

# Stochastic molecular dynamics in systems with multiple time scales and memory friction

Mark E. Tuckerman<sup>a)</sup> and Bruce J. Berne

*Department of Chemistry, Columbia University, New York, New York 10027*

(Received 11 February 1991; accepted 11 June 1991)

The generalized Langevin equation (GLE) has been used to model a wide variety of systems in which a subset of the degrees of freedom move on a potential of mean force surface subject to fluctuating forces and dynamic friction. When there is a wide separation in the time scales for motion on the potential surface and for relaxation of the friction kernel, direct integration of the GLE is very costly in CPU time. In this paper we introduce an integrator based on our previous work using numerical analytical propagator algorithm (NAPA) and reference system propagator algorithm (RESPA) that greatly accelerates such simulations. We also discuss sampling methods for the random force. Accuracy of this algorithm is assessed by comparisons with an analytically solvable example. Introducing dynamic friction kernels determined from full molecular dynamics (MD) simulations allows us to compare the accuracy of the GLE simulations with full scale molecular dynamics simulations.

## I. INTRODUCTION

A wide variety of problems involving molecular motion in liquids can be formulated in terms of the generalized Langevin equation (GLE).<sup>1-3</sup>

$$\mu\ddot{q} = \frac{\partial W(q)}{\partial q} - \int_0^t d\tau \zeta(t-\tau)\dot{q}(\tau) + R(t), \quad (1.1)$$

where  $q$  and  $\dot{q}$  are, respectively, the position and momentum of a degree of freedom of interest,  $W(q)$  is the potential of mean force on this coordinate,  $\zeta(t)$  is the memory friction, and  $R(t)$  is the random force. The second fluctuation-dissipation theorem<sup>4</sup> gives a relationship between the friction kernel and the covariance function of the random force,

$$\zeta(t) = \beta \langle R(0)R(t) \rangle, \quad (1.2)$$

where  $\beta = (kT)^{-1}$ . To proceed it is necessary to know how the dynamic friction kernel varies with time. What is often needed is the friction kernel on a molecular bond or some more complicated reaction coordinate. Recently, we have devised a method for extracting  $\zeta(t)$  from molecular dynamics simulations and have used this method to determine the friction kernel on vibrational coordinates.<sup>5,6</sup> The explicit dynamic friction kernels obtained in this way can now be used in stochastic simulations based on Eq. (1.1). We and others have also devised methods for determining the potential of mean force  $W(q)$ . It is usually assumed that the random force is a Gaussian random process. If this assumption is made it is possible to integrate Eq. (1.1) and to simulate vibrational dephasing and energy transfer as well as activated barrier crossing among other things. Such a stochastic simulation scheme consists of two parts. First, one performs full molecular dynamics to obtain the dynamic friction kernel and the potential of mean force and then one integrates the stochastic integro-differential equations, Eq. (1.1).<sup>7,8</sup>

In vibrational relaxation it is usually the case that the oscillator has a high frequency compared to the frequencies characterizing the solvent motion. It is then necessary to

treat one of the most pervasive problems in the molecular dynamics literature; that is, the problem of treating systems with high and low frequency motions. Recently we have devised simple new integrators NAPA (numerical analytical propagator algorithm) and RESPA (reference system propagator algorithm) to treat systems with multiple time scales or disparate frequencies numerically.<sup>9-11</sup> These integrators have allowed us to determine the dynamic friction  $\zeta(t)$  on a homonuclear diatomic bond from simulation.<sup>12</sup> For the first time it becomes possible to evaluate this bond friction and to assess the accuracy of various theoretical approximations for it. Equation (1.1) when applied to the vibrational relaxation of a very stiff oscillator will also have a separation of time scales because the oscillator frequency will be very large compared to the relaxation rate of  $\zeta(t)$ . In this paper we introduce a method for integrating Eq. (1.1) that incorporates NAPA. For Gaussian Markov processes, the method is a direct one, useful for integrating the nonlinear GLE without resorting to Fourier transform. For non-Markovian processes, this direct sampling method is possible, but impractical, and one must resort to Fourier transform methods.<sup>13-15</sup>

The algorithm is tested on an analytically solvable system, namely, the GLE with exponential friction kernel.<sup>16</sup> The new method is then used to compare the predictions of the GLE with those of full scale molecular dynamics on a harmonic oscillator. It is found that molecular dynamics (MD) and the GLE give almost identical answers for the case considered here. The dynamical test system consists of a harmonic oscillator dissolved in Lennard-Jones (LJ) (6-12) fluid with which the diatomic interacts through a site-site LJ potential with the solvent atoms with the same parameters as used for the solvent-solvent potential. We report elsewhere on the use of these methods to study vibrational dephasing and energy relaxation.

## II. SAMPLING THE RANDOM FORCE

The random force in Eq. (1.1) is usually assumed to be a Gaussian random process. To proceed we make this assumption

<sup>a)</sup> Ph.D. student in the Department of Physics, Columbia University.

tion, recognizing that it may not be completely valid. In fact, by comparing the results of the stochastic simulation to full scale molecular dynamics we will be able to assess this approximation.

According to Doobs theorem<sup>17</sup> a Gaussian random process is Markovian if and only if the autocorrelation function of this process is an exponential function of the time. For a stationary Gaussian Markov process defined over a time  $T$  we can divide the time  $T$  into  $P$  discrete time steps  $\Delta t$ . In this case

$$P_{P+1}(R_P, P\Delta t; R_{P-1}, (P-1)\Delta t; \dots, R_0, 0) = \prod_{j=1}^P K_2(R_j, \Delta t | R_{j-1}, 0) P_1(R_0, 0), \quad (2.1)$$

where  $P_{P+1}$  is the joint probability distribution that the random force will have the sequence of values  $\{R_0, R_1, R_2, \dots, R_P\}$  at the corresponding times  $\{0, \Delta t, 2\Delta t, \dots, P\Delta t\}$ , where  $K_2(R_j, \Delta t | R_{j-1}, 0)$  is the conditional probability distribution that the random force will have the value  $R_j$  at time  $j\Delta t$  given that it had the value  $R_{j-1}$  at time  $(j-1)\Delta t$  and  $P_1(R_0, 0)$  is the probability distribution that the random force will have the value  $R_0$  at the initial instant  $t = 0$ .  $P_1$  and  $K_2$  have the form

$$P_1(R_0, 0) = [4\pi\sigma_0^2]^{-1/2} \exp[-R_0^2/2\sigma_0^2], \quad (2.2)$$

$$K_2(R_j, \Delta t | R_{j-1}, 0) = [4\pi\sigma(\Delta t)^2]^{-1/2} \times \exp[-(R_j - R_{j-1}\psi(\Delta t))^2/2\sigma^2(\Delta t)], \quad (2.3)$$

where

$$\psi(t) \equiv \frac{\langle R(0)R(t) \rangle}{\langle R(0)R(0) \rangle} = \frac{\zeta(t)}{\zeta(t=0)} \quad (2.4)$$

is the normalized correlation function of the random force,

$$\sigma_0^2 \equiv \langle R_0^2 \rangle = kT\zeta(0) \quad (2.5)$$

is the mean-square value of the random force and

$$\sigma^2(\Delta t) = \sigma_0^2(1 - \psi^2(\Delta t)). \quad (2.6)$$

Using Box-Müller sampling<sup>18</sup> we can sample  $R_0$  from  $P_1(R_0)$  and  $(R_1 - R_0\psi(\Delta t))$  from  $K_2(R_1, \Delta t | R_0, 0)$ . Knowing  $R_0$  and  $\psi(\Delta t)$  allows us to determine  $R_1$ . Likewise we can sample  $(R_2 - R_1\psi(\Delta t))$  from  $K_2(R_2, \Delta t | R_1, 0)$ . Knowing  $R_1$  and  $\psi(\Delta t)$  allows us then to determine  $R_2$ . Continuing this procedure we can sample  $\{R_1, \dots, R_P\}$ .

Once we have sampled the random force we must sample an initial state for the oscillator from the distribution function,

$$f(q, \dot{q}) = Q^{-1} e^{-\beta[\frac{1}{2}\mu\dot{q}^2 + W(q)]}, \quad (2.7)$$

where  $Q$  is the canonical partition function. Again  $\dot{q}$  can be sampled from the Maxwell distribution using Box-Müller sampling and  $q$  can be sampled from  $\exp[-W(q)/kT]$  using Monte Carlo sampling techniques.<sup>18</sup> In Sec. III we show how Eq. (1.1) can be integrated.

The foregoing applies only to a one-dimensional random force with exponential time correlation function. It is not difficult to generalize this to an  $n$ -dimensional random force with correlation function matrix

$$\zeta_{\alpha\beta}(t) = \langle R_\alpha(0)R_\beta(t) \rangle \quad (2.8)$$

which satisfies the conditions of Doobs theorem so that the multivariate set of random forces is a Gaussian Markov process;<sup>17</sup> that is,

$$\zeta_{\alpha\beta}(t + \tau) = \zeta_{\alpha\gamma}(t)\zeta_{\gamma\beta}(\tau). \quad (2.9)$$

This requires diagonalizing the multivariate Gaussians in each  $K_2$  and Box-Müller sampling the normal coordinates then transforming back to the original random forces.

As has been shown in numerous studies, the dynamic friction calculated from molecular dynamics simulations of real liquids is nonexponential. This means that the random force is non-Markovian. It is possible to derive the joint probability distribution function for such a system. Its characteristic function has the form

$$W(q_n, t_n; \dots; q_0, t_0) = \exp[-q_i\psi(t_j - t_i)q_j], \quad (2.10)$$

where  $\psi(t)$  is the normalized autocorrelation function of the random force as before. Fourier inversion then gives the joint distribution function  $P_n$  of the random force. This is a multivariate Gaussian and involves the inverse of the correlation function matrix which couples the random forces at all times within the correlation time of the random force. One way to proceed with this is to transform to  $n$  normal coordinates, sample these using Box-Müller and then transform back. Unfortunately this will be very computationally intensive because this matrix has dimension  $n \times n$ , and  $n$  will be very large. Another method is to use Metropolis Monte Carlo sampling<sup>18</sup> of this distribution. This is perfectly realizable but rather than present this method here we shall adopt the method pioneered by Rice<sup>17,13,14</sup> and exploited by McCammon *et al.*<sup>15</sup> This involves the representation of the random force by a Fourier series.

If the total time is  $T = P\Delta t$  and  $R(t)$  is assumed to be periodic with period  $T$  we can express  $R(t)$  as

$$R(t) = \sum_{k=1}^P \{a_k \sin \omega_k t + b_k \cos \omega_k t\}, \quad (2.11)$$

where  $t = \{0, \Delta t, 2\Delta t, \dots, P\Delta t\}$  and where  $\omega_k = 2\pi k/P\Delta t$ . If  $R(t)$  is a Gaussian random process the  $\{a_k, b_k\}$  will also be Gaussian random processes.

$$P(\{a_k, b_k\}) = \prod_{k=1}^P [4\pi\sigma_k^2]^{1/2} \exp\left[-\sum_{k=1}^N \frac{(a_k^2 + b_k^2)}{2\sigma_k^2}\right], \quad (2.12)$$

where

$$\sigma_k^2 = \frac{1}{P} \sum_{j=0}^{P-1} \langle R(0)R(j\Delta t) \rangle \exp[-2\pi ijk/P]. \quad (2.13)$$

To proceed one samples a set of  $P a_k$ 's and  $P b_k$ 's from the above Gaussian distribution using the specified force autocorrelation function to determine the  $\sigma_k^2$ 's. This set can be used to construct a realization of the random force using Eq. (2.11) which is easily evaluated using fast-Fourier-transform (FFT) methods. Sampling many realizations of the random force in this way allows us to generate an ensemble of random force realizations and to apply these to the integration of the equations of motion.

The direct sampling method based on Eqs. (2.2) and (2.3) or the Rice method both give a full time history or

realization of the random force  $R(t)$ . Repeated application of these sampling methods give  $N$  independent samplings of the random force, thus giving an ensemble of random forces each consisting of  $P$  time steps. Let us denote the  $\alpha$ th realization by  $\{R_0^\alpha, R_1^\alpha, R_2^\alpha, \dots, R_P^\alpha\}$ . The autocorrelation function of the random force can now be calculated as a combined ensemble and time average

$$\langle R(0)R(t = n\Delta t) \rangle = \frac{1}{N} \sum_{\alpha=1}^N \frac{1}{(P-n+1)} \times \sum_{m=0}^P R_m^\alpha R_{m+n}^\alpha. \quad (2.14)$$

If  $N$  and  $P$  are large enough, we should recover

$$\langle R(0)R(t = n\Delta t) \rangle = kT\zeta(t = n\Delta t). \quad (2.15)$$

Note that Eq. (2.14) can be used to compute the autocorrelation function of any observable.

There are two extreme possibilities for carrying out this simulation. One long trajectory of length  $NP$  can be generated and the time averaged correlation functions can be determined or  $N$  short trajectories each of  $P$  time steps can be generated and the time correlation functions of the dynamical variables can be both time averaged and ensemble averaged over the  $N$  realizations of the random force. In the former case only one FFT is used but it contains  $NP$  Fourier components. In the latter case  $n$  FFTs are executed each with  $P$  coefficients. If the computer is a vector machine with large memory the method involving only one FFT would be faster but memory is often limited and the second method would then be preferable.

An alternative to the FFT method for obtaining a random force sequence has been recently developed by Harris *et al.*<sup>19</sup> We do not employ this method in any of our simulations, but we mention it here for completeness. The method models the random force as an autoregressive process of order  $P$  with a Gaussian white noise source. A set of  $P$  autoregressive parameters are obtained from the friction kernel by inverting a set of Yule-Walker equations. This process yields an approximation to the friction kernel which is a sum of  $P$  exponentials. Given this form of the friction kernel, a computationally efficient algorithm can be devised which incorporates NAPA or RESPA for integrating the GLE. A difficulty with which method lies in the fact that in order to obtain results accurate at long times, one needs to choose  $P$  very large. However, accurate results over short times can be obtained with a relatively small value of  $P$ . There are situations when the long time behavior of the friction kernel is very important. For example, the dephasing rate of a nonlinear oscillator is determined by the full time dependence of the friction and any error incurred will produce errors in predictions of this quantity.

### III. INTEGRATION OF THE GLE

In this section we describe an algorithm for integrating the GLE based on the velocity Verlet<sup>20,21</sup> scheme, a method close in spirit to that presented by Berkowitz *et al.*<sup>15</sup> This method can be easily adapted for use with NAPA<sup>9</sup> when the dynamics of high frequency oscillators is sought. Recall that the velocity Verlet algorithm for evolving the position  $q$  and

velocity  $\dot{q}$  is given by the following equations:

$$q_{n+1} = q_n + \Delta t \dot{q}_n + \frac{(\Delta t)^2}{2\mu} F_n, \quad (3.1)$$

$$\dot{q}_{n+1} = \dot{q}_n + \frac{\Delta t}{2\mu} (F_n + F_{n+1}), \quad (3.2)$$

where  $q_n = q(n\Delta t)$ , etc., and  $F_n$  is the force at the  $n$ th time step. The force is

$$F(t) = f(q) - \int_0^t d\tau \zeta(\tau) \dot{q}(t - \tau) + R(t), \quad (3.3)$$

where  $f(q) = -\partial W(q)/\partial q$  is the force due to the potential of mean force, and  $R(t)$  is the random force. Writing Eq. (3.3) in discrete notation gives

$$F_n = f(q_n) - \Delta t \sum_{k=0}^n w_k \zeta_k \dot{q}_{n-k} + R_n, \quad (3.4)$$

where  $w_k$  are a set of weights association with the numerical integration method for the memory integral (e.g.,  $w_0 = w_n = \frac{1}{2}$ ,  $w_i = 1$  otherwise, if the trapezoidal rule is used). Substitution of Eq. (3.4) into Eq. (3.1) gives a direct method for obtaining the positions. The velocity equation Eq. (3.2) requires  $F_{n+1}$  given by

$$F_{n+1} = f(q_{n+1}) - \Delta t \sum_{k=0}^{n+1} w_k \zeta_k \dot{q}_{n+1-k} + R_{n+1}, \quad (3.5)$$

which involves  $\dot{q}_{n+1}$  when  $k=0$ . We separate out this term by writing

$$F_{n+1} = F'_{n+1} - \Delta t w_0 \zeta_0 \dot{q}_{n+1}. \quad (3.6)$$

Substituting Eq. (3.6) and Eq. (3.4) into Eq. (3.2) and solving for  $\dot{q}_{n+1}$  gives the result

$$\dot{q}_{n+1} = \frac{\dot{q}_n + (\Delta t/2\mu)(F_n + F'_{n+1})}{1 + (\Delta t)^2 w_0 \zeta_0 / 2\mu}. \quad (3.7)$$

We now discuss the application of NAPA<sup>9</sup> to the integration of the GLE. The GLE reads

$$\mu \ddot{q} = f(q) - \int_0^t d\tau \zeta(\tau) \dot{q}(t - \tau) + R(t). \quad (3.8)$$

The trajectory  $q(t)$  is written in the form

$$q(t) = q_0(t) + \delta(t), \quad (3.9)$$

where the reference trajectory  $q_0(t)$  is chosen to satisfy

$$\mu \ddot{q}_0 = f(q_0). \quad (3.10)$$

This results in the following equation for  $\delta$

$$\mu \ddot{\delta} = f(q_0 + \delta) - f(q_0) - \int_0^t d\tau \zeta(\tau) [\dot{q}_0(t - \tau) + \dot{\delta}(t - \tau)] + R(t). \quad (3.11)$$

At the beginning of each time step, we choose a new set of initial conditions on  $q_0$  and  $\delta$  which are

$$q_0(n\Delta t) = q(n\Delta t), \quad \dot{q}_0(n\Delta t) = \dot{q}(n\Delta t), \quad (3.12)$$

and

$$\delta(n\Delta t) = \dot{\delta}(n\Delta t) = 0. \quad (3.13)$$

We note that this choice is only possible for Eq. (3.11) be-

cause the integral in it involves only the sum  $\dot{q}_0 + \delta$  for which we can simply substitute the previously calculated values of  $\dot{q}$ . The solution to Eq. (3.10) subject to the initial conditions Eq. (3.12) can be obtained analytically (e.g., as in the case of a harmonic reference system) or numerically by integrating Eq. (3.10) for  $m$  little time steps  $\delta t \equiv \Delta t / m$  to obtain  $q_{0,n+1} \equiv q_0((n+1)\Delta t)$  and  $\dot{q}_{0,n+1} \equiv \dot{q}_0((n+1)\Delta t)$ . These are then fed into Eq. (3.11) which is then solved numerically subject to the initial conditions Eq. (3.13). The

velocity Verlet scheme discussed above can be easily adapted for Eq. (3.11). Let us define

$$G_n = -\Delta t \sum_{k=0}^n w_k \zeta_k (\dot{q}_{0,n-k} + \delta_{n-k}) + R_n \quad (3.14)$$

and, as before, we write

$$G_{n+1} = G'_{n+1} - \Delta t w_0 \zeta_0 (\dot{q}_{0,n+1} + \delta_{n+1}) \quad (3.15)$$

separating out the  $k=0$  term. Then, solving for  $\delta_{n+1}$  gives the following velocity Verlet scheme for NAPA:

$$\begin{aligned} q_{n+1} &= q_{0,n+1} + \delta_{n+1} = q_{0,n+1} + \frac{(\Delta t)^2}{2\mu} G_n, \\ \dot{q}_{n+1} &= \dot{q}_{0,n+1} + \delta_{n+1} = \dot{q}_{0,n+1} + \frac{\frac{\Delta t}{2\mu} [f(q_{n+1}) - f(q_{0,n+1}) + G_n + G'_{n+1} - \Delta t w_0 \zeta_0 \dot{q}_{0,n+1}]}{\left(1 + \frac{(\Delta t)^2 w_0 \zeta_0}{2\mu}\right)}. \end{aligned} \quad (3.16)$$

The initial conditions are now reset so that

$$\begin{aligned} q_0((n+1)\Delta t) &= q((n+1)\Delta t), \\ \dot{q}_0((n+1)\Delta t) &= \dot{q}((n+1)\Delta t), \end{aligned} \quad (3.17)$$

and

$$\delta((n+1)\Delta t) = \dot{\delta}((n+1)\Delta t) = 0 \quad (3.18)$$

and the process is repeated. The resetting of the initial conditions on  $\delta$  and  $\dot{\delta}$  to 0 at every step prevents the force term  $f(q_0 + \delta) - f(q_0)$  from becoming too large, thus allowing the trajectory to be generated with a larger time step than could be used with ordinary velocity Verlet integration. As we shall see, for the GLE, even small increases in the time step yield large savings in CPU time. The reason for this is that with a larger time step, not only can shorter trajectories (i.e., smaller  $P$ ) be used, but fewer points are needed for the friction kernel so that the time needed to evaluate the sums in Eqs. (3.4) and (3.14) is reduced.

The choice of reference system Eq. (3.10) will be good if, for example,  $f(q)$  represents an oscillator with frequency large compared to  $\sqrt{\zeta(0)/\mu}$ . However, if  $\sqrt{\zeta(0)/\mu}$  is comparable to or much larger than the frequency of the oscillator, another choice of reference system proves useful. Let  $q_0(t)$  satisfy

$$\mu \ddot{q}_0 = f(q_0) - \zeta(0) q_0. \quad (3.19)$$

This case has been referred to in the literature as the dynamic caging regime.<sup>3,22-25</sup> With this choice of reference system, the equation for  $\delta$  becomes

$$\begin{aligned} \mu \ddot{\delta} &= f(q_0 + \delta) - f(q_0) + \zeta(0) q_0 \\ &\quad - \int_0^t d\tau \zeta(\tau) [\dot{q}_0(t-\tau) + \dot{\delta}(t-\tau)] + R(t). \end{aligned} \quad (3.20)$$

The integration of these equations proceeds exactly as described above. The advantage of this reference system can be seen immediately for the case that  $f(q)$  has a harmonic term  $-\mu\omega^2 q$ . Then the reference system equation will have a harmonic term with an effective frequency

$$\Omega^2 = \omega^2 + \frac{\zeta(0)}{\mu}. \quad (3.21)$$

If  $\zeta(0)/\mu \gg \omega^2$ , then the vibrational frequency is determined primarily by  $\zeta(0)/\mu$  and is treated explicitly in the reference system.

#### IV. DYNAMIC FRICTION WITH EXPONENTIAL DECAY

As a test of the algorithm, we consider the stochastic dynamics of a particle in a harmonic well with exponential friction. The potential of mean force is then taken to be

$$W(q) = \frac{1}{2} \mu \tilde{\omega}^2 q^2, \quad (4.1)$$

where  $\mu$  is the reduced mass, and  $\tilde{\omega}$  is the solvent shifted vibrational frequency. The friction is taken to be of the form<sup>16</sup>

$$\zeta(t) = A e^{-\alpha t}. \quad (4.2)$$

With a harmonic potential and exponential friction kernel, all the correlation functions can be obtained analytically and compared with the simulation results. In particular, we study the autocorrelation functions of velocity, position, energy, and the random force.

The parameters used in this study are  $A = 406$ ,  $\alpha = 20.3$ ,  $\tilde{\omega} = 60$ ,  $\mu = 0.5$ , and the temperature is taken to be  $T = 2.5$ . These parameters are chosen to parallel those used in the molecular dynamics simulation which we discuss in Sec. V, and hence we express these quantities in Lennard-Jones reduced units. The values of  $A$  and  $\alpha$  were chosen so that the exponential decay is qualitatively similar to the decay of the true friction obtained from MD. We integrate the GLE using the velocity Verlet scheme described in the Sec. III both with and without NAPA to compare the CPU times. When NAPA is used, since the potential Eq. (4.1) is harmonic, the reference equation (3.10) becomes

$$\ddot{q}_0 = -\tilde{\omega}^2 q_0, \quad (4.3)$$

which when solved subject to the initial conditions Eq. (3.12) gives the reference trajectory

$$q_0(t) = q(n\Delta t)\cos \tilde{\omega}t + \frac{\dot{q}(n\Delta t)}{\tilde{\omega}} \sin \tilde{\omega}t \quad (4.4)$$

for  $n\Delta t \leq t \leq (n+1)\Delta t$ . In this comparative study, the random force is sampled directly using Eqs. (2.2) and (2.3). Since energy is not conserved in a stochastic simulation, another criterion is needed for determining a good time step. For exponential friction, we may use a direct comparison with the analytical results. However, in general, one may use the size of the fluctuations in the average kinetic or potential energies as a measure of the accuracy associated with a given time step. To obtain the same accuracy in the simulations, correlation functions are computed using Eq. (2.14) with  $P = 16384$ ,  $\Delta t = 5 \times 10^{-4}$  and 1400 friction points [ $\zeta_k$ 's as in Eq. (3.4)] without NAPA and  $P = 8192$ ,  $\Delta t = 1 \times 10^{-3}$  and 700 friction points [cf. Eq. (3.14)] with NAPA. In both cases,  $N = 5000$ . We find that NAPA gives a factor of 4 saving in CPU time, even though only a factor of 2 is gained in the time step. When higher frequency oscillators are involved, we have seen that NAPA will allow a larger time step factor over Verlet,<sup>9</sup> and we expect that larger time step improvements will have even more dramatic savings in CPU time.

At this point, we introduce a further time saving feature related to the evaluation of the memory integral. In principle, one should use the entire velocity history in evaluating the memory integral in Eq. (3.8) at time  $t$ . Often, however, the friction kernel is a rapidly decaying function of time, so that for large  $t$ , the early history of the trajectory will contribute negligibly to the integral. Therefore, in practice, it proves useful to choose some upper cutoff time  $t_{\max}$  above which the contributions to the integral are small, and write

$$\int_0^t d\tau \zeta(\tau) \dot{q}(t-\tau) \approx \int_0^{t_{\max}} d\tau \zeta(\tau) \dot{q}(t-\tau). \quad (4.5)$$

Obviously, the interval  $[0, t_{\max}]$  should be chosen large compared to the decay time of the friction. For the  $A$  and  $\alpha$  considered here, an interval of  $t_{\max} = 0.7$  is sufficient. When the friction decays rapidly, this approximation proves to be a significant time saver.

For comparison with analytical results, we integrate the GLE using the NAPA scheme with the same  $P$ ,  $N$ , etc. given above. The random force is sampled directly using Eqs. (2.2) and (2.3). The figures below compare the results of these simulations with the analytical results. Before proceeding, however, we note that the energy autocorrelation function is extremely difficult to obtain analytically, even for exponential friction. We have recently derived an expression for the energy autocorrelation function for a harmonic potential of mean force which relates it to other simpler correlation functions. The derivation will be presented in a forthcoming paper;<sup>26</sup> for now, we simply present the result. The normalized correlation function can be shown to be

$$\begin{aligned} C_{\delta E}(t) &\equiv \frac{\langle \delta E(0) \delta E(t) \rangle}{\langle \delta E^2 \rangle} \\ &= \frac{1}{2} C_{\dot{q}}^2(t) + \frac{1}{2} C_q^2(t) + \frac{1}{\tilde{\omega}^2} [\dot{C}_q(t)]^2, \end{aligned} \quad (4.6)$$

where  $C_{\dot{q}}(t)$  and  $C_q(t)$  are the normalized velocity and posi-

tion autocorrelation functions, respectively, and  $\delta E(t) \equiv E(t) - \langle E \rangle$ . The derivation of this expression is based solely on the GLE and the assumption that the random force is a Gaussian random process.

In Fig. 1(a), we show the comparison between the dynamic friction  $\zeta(t)$  used in the sampling function for the random force and the correlation function  $\langle R(0)R(t) \rangle / kT$  computed from the GLE simulation [cf. Eq. (1.2)]. The inset shows the difference between the simulation result and the true  $\zeta(t)$ . That is,

$$\Delta \zeta(t) \equiv |\zeta_{\text{GLE}}(t) - \zeta_{\text{true}}(t)|. \quad (4.7)$$

We see from the figure that the correlation function for the direct sampling method gives excellent agreement with the  $\zeta(t)$  used to sample the random force. In Fig. 1(b), we show the energy autocorrelation function for the GLE simulation plotted against the analytical result obtained from Eq. (4.6). Again, the inset shows the difference between the simulation result and the analytical form. Similarly, in Figs. 2(a) and 2(b), we show the autocorrelation functions of the velocity and position, respectively for the GLE simulation together with the corresponding analytical forms. The insets again show the differences. We also ran the same GLE simulation using the FFT method in Eqs. (2.11)–(2.13) to sample the random force to check the agreement between the two schemes, and we find that the agreement is very good. The close agreement between numerical and analytical results shows that the GLE can be accurately integrated using the simple method described in Sec. III.

To test the efficiency of the caging reference system Eq.

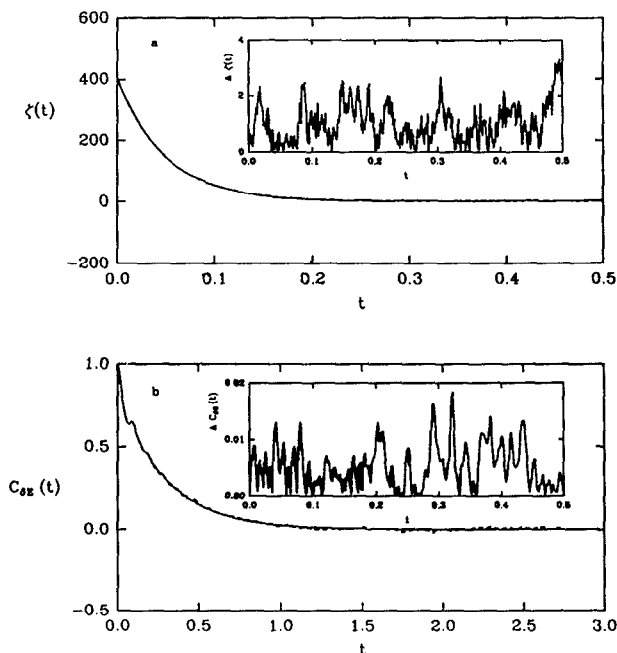


FIG. 1. (a) The autocorrelation function of the random force divided by  $kT$  (solid line) plotted against the true exponential friction (dashed line). The inset shows the absolute value of the difference between the autocorrelation function and the true friction according to Eq. (4.7). (b) The energy autocorrelation function computed from the GLE simulation (solid line) plotted against the analytical result (dashed line) [cf. Eq. (4.6)]. The inset shows the absolute value of the difference between the numerical result and the analytical result.

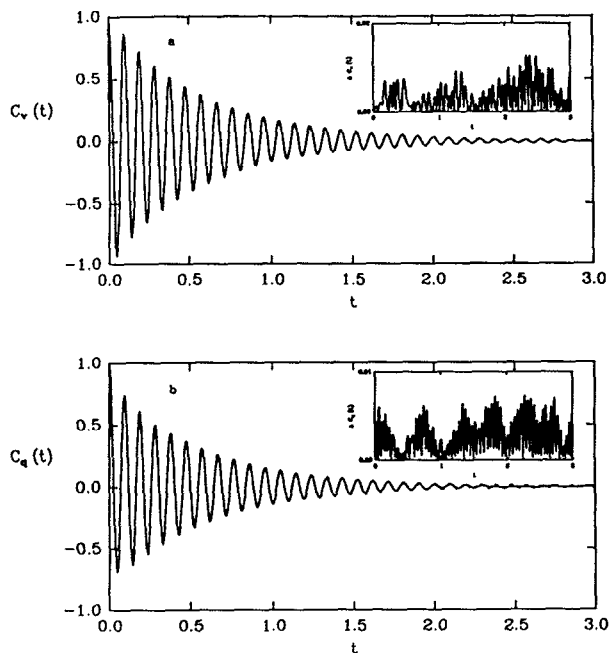


FIG. 2. (a) The autocorrelation function of velocity computed from the GLE simulation (solid line) plotted against the analytical result (dashed line). The inset shows the absolute value of the difference between the numerical result and the analytical result. (b) The position autocorrelation function computed from the GLE simulation (solid line) plotted against the analytical result (dashed line). The inset shows the absolute value of the difference between the numerical result and the analytical.

(3.19), we carry out a comparative study with  $\bar{\omega} = 20$ , and  $A = 1800$ . This gives an effective frequency of about  $\Omega = 63$  according to Eq. (3.21). The reference system trajectory for  $n\Delta t \leq t \leq (n+1)\Delta t$  is given by

$$q_0(t) = q(n\Delta t) \cos \Omega t + \frac{\dot{q}(n\Delta t)}{\Omega} \sin \Omega t. \quad (4.8)$$

Based on the case discussed above of  $\bar{\omega} = 60$  and  $A = 406$  which has roughly the same effective frequency, we expect the time step improvement to be a factor of 2 which is indeed what we find. This gives rise to a factor of 4 saving in CPU time. It is worth mentioning that the caging reference system will yield large savings only in cases of extremely large  $\zeta(0)/\mu$  which occur much more rarely in real systems than cases of extremely large  $\bar{\omega}$ .

To test the time step improvement with NAPA in an extreme situation, we carry out a similar comparative study using  $\bar{\omega} = 300$  and  $A = 406$ . Again, for the NAPA simulation, we use  $\Delta t = 1 \times 10^{-3}$  and 700 friction points. We run 500 trajectories each of length 40 000 time steps and find that the fluctuations in the average kinetic energy are around 5%. To obtain the same accuracy in the average kinetic energy without NAPA would require a time step of  $\Delta t = 3 \times 10^{-4}$ , 2333 friction points, and trajectories 133 333 time steps long. The result is a factor of 11 saved in CPU time using NAPA. In general, there are two aspects of the computation which primarily determine the CPU time required for the simulation: the number of time steps and the number of friction points in the memory integral. From the number of

time steps alone, we gain a factor of  $\Delta t(\text{NAPA})/\Delta t(\text{Verlet})$  improvement using NAPA over the ordinary Verlet simulation. Since the number of friction points used also goes as the ratio of time steps, we gain roughly another factor of  $\Delta t(\text{NAPA})/\Delta t(\text{Verlet})$  from the evaluation of the memory integral. These considerations lead us to conclude that the ratio of the CPU time for the Verlet simulation to the CPU time for the NAPA simulation will be

$$\frac{\text{CPU}(\text{Verlet})}{\text{CPU}(\text{NAPA})} = \left( \frac{\Delta t(\text{NAPA})}{\Delta t(\text{Verlet})} \right)^2. \quad (4.9)$$

This ratio may be reduced, however, by noting that if the decay of the friction kernel is slow compared to the time step, then not every friction point is needed to evaluate the memory integral in the Verlet simulation. In other words, the memory integral at the  $n$ th step may be well approximated by

$$\int_0^t d\tau \zeta(\tau) \dot{q}(t-\tau) \approx \sum_{k=0}^{n/\nu} w_{\nu k} \zeta_{\nu k} \dot{q}_{n-\nu k}, \quad (4.10)$$

which requires only every  $\nu$ th point. By choosing  $\nu = 3$  for this case, and using in addition the approximation in Eq. (4.5), the CPU time required to obtain the same accuracy without NAPA is cut by more than a factor of 2. The saving in CPU time with NAPA, however, is still a factor of 5 for this case which shows that NAPA can still yield substantial improvements in CPU time.

## V. DYNAMIC FRICTION FROM MOLECULAR DYNAMICS

The validity of the GLE approach to molecular dynamics can be tested by a direct comparison with an actual molecular dynamics simulation. For this comparison, we choose to carry out an MD simulation of a Lennard-Jones fluid in which is buried a LJ diatomic molecule with harmonic bond. The fluid plus diatomic comprises a system of 64 particles with temperature and density chosen to be 2.38 and 1.05, respectively, in reduced units. The harmonic bond potential is of the form

$$U(r) = \frac{1}{2} \mu \omega_0^2 (r - r_0)^2, \quad (5.1)$$

where  $r$  is the relative distance between the atoms,  $\mu$  is the reduced mass,  $\omega_0$  and  $r_0$  are the bare frequency and bond length, respectively. This system is similar to those discussed in Ref. 12. Since the bond frequency is higher than the peak of the spectral density of the fluid (around 20 for this temperature and density), the potential of mean force which goes into the GLE can be represented to a good approximation as a harmonic potential of the form

$$W(r) = \frac{1}{2} \mu \bar{\omega}^2 (r - \langle r \rangle)^2, \quad (5.2)$$

where  $\langle r \rangle$  is the equilibrium bond length averaged over a canonical ensemble, and  $\bar{\omega}$  is the renormalized frequency which can be computed from

$$\bar{\omega}^2 = \frac{kT}{\mu \langle (r - \langle r \rangle)^2 \rangle}. \quad (5.3)$$

The friction kernel is determined from the MD simulation using the method of Straub, Brokovec, and Berne,<sup>5</sup> which is based on the method of Harp and Berne.<sup>27</sup>

The MD simulation was carried out for  $3 \times 10^6$  steps

with a time step of  $1.75 \times 10^{-3}$ , using NAPA on the equation of motion for the bond length. As a simplification of the full dynamics, the rotational degrees of freedom of the diatomic are frozen out by fixing its orientation in space. This is done so that the NAPA reference equation can be solved analytically. The autocorrelation functions of the velocity, position and energy are computed for this run for comparison with the corresponding GLE simulation. Using Eq. (5.3), we calculate the renormalized frequency to be 59.6. To solve the GLE, the friction and renormalized frequency computed from the MD run are substituted into the equation, and the GLE is integrated using NAPA. The random force is sampled using the Fourier method described in Eq. (2.11) with  $P = 8192$ ,  $\Delta t = 1 \times 10^{-3}$  and 1000 friction points. A set of 500 initial conditions were sampled from the equilibrium distribution, and Eq. (2.14) was used to compute the autocorrelation functions in the GLE simulation.

In Figs. 3(a) and 3(b), we show the comparison of the autocorrelation functions of the velocity and position for the GLE and for MD. We see that the agreement is very close. The close agreement between MD and the GLE for this case is not unexpected since the potential is harmonic, and as we have seen from the difference between  $\omega_0$  and  $\tilde{\omega}$ , the frequency is not shifted much by the solvent. In Figs. 4(a) and 4(b), we show the autocorrelation functions of the random force and the energy, respectively. For consistency, it is necessary to use the energy of the bond in the form

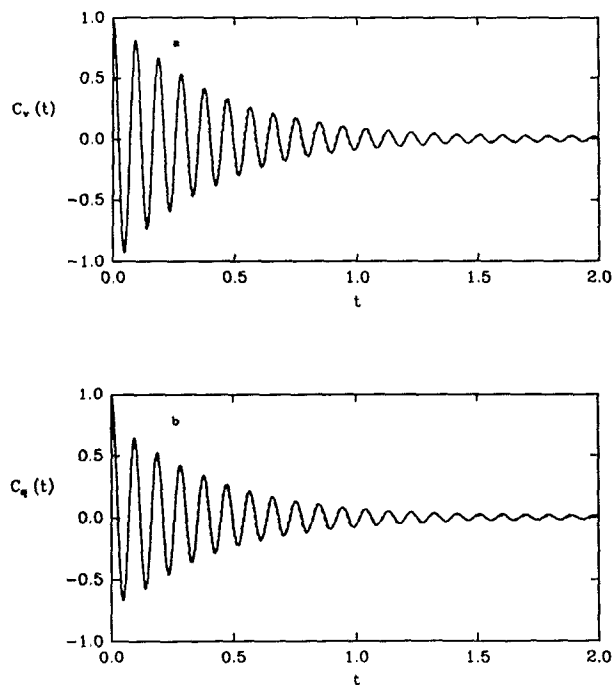


FIG. 3. (a) The autocorrelation function of velocity computed from the MD simulation (solid line) plotted against the GLE simulation result (dashed line). (b) The position autocorrelation function computed from the MD simulation (solid line) plotted against the GLE simulation result (dashed line).

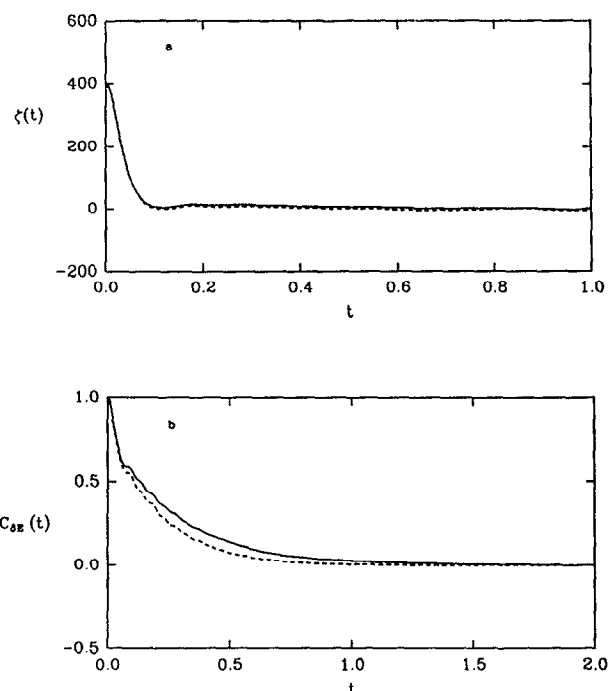


FIG. 4. (a) The true dynamic friction kernel computed from the MD simulation (solid line) plotted against the autocorrelation function of the random force divided by  $kT$  as computed in the GLE simulation (dashed line). (b) The energy autocorrelation function computed from the MD simulation (solid line) plotted against the GLE simulation result (dashed line).

$$E = \frac{p_r^2}{2\mu} + \frac{1}{2} \mu \tilde{\omega}^2 (r - \langle r \rangle)^2, \quad (5.4)$$

when computing the energy autocorrelation function for both the MD and GLE simulations. Thus, in the MD case, the renormalized frequency average bond length must be obtained before the energy can be computed. We see that the energy decay predicted by the GLE is slightly faster than the MD result. In this connection it is worth stating that the energy correlation function provides a test of the assumption that the random force is a Gaussian random process. In molecular systems the force can be resolved into a short range part and a long range part. We expect that the short range part will reflect energetic binary collisional events and will behave more like a Poisson process than a Gaussian process, whereas the long range part will be a superposition of a large number of terms representing the long range force from all of the neighboring atoms and will, by the central limit theorem, behave like a Gaussian random process. For a high frequency oscillator vibrational energy transfer to the bath will be through both the short range highly energetic collisions and the longer range force. To the extent that the short range impulsive collisions contribute we expect that the stochastic simulation of the energy relaxation will not agree with the full molecular dynamics. We will return to this question in a forthcoming publication devoted to the question of vibrational relaxation in liquids.

## VI. CONCLUSION

In problems where there is a separation of time scales, the methodology embodied in NAPA and RESPA can considerably accelerate simulations. In this paper, we have shown how this can be done for stochastic simulations using GLE's. The systems considered here involve high frequency vibrational motion damped by slowly decaying dynamic friction. Depending on the separation in the time scales, one can achieve accelerations on the order of 5 to 10 or more. Although not treated here, there is a class of problems for which NAPA and RESPA algorithms can lead to very large accelerations. Consider the case of an aqueous dispersion of polyballs (charged colloidal particles like polystyrene latex spheres). Each sphere will experience a frictional force due to the solvent plus long range forces due to Coulomb interactions with the other polyballs and due to hydrodynamic interactions (e.g., Oseen tensor). In such cases, one does not usually have a separation of time scales in the sense discussed above, but one must compute many forces because of the long range interactions. Nevertheless, the set of LEs or GLEs can be integrated efficiently using the RESPA methods outlined in Ref. 11. If there is also a separation in time scales, then a double RESPA or combined RESPA/NAPA method can be used. Recent computations using this approach to treat full MD simulations have resulted in accelerations approaching a factor of 50.<sup>28</sup>

## ACKNOWLEDGMENT

This work was supported by a grant from the National Science Foundation and from the Petroleum Research Fund, administered by the American Chemical Society.

<sup>1</sup> B. J. Berne and R. Pecora, *Dynamic Light Scattering* (Wiley-Interscience, New York, 1976).

- <sup>2</sup> P. Hänggi, P. Talkner, and M. Borkovec, *Rev. Mod. Phys.* **62**, 250 (1990).
- <sup>3</sup> J. T. Hynes, *Theory of Chemical Reaction Dynamics*, edited by M. Baer (Chemical Rubber, Boca Raton, 1985), p. 171.
- <sup>4</sup> H. Mori, *Prog. Theor. Phys.* **33**, 423 (1965).
- <sup>5</sup> J. E. Straub, M. Borkovec, and B. J. Berne, *J. Phys. Chem.* **91**, 4995 (1987).
- <sup>6</sup> J. Straub, B. J. Berne, and Benoit Roux, *J. Chem. Phys.* **93**, 6804 (1990).
- <sup>7</sup> Z. Schuss, *Theory and Applications of Stochastic Differential Equations* (Wiley, New York, 1980).
- <sup>8</sup> L. M. Delves and J. L. Mohamed, *Computational Methods for Integral Equations* (Cambridge University, New York, 1985).
- <sup>9</sup> M. Tuckerman, G. Martyna, and B. J. Berne, *J. Chem. Phys.* **93**, 1287 (1990).
- <sup>10</sup> M. Tuckerman, B. J. Berne, and A. Rossi, *J. Chem. Phys.* **94**, 1465 (1991).
- <sup>11</sup> M. Tuckerman, G. Martyna, and B. J. Berne, *J. Chem. Phys.* **94**, 6811 (1991).
- <sup>12</sup> B. J. Berne, M. E. Tuckerman, John E. Straub, and A. L. R. Bug, *J. Chem. Phys.* **93**, 5084 (1990).
- <sup>13</sup> S. O. Rice, *Bell. Tel. J.* **23**, 282 (1944).
- <sup>14</sup> S. O. Rice, *Bell. Tel. J.* **25**, 46 (1945).
- <sup>15</sup> M. Berkowitz, J. D. Morgan, and J. A. McCammon, *J. Chem. Phys.* **78**, 3256 (1983).
- <sup>16</sup> B. J. Berne, J. P. Boon, and S. A. Rice, *J. Chem. Phys.* **45**, 1086 (1966).
- <sup>17</sup> M. C. Wang and G. E. Uhlenbeck, *Rev. Mod. Phys.* **17**, 323 (1945).
- <sup>18</sup> M. H. Kalos and P. A. Whitlock, *Monte Carlo Methods* (Wiley, New York, 1986).
- <sup>19</sup> D. E. Smith and C. B. Harris, *J. Chem. Phys.* **92**, 1304 (1990).
- <sup>20</sup> W. C. Swope, H. C. Andersen, P. H. Berens, and K. R. Wilson, *J. Chem. Phys.* **72**, 4350 (1980).
- <sup>21</sup> M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Oxford University, Oxford, 1989).
- <sup>22</sup> J. E. Straub, M. Borkovec, and B. J. Berne, *J. Chem. Phys.* **84**, 1788 (1986); Addendum and Erratum, *ibid.* **86**, 1079 (1986).
- <sup>23</sup> J. P. Bergsma, B. J. Gertner, K. R. Wilson, and J. T. Hynes, *J. Chem. Phys.* **86**, 1356 (1987).
- <sup>24</sup> B. J. Gertner, J. P. Bergsma, K. R. Wilson, S. Lee, and J. T. Hynes, *J. Chem. Phys.* **86**, 1377 (1987).
- <sup>25</sup> B. J. Gertner, K. R. Wilson, and J. T. Hynes, *J. Chem. Phys.* **90**, 3537 (1989).
- <sup>26</sup> M. Tuckerman and B. J. Berne (unpublished).
- <sup>27</sup> B. J. Berne and G. D. Harp, *Adv. Chem. Phys.* **17**, 63 (1970).
- <sup>28</sup> M. Tuckerman and B. J. Berne (unpublished comment).