refinement using simulated annealing techniques in either NMR or crystallographic studies. Such applications of the ARM and the DOMC method are already in progress.

We have demonstrated the efficiency of these methods by applying them to a small molecule that exhibits many of the characteristics that make simulations of larger molecules difficult. However, the structure of proteins presents certain problems that require specialized techniques beyond the scope of the present paper. Simulations of larger molecules, including progress on simulated annealing to find protein conformations, are planned to be discussed elsewhere.

## ACKNOWLEDGMENTS

We would like to acknowledge support by the National Science Foundation Grant No. DMR-9009475 (D.B. and R.H.S.) and NIH Grant No. GM25671 and the National Science Foundation Grant No. ASC-9015310 (S.K.). We also thank Charles Brooks, Stephen Garoff, John Nagle, and John Rosenberg for helpful comments and valuable discussions.

## APPENDIX: $d=2$ AND $d=3$ DOMC METHOD

To extend DOMC equations to higher dimensions, we first consider an effective anisotropic simple harmonic oscillator of the form

$$\mathcal{H}=\tfrac{1}{2}\sum_{i,j} k_{ij}x_i x_j \ ,$$

where $k_{ij}$ represents the spring constants. We use a transformation matrix $\mathbf{D}$ to generate moves $\{\eta_i\}$ in an ellipsoid (or an ellipse for $d=2$) given by

$$\eta_i=\sum_{j=1}^{3} D_{ij}\xi_j \ .$$

$\{\xi_j\}$ is a random vector chosen from a unit sphere (unit circle for $d=2$). The optimum jump size scales as the contours of constant energy and is determined by a dimensionless parameter $F$ given by

$$F^2=\tfrac{1}{2}\beta\mathbf{D}^t\cdot\mathbf{k}\cdot\mathbf{D} ,$$

where $\mathbf{D}^t$ is the transpose of the transformation matrix $\mathbf{D}$ and $\beta=1/k_BT$. $F$ is a scale factor chosen to optimize the efficiency of the simulation. The matrix $\mathbf{k}$ is determined from the simulation using

$$[\Delta E\eta_l\eta_m]=\tfrac{1}{2}\sum_{i,j}k_{ij}[\eta_i\eta_j\eta_l\eta_m] .$$

In this linear system of equations, $\Delta E$ denotes the change of energy for an attempted move $\{\eta_i\}$ and the square brackets indicate an average over all attempted moves, whether or not they were accepted [26]. The matrix $\mathbf{D}$ is then obtained from

$$D_{in}=F\left[\frac{2}{\beta\lambda_n}\right]^{1/2}V_{in} ,$$

where $\lambda_n$ is an eigenvalue of $\mathbf{k}$ and $\mathbf{V}$ is the corresponding normalized eigenvector. $\mathbf{D}$ is then updated every cycle to adapt to the changing local environment of each atom.

*Present address: Department of Biomedical Engineering, Boston University, Boston, MA 02215.

[1] C. L. Brooks III, M. Karplus, and B. M. Pettitt, *Proteins: A Theoretical Perspective of Dynamics, Structure, and Thermodynamics,* Advances in Chemical Physics Vol. LXXI (Wiley, New York, 1988).
[2] J. A. McCammon and S. C. Harvey, *Dynamics of Proteins and Nucleic Acids* (Cambridge University Press, New York, 1987).
[3] See, for example, M. Karplus and J. A. McCammon, Annu. Rev. Biochem. **52**, 263 (1983), and references therein.
[4] U. Burkert and N. L. Allinger, *Molecular Mechanics* (American Chemical Society, Washington, DC 1982).
[5] K. D. Gibson and H. A. Scheraga, Proc. Natl. Acad. Sci. U.S.A. **58**, 420 (1967).
[6] M. Levitt, J. Mol. Biol. **168**, 595 (1983).
[7] B. R. Brooks, R. Bruccoleri, B. Olafson, D. States, S. Swaminathan, and M. Karplus, J. Comput. Chem. **4**, 187 (1983).
[8] S. J. Weiner, P. A. Kollman, D. A. Case, U. C. Singh, C. Ghio, G. Alagona, S. Profeta, and P. Weiner, J. Am. Chem. Soc. **106**, 765 (1984).
[9] H. C. Andersen, J. Chem. Phys. **72**, 2384 (1980).
[10] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak, J. Chem. Phys. **81**, 3684 (1984).
[11] J. M. Haile and S. Gupta, J. Chem. Phys. **79**, 3067 (1983).
[12] L. V. Woodcock, Chem. Phys. Lett. **10**, 257 (1971).
[13] S. Nosé, Mol. Phys. **52**, 255 (1984).
[14] S. Nosé, J. Chem. Phys. **81**, 511 (1984).
[15] C. W. Gear, *Numerical Initial Value Problems in Ordinary Differential Equations* (Prentice-Hall, Englewood Cliffs, NJ, 1971).
[16] L. Verlet, Phys. Rev. **159**, 98 (1967).
[17] C. H. Bennett, J. Comput. Phys. **19**, 267 (1975).
[18] A. A. Markov, Izv. Adad. Nauk SPB **VI**, 61 (1907).
[19] S. Duane, A. D. Kennedy, B. J. Pendleto, and D. Roweth, Phys. Lett. B **195**, 216 (1987).
[20] D. W. Heermann, P. Nielaba, and M. Rovere, Comput. Phys. Commun. **60**, 311 (1990).
[21] S. H. Northrup and J. A. McCammon, Biopolymers **19**, 1001 (1980).
[22] S. L. Shumway and J. P. Sethna [Bull. Am. Phys. Soc. **35**, 500 (1990)] present a novel technique for optimizing MC moves using almost Markov simulations. Their specific technique is quite different from ours, but there are similarities between their approach and ours. See also S. L. Shumway and J. P. Sethna, Phys. Rev. Lett. **67**, 995 (1991).
[23] H. Müller-Krumbhaar and K. Binder, J. Stat. Phys. **8**, 1 (1973).
[24] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Clarendon, Oxford, 1987), p. 122.
[25] A. Corana, M. Marchesi, C. Martini, and S. Ridella, Assoc. Comput. Mach. Trans. Math. Soft. **13**, 262 (1987).
[26] Charles Brooks has pointed out that the straight averages of $[(\Delta x)^2]$ and $[\eta_i\eta_j\eta_l\eta_m]$ (see Appendix) can be calculated exactly by integrating over the distribution function from which the MC steps are sampled (private communication).
[27] D. C. Rapaport and H. A. Scheraga, Macromolecules **14**, 1238 (1981).
[28] T. Noguti and N. Go, Biopolymers **24**, 527 (1985).
[29] A. Kolinski, J. Skolnick, and R. Yaris, Biopolymers **26**, 937 (1987).
[30] A. Warshel and M. Levitt, J. Mol. Biol. **103**, 227 (1976).
[31] W. Kabsch, Acta Crystallogr. Sect. A **32**, 922 (1976).
[32] A. D. Machlan, J. Mol. Biol. **128**, 49 (1979). See Appendix.
[33] A. M. Ferrenberg and R. H. Swendsen, Phys. Rev. Lett. **63**, 1195 (1989), and references therein.