The notion of error in Langevin dynamics. I. Linear analysis

Bimal Mishra and Tamar Schlick[a)]
Department of Chemistry and the Courant Institute of Mathematical Sciences, New York University, New York, New York 10012

(Received 28 September 1995; accepted 18 March 1996)

The notion of error in practical molecular and Langevin dynamics simulations of large biomolecules is far from understood because of the relatively large value of the timestep used, the short simulation length, and the low-order methods employed. We begin to examine this issue with respect to equilibrium and dynamic time-correlation functions by analyzing the behavior of selected implicit and explicit finite-difference algorithms for the Langevin equation. We derive: local stability criteria for these integrators; analytical expressions for the averages of the potential, kinetic, and total energy; and various limiting cases (e.g., timestep and damping constant approaching zero), for a system of coupled harmonic oscillators. These results are then compared to the corresponding exact solutions for the continuous problem, and their implications to molecular dynamics simulations are discussed. New concepts of practical and theoretical importance are introduced: scheme-dependent perturbative damping and perturbative frequency functions. Interesting differences in the asymptotic behavior among the algorithms become apparent through this analysis, and two symplectic algorithms, ''LIM2'' (implicit) and ''BBK'' (explicit), appear most promising on theoretical grounds. One result of theoretical interest is that for the Langevin/implicit-Euler algorithm (''LI'') there exist timesteps for which there is neither numerical damping nor shift in frequency for a harmonic oscillator. However, this idea is not practical for more complex systems because these special timesteps can account only for one frequency of the system, and a large damping constant is required. We therefore devise a more practical, delay-function approach to remove the artificial damping and frequency perturbation from LI. Indeed, a simple MD implementation for a system of coupled harmonic oscillators demonstrates very satisfactory results in comparison with the velocity-Verlet scheme. We also define a probability measure to estimate individual trajectory error. This framework might be useful in practice for estimating rare events, such as barrier crossing. To illustrate, this concept is applied to a transition-rate calculation, and transmission coefficients for the five schemes are derived. [S0021-9606(96)50224-2]

I. INTRODUCTION

Molecular dynamics (MD) simulations have become a powerful tool for analyzing the properties of many molecular systems. Besides providing insight into kinetic pathways, such simulations can often sample the energetically accessible configurations of a system more efficiently in comparison to the commonly used Metropolis/Monte Carlo algorithms. From these configurational ensembles, various transport coefficients and dynamic correlation functions can be calculated.

Often, the equations of motion are modified by coupling the system to an additional degree of freedom[1,2] or to a heat bath[3] in the form of Langevin equation. This makes possible the study of a system's dynamics in a suitable statistical framework. Such a modified framework of the governing equations of motion can also serve a different purpose, such as overcoming the inherent instability (e.g., drifts in energy) of microcanonical MD algorithms due to the truncation errors induced by a nonzero timestep.[4] This problem is particularly severe for systems with long relaxation times. Applications of MD today extend from liquid and small molecular systems that began over 50 years ago[5] to proteins in solution and complex membrane systems.

In canonical ensembles, the target system is coupled to a surrounding heat bath with which the system is free to exchange energy, thereby maintaining a state of equilibrium. In the Langevin dynamics formalism, the explicit solvent degrees of freedom are eliminated from the nonequilibrium thermodynamic description with the help of Mori–Zwanzig projection operator technique.[6,7] The result is a set of stochastic differential equations describing the dynamic state of the solute. For example, for a molecular system whose phase-space coordinates are $x(t)$ and $p(t)$, the Langevin equation (in its simplest form) can be written as

$$\begin{aligned} \frac{dx}{dt} &= M^{-1}p, \\ \frac{dp}{dt} &= -g_E(x) - \gamma p + r(t), \end{aligned} \tag{1}$$

where M is the diagonal mass matrix, $g_E(x)$ is the gradient of the potential energy $E(x)$, γ is the damping constant, and $r(t)$ is the random force vector.

The Langevin equation, like most coupled nonlinear differential equations, must be solved numerically. There has certainly been considerable mathematical analysis on nu-

[a)]Howard Hughes Medical Institute. Author to whom all correspondence should be addressed.

merical solution of the Langevin equation[8–10] with regards to accuracy and stability, but the results are often of little practical use to simulators of biomolecular dynamics. This is because the available mathematical theory focuses on the limit $\Delta t \to 0$ (where Δt is the timestep), whereas large-scale MD simulations today use rather large timesteps with respect to the fastest period of the system. (e.g., one ninth to one quarter). Furthermore, such simulations are rather short in total time with respect to relaxation times of the slowest modes, and they use low-order algorithms (e.g., 2).

Clearly, in stochastic dynamics, there is no unique trajectory but rather an ensemble of trajectories. Even in MD ($\gamma=0$), different trajectories result from using: various algorithms, same algorithms but different timesteps, different starting points,[11] and different pseudorandom-number generators or seeds. An important question that arises is how to evaluate a generated trajectory. This question has long been realized in the chemical dynamics community of small systems.[12] For biomolecules, in the absence of direct experimental data for comparison, the convention to date has been to compare results to a simulation at a small timestep. But is this comparison adequate? How can it be supplemented? Are there rigorous mathematical tests that can be performed? These are difficult questions, and it is clear that the answers must be based on statistical considerations. However, both ways to evaluate static means as well as dynamic functions must be constructed.

In Langevin dynamics simulations, we have two major sources of error: systematic and random. The first arises from the limited accuracy of the integration algorithm (discretization error), the finite precision of the computer (roundoff error), the errors in the energy and force evaluations (e.g., approximations to trigonometric expressions, truncated multipole expansions), and so on. The ''random'' error component comes from the finite length of the trajectory (i.e., finite ensemble averages), and many other aspects of the calculation that can lead to spurious results (e.g., different equilibration procedures, finite length of the pseudorandom number generator).

Many of the above considerations must be tested heuristically, e.g., by running a simulation five times longer and varying parameters, conditions, and protocols. While these requests may seem trivial, the high cost of biomolecular simulations in terms of computer time (i.e., months to generate a nanosecond trajectory of a protein in solution) has limited these checks in practice.

The problem of trajectory assessment arises in this context because the nonequilibrium processes simulated in biomolecular MD are associated with large, chaotic systems; there are both *deterministic* and *stochastic* features; simulation time is relatively *short*; the force fields are *approximate*; and experimental data are *limited*. Furthermore, we are interested in both local (detailed kinetics) and global (sampling) features. Therefore, one can imagine that different models, in combination with different integration or propagation methods, could be designed to address different aspects of dynamics problems for macromolecules. Using dynamic simulations for statistical averaging in phase space is certainly appropriate, though special care must be exercised for highly correlated data. When detailed kinetics, such as transition pathways and rates are of interest, global aspects (i.e., ensemble properties in the framework of statistical mechanics) are also required.

How can we know whether a simulated trajectory is ''representative'' in some sense? Elofsson and Nilsson[13] asked how ''consistent'' MD simulations are by comparing 30 protein simulations differing in solvent representation and protocols; they found great sensitivity of overall fluctuations to the starting structure and suggested that several shorter simulations span conformation space better than one long one. More recently, Auffinger *et al.*[11] demonstrated the divergence of ten 100 ps trajectories of tRNA in solvent and salt—from the initial x-ray structure as well as from one another—when initial conditions and parameters were varied. Consistency problems emerged (e.g., results could be *worse*, surprisingly, when equilibration time was *extended*.) In addition, the authors emphasized the *inadequacy* of energy conservation and root-mean-square fluctuations (from one available experimental structure) alone; they suggested that multiple MD simulations be generated to evaluate the consistency of results in general. While it is likely that the specific problems above are aggravated by the complexity of electrostatics in nucleic acids, they are still typical of biomolecular dynamics, as practitioners well know. Adequate evaluation of biomolecular simulations will undoubtedly increase in urgency in the coming years as longer simulations of larger systems will be possible. Commensurate refinement of algorithmic and simulation protocols is expected.

To begin an investigation of the notion of error, we present a comparative study of selected finite-differencing algorithms for Langevin dynamics simulations. The basic themes we explore are twofold: the effect of different integrators on the physical properties of interest, and the general notion of error in this stochastic framework. The concept of error in finite dynamic simulations using various algorithms is particularly important, especially when one is interested in long-time processes of complex systems. Different algorithms can be used for different purposes, and obvious evaluation criteria are not available.[14]

Over the last 20 years, a variety of integration algorithms have been proposed and their relative merits and accuracies discussed.[15–19] However, much of these analyses focused on explicit methods, since they are easier to implement for complex nonlinear forces, as well as on local error only. It is well known that explicit methods are stable only for small timesteps; they become unstable at some critical value of Δt and are thus problematic for systems with multiple timescales.[20] We emphasize, though, that ''small'' in our context of explicit integrators is rather large in mathematical analysis.

Implicit algorithms, on the other hand, tend to be mostly unconditionally stable (A-stable).[20] Roughly speaking, this means that there are no stability restrictions on the model problem $y' = \lambda y$ (whose solution is $Ae^{\lambda t}$) as long as $\mathrm{Re}(\lambda) < 0$ for nonnegative timesteps. Implicit algorithms can, however, introduce numerical damping[21] and are computa-

tionally very demanding.[19] Symplectic methods have recently gained favorable attention for their good performance, but resonance problems have been noted.[22,23] Thus, in order to adequately apply a particular algorithm to a specific problem at hand, a comparative study of local stability, accuracy, and statistical behavior is required for the range of timesteps used in molecular and Langevin simulations. This paper begins an examination in this direction.

For illustration of these concepts, we consider five finite-differencing algorithms. Three of them have been discussed by Zhang and Schlick in detail.[19] ''LI'', for Langevin/implicit-Euler; and ''LIM2'' and ''MID1'' which are semi-implicit. The other two are explicit: ''BBK'', a Verlet-like algorithm presented by Brooks, Brünger, and Karplus;[18] and an explicit-Euler version, ''LE''; their statistical properties and various limiting cases have been analyzed by Pastor *et al.*[24] The analysis offered here extends to new questions and also presents a unified approach for examining many properties of interest, particularly for MD applications. Extensions to other algorithms and more complex potentials are natural in this framework.

In Sec. II, we discuss the issue of stability and accuracy by deriving general solution expressions in the linear force case. We define two concepts: *perturbed frequency* and *perturbed damping*, both of which are scheme dependent. In particular, for algorithms where numerical damping is important, we show that a choice of timestep, as a function of a frequency and a damping constant, can be made (at least for the case of quadratic potentials) to yield no numerical damping as well as no shift in frequency. This formula is possible only for LI, but has limited practical use because of the large damping constant required, and the restriction to a single frequency. Therefore, we also propose a ''delay-function'' approach to remove artificial damping and frequency perturbation from LI which also works for MD ($\gamma=0$). Application to a system of coupled oscillators demonstrates very satisfactory results in comparison to velocity Verlet.

In Sec. III, the concept of stationary states is introduced, and we derive for each algorithm selected averages for a harmonic oscillator system. In Sec. IV, the diffusion constant is calculated for each algorithm as a limiting case (zero frequency). The possible ways of estimating a trajectory error are discussed in Sec. V by introducing a probability measure. For an illustration, this measure is applied to rate-constant calculations in barrier crossing events. We conclude with a brief summary of the findings and discussion of future work.

II. ANALYSIS OF STABILITY AND ACCURACY

When applying numerical integrators to a nonlinear set of equations, it is important to know their region of stability. Stability of numerical methods is strongly related to the ''stiffness'' of the problem. Intuitively, this characteristic means that the system's dynamics is governed by at least two processes that occur on varying time scales. The existence of noise in the stochastic case complicates matters further (see, for example, Fig. 1 in Ref. 25), since it adds other physical and numerical considerations not present in the corresponding nonstochastic differential equation. In practice, the analysis of long-time behavior of numerical methods for initial value problems starts with the study of a linear reference problem. This is because, unfortunately, there is no general global method for analysis of a nonlinear system. However, many insights can already be gained from the linear analysis.

A. The Langevin equation for coupled harmonic oscillators

Let us consider the Langevin equation for the position x:

$$M\ddot{x}+g_E(x)+\gamma M\dot{x}=r(t), \tag{2}$$

where the dot superscripts denote differentiation with respect to time. The random force $r(t)$ is a stationary, Gaussian white noise characterized by mean and covariance matrix, respectively,

$$\langle r(t)\rangle=0$$

and

$$\langle r(t)r(t_1)^T\rangle=2\gamma k_B T M\,\delta(t-t_1).$$

Here, T is the temperature of the heat bath, k_B is Boltzmann's constant, and δ is the Dirac delta function.

If the force is linearized, the governing potential has the bilinear form

$$M^{-1}g_E(x)=(\tfrac{1}{2}x^TAx)'=Ax,$$

where A is a symmetric positive definite $n\times n$ matrix, and the prime superscript denotes differentiation with respect to x. This is the case for a system of coupled harmonic oscillators.

B. Mass scaling and unitary transformation

Let us apply the following transformations to simplify the Langevin equation further:

$$x=M^{-1/2}Z,$$

$$G=M^{-1/2}r,$$

$$B=M^{-1/2}AM^{-1/2}.$$

The Langevin equation then becomes

$$\ddot{Z}(t)+\gamma\dot{Z}(t)+BZ(t)=G(t), \tag{3}$$

where $G(t)$ has the statistical properties:

$$\langle G(t)\rangle=0$$

and

$$\langle G(t)G(t_1)^T\rangle=2\gamma k_B T\delta(t-t_1).$$

We now proceed by applying unitary transformations. We can transform the matrix B into a diagonal matrix D on the basis of the orthogonal matrix $T(T^T=T^{-1})$, to obtain

$$D=TBT^{-1}.$$

Under this transformation, the coordinates and the random force become

$$Q=TZ$$

TABLE I. The five discretizations for Langevin dynamics. The first data column gives expression for $Q,\dot{Q},\ddot{Q}$ for each integration scheme. For the origin of these definitions, see Ref. 19 for LI, LIM2, and MID1; and Ref. 24 for BBK and LE. In the second data column, we list the corresponding values of μ, ν, and σ [see Eq. (10) of the text]. To simplify the expressions, we use the definitions $\delta=\gamma\Delta t$ and $\epsilon=\omega\Delta t$, where γ is the damping coefficient, ω is the frequency, and Δt is the timestep.

Algorithm	$Q,\ \dot{Q},\ \ddot{Q}$	$\mu,\ \nu,\ \sigma$
LI	Q^{n+1}	$1+\delta+e^2,$
	$[Q^n-Q^{n-1}]/\Delta t,$	$-2-\delta,$
	$[Q^{n+1}-2Q^n+Q^{n-1}]/\Delta t^2$	1
LIM2	$[Q^{n+1}+Q^{n-1}]/2,$	$1+\frac{\delta}{2}+\frac{\epsilon^2}{2},$
	$[Q^{n+1}-Q^{n-1}]/2\Delta t,$	$-2,$
	$[Q^{n+1}-2Q^n+Q^{n-1}]/\Delta t^2$	$1-\frac{\delta}{2}+\frac{\epsilon^2}{2}$
MID1	$[Q^{n+1}+Q^n]/2,$	$1+\frac{\delta}{2}+\frac{\epsilon^2}{4},$
	$[V^{n+1}+V^n]/2=[Q^{n+1}-Q^n]/\Delta t,$	$-2+\frac{\epsilon^2}{2},$
	$[V^{n+1}-V^n]/\Delta t$	$1-\frac{\delta}{2}+\frac{\epsilon^2}{4}$
BBK	$Q^n,$	$1+\frac{\delta}{2},$
	$[Q^{n+1}-Q^{n-1}]/2\Delta t,$	$-2+\epsilon^2,$
	$[Q^{n+1}-2Q^n+Q^{n-1}]/\Delta t^2$	$1-\frac{\delta}{2}$
LE	$Q^n,$	$1+\delta,$
	$[Q^{n+1}-Q^n]/\Delta t,$	$-2-\delta+\epsilon^2,$
	$[Q^{n+1}-2Q^n+Q^{n-1}]/\Delta t^2$	1

and

$$F=TG.$$

We have now converted Eq. (3) into a system of N uncoupled independent equations for each normal mode Q_j:

$$\ddot{Q}_j+\gamma\dot{Q}_j+\omega_j^2Q_j=F_j(t). \tag{4}$$

The corresponding random force characteristics are

$$\langle F_j\rangle=0$$

and

$$\langle F_j(t)F_k(t_1)\rangle=2\gamma k_BT\delta_{jk}\delta(t-t_1).$$

For notational simplicity, we drop the subscript j in future discussions.

C. Discretizations of the normal-mode equation

Consider now a numerical discretization of Eq. (4). We denote by Q^n the difference-equation approximation to Q at time $n\Delta t$. This approximation is integrator dependent. Therefore, we are interested in examining the different propagation patterns. For each of the algorithms considered, we define the variables Q, $\dot{Q}$, and $\ddot{Q}$ in a unified framework (Table I). See Appendix A for details regarding MID1.

Very briefly, ''LI'' is the first-order implicit-Euler algorithm applied to the Langevin equation; it exhibits numerical damping which is well understood,[19,26] but is unconditionally stable. ''LIM2'' and ''MID1'' are second-order *symplectic*,[27] semi-implicit algorithms. The implicit-midpoint algorithm, ''MID1'', in particular, was found to perform well at larger timesteps but exhibit resonance.[23] Verlet, other symplectic algorithms, and possibly other integrators exhibit resonance also.[23] ''BBK'' is a second-order, explicit algorithm that reduces to well known symplectic Verlet algorithm when $\gamma\to0$.[24] ''LE'' is first-order explicit-Euler algorithm, differing from ''BBK'' in the definition of the velocity.

The recursion relations for each algorithm are given below (we set $\delta=\gamma\Delta t$, not to be confused with delta function):

$$\text{LI:}\quad (1+\delta)x^{n+1}-(2+\delta)x^n+x^{n-1}=\Delta t^2M^{-1}[-g_E(x^{n+1})+r^n], \tag{5}$$

$$\text{LIM2:}\quad (1+\delta/2)x^{n+1}-2x^n+(1-\delta)x^{n-1}=\Delta t^2M^{-1}\left[-g_E\left(\frac{x^{n+1}+x^{n-1}}{2}\right)+r^n\right], \tag{6}$$

$$\text{MID1:}\quad \frac{x^{n+1}-x^{n-1}}{\Delta t}=\frac{v^{n+1}+v^{n-1}}{2},\qquad \frac{v^{n+1}-v^{n-1}}{\Delta t}=M^{-1}g_E\left(\frac{x^{n+1}+x^{n-1}}{2}\right)-\gamma\left(\frac{x^{n+1}-x^n}{\Delta t}\right)+r^n, \tag{7}$$

$$\text{BBK:}\quad (1+\delta/2)x^{n+1}-2x^n+(1-\delta/2)x^{n-1}=\Delta t^2M^{-1}[-g_E(x^n)+r^n], \tag{8}$$

$$\text{LE:}\quad (1+\delta)x^{n+1}-(2+\delta)x^n+x^{n-1}=\Delta t^2M^{-1}[-g_E(x^n)+r^n]. \tag{9}$$

Note that x^{n+1} appears in both sides of the equations for LI, LIM2, and MID1. This implicit relation can be solved by reformulating the solution of the nonlinear equation for x^{n+1} into a minimization subproblem; see details in Ref. 19. Note also that MID1 differs from the other algorithms in that the position and velocity are coupled by a matrix relation; for the other integrators, the propagation formulae involve positions only. Finally, note that all right-hand sides except for MID1 reduce to the same discretization for $\ddot{x}=(x^{n+1}-2x^n+x^{n-1})/\Delta t^2$ when $\gamma=0$. The systematic forces are evaluated at the previous point for the explicit algorithms (BBK and LE), at the new point for LI, and at midpoint for MID1 and LIM2.

The discretizations listed in Table I for the normalized position, velocity, and acceleration for each algorithm can be considered within a unified representation for the dynamics. Namely, by assuming a three-step method, either explicit or implicit, we write Eq. (4) with the exception of MID1 (see Appendix A) as

$$\mu Q^{n+1}+\nu Q^n+\sigma Q^{n-1}=\Delta t^2F^n. \tag{10}$$

Here μ, ν, and σ depend on the choice of finite-differencing scheme and are listed in Table I. Note that these coefficients

are functions of Δt, γ, and ω. The random force F^n above follows the Gaussian statistics with mean and variance, respectively, given by

$$\langle F^n \rangle = 0, \quad \langle F^n F^m \rangle = \frac{2\gamma kT\delta_{nm}}{\Delta t}.$$

The solution of Eq. (10) and derivation of long-time values for a free particle were already published by Pastor, Brooks, and Szabo (Ref. 24, appendix).

D. Solution of the discretized Langevin equation

Equation (10) is a linear, inhomogeneous difference equation with specified initial conditions Q^0 and Q^1. It can readily be solved by the method of variation of parameters.[28] Briefly, this method first determines solution of the homogeneous equation $L(y)=0$, where L is the difference operator. For example, for a second-order difference equation, the solution might look like $y=A_1\alpha_1^n+A_2\alpha_2^n$, where A_1 and A_2 are functions to be determined. The method of variation of parameters for $L(y)=f$ determines A_1' and A_2', after which A_1 and A_2 are obtained by integration. That system of equations is

$$A_1'\alpha_1+A_2'\alpha_2=0, \quad A_1'\alpha_1'+A_2'\alpha_2'=f.$$

The second equation is obtained by imposing the former condition, applying the operator L to y, and setting the result equal to f.

For the homogeneous analog of Eq. (10), we assume a solution of form $Q^n=\alpha^n$, where α is the function to be determined. We then obtain a quadratic equation in α,

$$\mu\alpha^2+\nu\alpha+\sigma=0, \quad \mu\neq 0. \tag{11}$$

The solution, α, has at most two roots:

$$\alpha_\pm=\frac{-\nu}{2\mu}\pm\frac{1}{2\mu}\sqrt{\nu^2-4\mu\sigma}. \tag{12}$$

Note that if we write α as a complex number $\alpha_R+i\alpha_I$, where $i=\sqrt{-1}$, we must know the sign of $\nu^2-4\mu\sigma$ for each scheme.[25]

The form of α is an important characteristic of each integration scheme because it determines how the solution evolves in time ($Q^n=A_1\alpha_+^n+A_2\alpha_-^n$). For example, for the homogeneous case, writing α as $\alpha_R+i\alpha_I$, we obtain $|\alpha^n|=(\alpha_R^2+\alpha_I^2)^{n/2}$. For the implicit-Euler discretization, $|\alpha|=[1+\gamma\Delta t+\omega^2\Delta t^2)]^{-1/2}$, which is less than or equal to 1 (for γ and Δt positive) for an underdamped oscillator; hence the scheme is unconditionally stable. For BBK, $|\alpha|=[(1-\gamma\Delta t/2)/(1+\gamma\Delta t/2)]^{1/2}$ as long as $\Delta t \leqslant (2/\omega)\sqrt{1-(\gamma/2\omega)^2}$. This is the timestep constraint for stability.

In general, α is a complex number. Writing $\alpha=e^{\beta\Delta t}$ and $\beta=\beta_R+i\beta_I$ and substituting these expressions in Eq. (11), we obtain the following system of equations for β_R and β_I:

$$[\mu e^{\beta_R\Delta t}-\sigma e^{-\beta_R\Delta t}]\sin(\beta_I\Delta t)=0 \tag{13}$$

and

$$[\mu e^{\beta_R\Delta t}+\sigma e^{-\beta_R\Delta t}]\cos(\beta_I\Delta t)+\nu=0. \tag{14}$$

Analysis of these equations determines the regions of stability for each algorithm. There are two possibilities for the above relations to hold: (i) $\sin(\beta_I\Delta t)=0$, pertaining to the overdamped oscillator case ($\gamma>2\omega$) and (ii) $\mu e^{\beta_R\Delta t}-\sigma e^{-\beta_R\Delta t}=0$, the underdamped case ($\gamma<2\omega$). We discuss each case in turn.

Case i (overdamped oscillator):

Since $\beta_I\Delta t=m\pi$ for $m=0,\pm1,\pm2,...$, the value of $\cos(\beta_I\Delta t)$ is -1 for m odd and $+1$ for m even. When m is even, from Eq. (14) we have

$$\mu e^{\beta_R\Delta t}+\sigma e^{-\beta_R\Delta t}+\nu=0,$$

a quadratic equation in $e^{\beta_R\Delta t}$. We arrive at

$$\beta_{R\pm}=\frac{1}{\Delta t}\ln\left[\frac{-\nu\pm\sqrt{\nu^2-4\mu\sigma}}{2\mu}\right].$$

Since β_R is a real number, stability requires the following two conditions:

$$\nu^2-4\mu\sigma>0, \tag{15}$$

$$\frac{-\nu\pm\sqrt{\nu^2-4\mu\sigma}}{2\mu}>0. \tag{16}$$

When m is odd, we obtain from Eq. (14) the relation

$$\mu e^{\beta_R\Delta t}+\sigma e^{-\beta_R\Delta t}-\nu=0,$$

which gives the value of β_R,

$$\beta_{R\pm}=\frac{1}{\Delta t}\ln\left[\frac{\nu\pm\sqrt{\nu^2-4\mu\sigma}}{2\mu}\right].$$

The first stability criterion is identical to the above case ($\nu^2-4\mu\sigma>0$) and the second is

$$\frac{\nu\pm\sqrt{\nu^2-4\mu\sigma}}{2\mu}>0. \tag{17}$$

Case ii (underdamped oscillator):

In the second possible scenario, we require [instead of $\sin(\beta_I\Delta t)=0$]

$$\mu e^{\beta_R\Delta t}-\sigma e^{-\beta_R\Delta t}=0.$$

The solution of this equation yields

$$\beta_R=\frac{1}{2\Delta t}\ln\frac{\sigma}{\mu}. \tag{18}$$

With this value of β_R, we can solve Eq. (14) for β_I to get

$$\beta_I=\frac{1}{\Delta t}\cos^{-1}\left[\frac{-\nu}{2\sqrt{\mu\sigma}}\right]. \tag{19}$$

Since β_R must be negative for asymptotic convergence, we require σ/μ in Eq. (18) to be positive and less than 1. In Eq. (19), the argument of the inverse cosine function must lie between -1 and $+1$. Taken together, we arrive at the following two stability criteria for the underdamped case:

$$\nu^2-4\mu\sigma<0, \tag{20}$$

$$0<\frac{\sigma}{\mu}<1. \tag{21}$$

TABLE II. Accuracy and stability criteria for the five integrators, $\gamma<2\omega$ case. The stability conditions are derived from Eqs. (20) and (21) of the text. The second data column shows the expressions for the perturbed damping constant [Eq. (27)] and the effective normalized frequency [Eq. (29)] for each algorithm. In the third data column, we list the nonperturbative solution of Eqs. (31) and (32) relevant only for LI. As in Table I, we used $\delta=\gamma\Delta t$ and $\epsilon=\omega\Delta t$.

Algorithm	Stability Condition	$\gamma_{\text{eff}}(\Delta t)$, $\omega_{\text{eff}}(\Delta t)$	Nonperturbative $\Delta t(\gamma)$
LI	A-Stable	$\frac{1}{\Delta t}\ln[1+\delta+\epsilon^2]$, $\frac{1}{\Delta t}\cos^{-1}\left[\frac{2+\delta}{2\sqrt{1+\delta+\epsilon^2}}\right]$	$e^{-\delta/2}\left(1+\frac{\delta}{2}\right)=\cos\left[\sqrt{e^{\delta}-\left(1+\frac{\delta}{2}\right)^2}\right]$
LIM2	A-Stable	$\frac{1}{\Delta t}\ln\left[\frac{2+\delta+\epsilon^2}{2-\delta+\epsilon^2}\right]$, $\frac{1}{\Delta t}\cos^{-1}\left[\frac{2}{\sqrt{(2+\epsilon^2)^2-\delta^2}}\right]$	$\Delta t=0$
MID1	A-Stable	$\frac{1}{\Delta t}\ln\left[\frac{4+2\delta+\epsilon^2}{4-2\delta+\epsilon^2}\right]$, $\frac{1}{\Delta t}\cos^{-1}\left[\frac{4-\epsilon^2}{\sqrt{(4+\epsilon^2)^2-4\delta^2}}\right]$	$\Delta t=0$
BBK	$\Delta t<\frac{2}{\omega}\sqrt{1-\left(\frac{\gamma}{2\omega}\right)^2}$	$\frac{1}{\Delta t}\ln\left[\frac{2+\delta}{2-\delta}\right]$, $\frac{1}{\Delta t}\cos^{-1}\left[\frac{2-\epsilon^2}{4-\delta^2}\right]$	$\Delta t=0$
LE	$\Delta t<\frac{2}{\omega}+\frac{\gamma}{\omega^2}$	$\frac{1}{\Delta t}\ln[1+\delta]$, $\frac{1}{\Delta t}\cos^{-1}\left[\frac{2+\delta-\epsilon^2}{2\sqrt{1+\delta}}\right]$	$\Delta t=0$

In Table II, we present the corresponding stability condition for each algorithm. Note that LI, LIM2 and MID1 are unconditionally stable. BBK requires

$$\Delta t<\frac{2}{\omega}\sqrt{1-\left(\frac{\gamma}{2\omega}\right)^2}, \tag{22}$$

and LE requires

$$\Delta t<\frac{2}{\omega}+\frac{\gamma}{\omega^2}. \tag{23}$$

See Table III for relevant limits for MD simulations. Bad behavior in practice, however, is realized for smaller timesteps than those dictated analytically (above) due to nonlinearity and other errors (see Sec. V). Both explicit algorithms require $\Delta t<2/\omega$ for the zero-damping case ($\gamma=0$).

TABLE III. Limiting timesteps for stability of the explicit algorithms. The stability criteria of Eqs. (20) and (21) in the text are used to illustrate the upper bounds for the timesteps (given in fs) corresponding to three γ values and $\omega=0.7$ fs^{-1}, characteristic of an O–H bond stretch (period of 9 fs).

Algorithm	$\gamma=0$	$\gamma=0.05$ fs^{-1}	$\gamma=0.5$ fs^{-1}
BBK	2.86	2.85	2.46
LE	2.86	2.96	3.88

Having found the solution for the homogeneous part, we can determine the solution to the inhomogeneous equation [Eq. (10)] by the method of variation of parameters. This is outlined in Appendix B.

E. Perturbed frequency and damping constant for the underdamped oscillator

Having derived the stability criteria from the solution of the linear inhomogeneous difference equation for various algorithms, we can now examine another interesting aspect of these solutions. The exponential solutions of form α^n exhibit a time dependence on functions in γ and ω. For example, for LI $|\alpha|^n\approx\exp\{[-\gamma/2+(\omega^2\Delta t)/2]t\}$ while for BBK $|\alpha|^n\approx\exp\{[-\gamma/2+(\gamma^3\Delta t^2)/24]t\}$. The $\gamma/2$ term corresponds to physical damping due to friction. The $(\omega^2\Delta t)/2$ term in the LI exponent corresponds to numerical damping. This damping, as we see, is both timestep and frequency dependent.[19,26] As ω and/or Δt increase, numerical damping becomes more severe. For BBK, the numerical damping term is of second order in Δt and is γ^3 dependent.

More generally, we can express this asymptotic behavior for each scheme by the following analysis. For the homogeneous solution of the corresponding differential equation [Eq. (4)], assuming a $Q(n\Delta t)=e^{n\Delta t\beta_0}$ type of solution, we obtain a quadratic equation in β_0,

$$\beta_0^2+\gamma\beta_0+\omega^2=0. \tag{24}$$

Solutions (at most two distinct roots) thus correspond to the following:

$$\beta_{0\pm}=\frac{-\gamma}{2}\pm\sqrt{\frac{\gamma^2}{4}-\omega^2}.$$

Note that the special case $\gamma=0$ yields $\beta_0=\pm i\omega$.

The solution to the inhomogeneous equation [Eq. (4)] can be found similarly as for discrete equations. We write the solution in terms of the coordinate and velocity, $V(t)$, at $t=0$:

$$\begin{aligned}Q(n\Delta t)=&\left(\frac{\beta_{0+}e^{\beta_{0-}t}-\beta_{0-}e^{\beta_{0+}t}}{\beta_{0+}-\beta_{0-}}\right)Q(0)\\&+\left(\frac{e^{\beta_{0+}t}-e^{\beta_{0-}t}}{\beta_{0+}-\beta_{0-}}\right)V(0)\\&+\int_0^{n\Delta t}d\tau\left(\frac{e^{\beta_{0+}\tau}-e^{\beta_{0-}\tau}}{\beta_{0+}-\beta_{0-}}\right)F(n\Delta-\tau).\end{aligned}\tag{25}$$

Since the difference-equation solution, Q^n, converges to the analytic solution, $Q(n\Delta t)$, as $\Delta t\to 0$, the following limit holds:

$$\lim_{\Delta t\to 0}\beta_{\pm}(\Delta t)=\beta_{0\pm}.\tag{26}$$

Thus, our solutions from Eqs. (13) and (14) for β_R and β_I contain the effective values of the damping constant and frequency as a function of the Δt. These are obtained by applying the above limit [Eq. (26)] to the real and imaginary parts of β. Since the real part of β_0 is $-\gamma/2$, we have from Eq. (18),

$$\gamma_{\rm eff}(\Delta t)=2\beta_R=\frac{1}{\Delta t}\ln\frac{\mu}{\sigma}.\tag{27}$$

Therefore,

$$\lim_{\Delta t\to 0}\gamma_{\rm eff}(\Delta t)=\gamma.\tag{28}$$

From the imaginary part of β [Eq. (19)], $\mathrm{Im}(\beta_0)=\frac{1}{2}(4\omega^2-\gamma^2)^{1/2}$, we have

$$\omega_{\rm eff}(\Delta t)=\beta_I=\frac{1}{\Delta t}\cos^{-1}\left[\frac{-\nu}{2\sqrt{\mu\sigma}}\right],\tag{29}$$

so

$$\lim_{\Delta t\to 0}\omega_{\rm eff}(\Delta t)=\tfrac{1}{2}\sqrt{4\omega^2-\gamma^2}.\tag{30}$$

We refer to the functions $\gamma_{\rm eff}$ and $\omega_{\rm eff}$ defined above as the *effective damping constant* and the *effective frequency* of the difference-equation solution, respectively. For a discussion of $\omega_{\rm eff}$ for the implicit-midpoint scheme, for example, see Ref. 23. These expressions for each algorithm are given in Table II and illustrated in Figs. 1 and 2 (discussed below).

Interestingly, all the algorithms exhibit a perturbed friction. For the implicit algorithms, the numerical damping is frequency dependent, whereas for explicit methods, it is not. In LI, numerical damping exists even when $\gamma=0$ (that is why the method has been proposed only for the Langevin model[21]), but for large γ numerical damping increases only logarithmically.

An intriguing question immediately arises from the above findings. Is it possible for the effective frequency and the effective damping constant to be equal to their associated asymptotic (true) values at some nonzero Δt? This is pos-

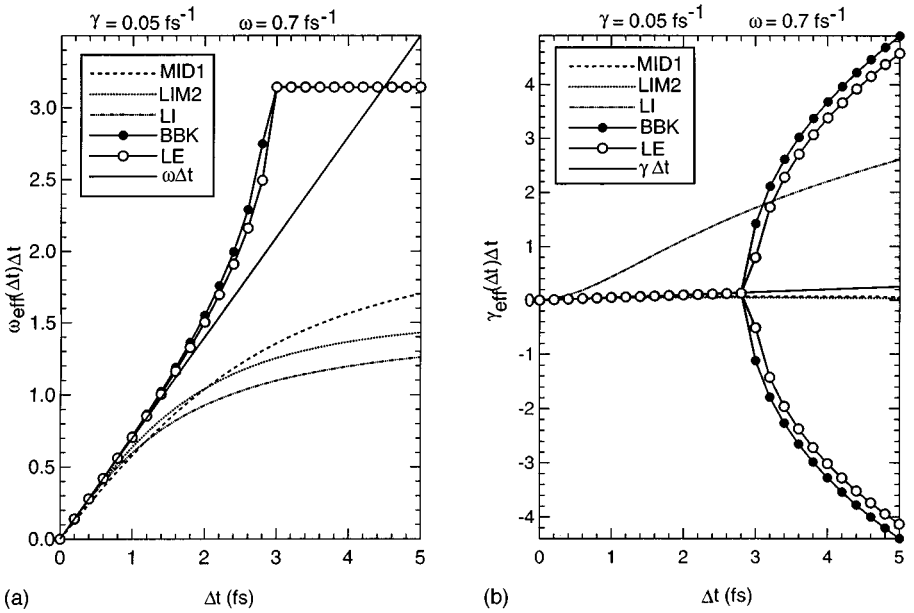

FIG. 1. Products of the effective frequency (a) and damping coefficient (b) with the timestep for a harmonic oscillator of frequency $\omega=0.7$ fs^{-1} and damping coefficient $\gamma=0.05$ fs^{-1} for the five schemes. The frequency used here corresponds approximately to an O–H stretch (period of 9.0 fs). The formulas for $\omega_{\rm eff}$ and $\gamma_{\rm eff}$ given in Eqs. (27) and (29), respectively.

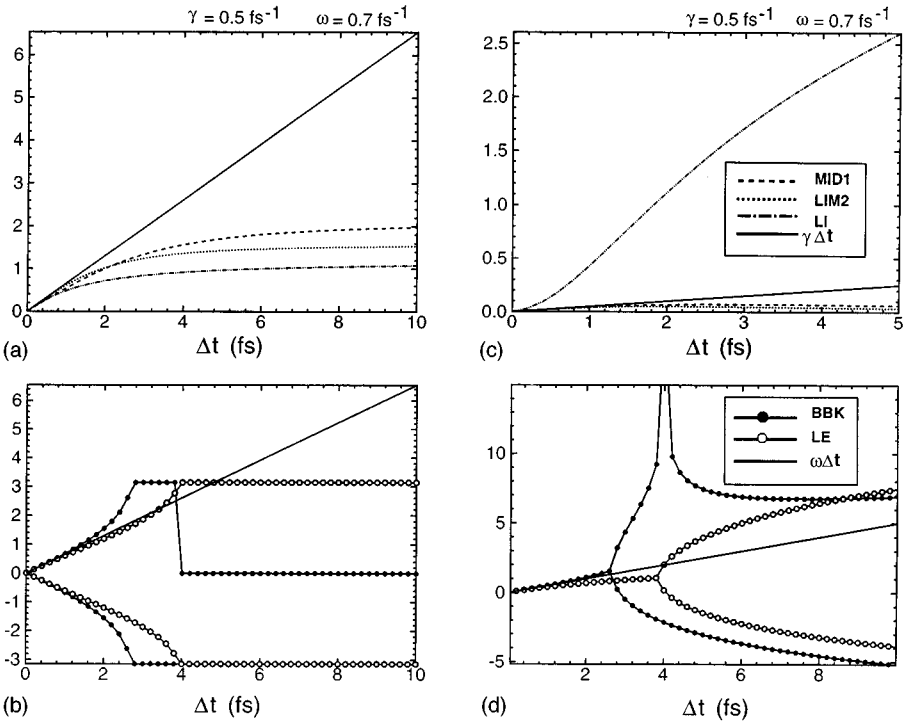

FIG. 2. Products of the effective frequency (a),(b) and damping coefficient (c),(d) with the timestep for a harmonic oscillator of frequency ω=0.7 fs^{-1} and damping coefficient γ=0.5 fs^{-1} for the five schemes. See Fig. 1 caption. The explicit and implicit algorithms are shown separately for clarity. Note the different scales used and the more complex behavior of the explicit algorithms in comparison to a smaller γ (Fig. 1).

sible if there exists $\Delta t>0$ for which both of the following equations in ω, γ, and, Δt are satisfied for the algorithm-dependent γ_{eff} and ω_{eff} expressions:

$$\gamma_{\text{eff}}(\Delta t)-\gamma=0, \tag{31}$$

$$\omega_{\text{eff}}(\Delta t)-\tfrac{1}{2}\sqrt{4\omega^2-\gamma^2}=0. \tag{32}$$

From the above two equations, we can eliminate ω to get a relation between Δt and γ, which may have zeros at values of Δt other than zero. If such nonzero solutions exist, then given a value of γ, we can choose a certain timestep Δt so that the algorithm will exhibit neither numerical damping nor perturbation of frequency! Indeed, the frequency perturbation is a serious artifact since it can lead to resonance,[23] and the numerical damping alters the overall motion of a system, even for a large biomolecule with an enormous range of vibrational modes.[29]

In the case of LI, from Eq. (31), Eq. (27), and Table I, we write

$$\mu=e^{\gamma\Delta t}\sigma,$$

or (since $\sigma=1$)

$$1+\gamma\Delta t+(\omega\Delta t)^2=e^{\gamma\Delta t}.$$

This relation implies

$$\omega\Delta t=\sqrt{e^{\gamma\Delta t}-(1+\gamma\Delta t)}. \tag{33}$$

Combining Eq. (29) with Eq. (32) yields

$$\frac{1}{\Delta t}\cos^{-1}\left[\frac{-\nu}{2\sqrt{\mu\sigma}}\right]=\frac{1}{2}\sqrt{4\omega^2-\gamma^2}$$

or

$$\cos\left[\frac{1}{2}\sqrt{4(\omega\Delta t)^2-(\gamma\Delta t)^2}\right]=\frac{-\nu}{2\sqrt{\mu\sigma}}.$$

We now use the expression for $(\omega\Delta t)^2$ above and the expression for μ, σ, and ν (Table I) to obtain the final relation:

$$\cos\sqrt{e^{\gamma\Delta t}-\left[1+\frac{\gamma\Delta t}{2}\right]^2}=e^{-(\gamma\Delta t)/2}\left[1+\frac{\gamma\Delta t}{2}\right]. \tag{34}$$

Thus, Eqs. (33) and (34) are satisfied by infinitely many $\{\omega\Delta t,\gamma\Delta t\}$ pairs. In Table IV, we show the first few pairs.

TABLE IV. Nonperturbative timesteps for LI. The first few $\{\omega\Delta t,\gamma\Delta t\}$ pairs are listed as solutions to the Eqs. (33) and (34).

$\omega\Delta t$	$\gamma\Delta t$
5.496 53	3.548 476
11.560 42	4.938 652
14.132 42	5.328 136
17.733 73	5.772 243
20.449 01	6.052 594
23.948 71	6.364 588
26.746 55	6.583 356
30.184 59	6.823 213
33.036 61	7.002 538

Similar manipulations can be performed for MID1 and LIM2. It turns out, however, that for these algorithms the cosine function becomes hyperbolic for nonnegative values of Δt, implying no solutions for nonzero Δt. For BBK and LE, the first equation does not contain ω, so for these algorithms the derivation is not possible. Thus, there are nonzero solutions only in the case of LI.

To implement this idea of nondamping, nonperturbative timesteps for LI, one would use a known ω to get Δt, from which γ would be specified. However, this approach cannot be used for a system with more than one frequency, and even for a single-frequency system γ turns out to be rather large. In order to overcome these difficulties in performing nonperturbative simulations with LI, we devise a different approach in the next subsection which, in principle, can be applied for all timesteps and damping constants.

F. Delay-function approach to LI

The artificial damping and frequency-perturbation of LI can be removed, at least in the case of coupled harmonic oscillators, by introducing a linear function Ψ into the LI discretization [see Eq. (5) with $M^{-1}g_E(x)=Ax$]:

$$\frac{x^{n+1}-2x^n+x^{n-1}}{\Delta t^2}+\gamma\frac{x^{n+1}-x^n}{\Delta t}+Ax^{n+1}+\Psi(x^n,x^{n-1}) = M^{-1}r^n. \tag{35}$$

Here, the delay function

$$\Psi(x^n,x^{n-1})=Px^n+Qx^{n-1} \tag{36}$$

is expressed in terms of symmetric matrices A, P, and Q. As done earlier, we recast the above equation in the general form

$$\mu x^{n+1}+\nu x^n+\sigma x^{n-1}=M^{-1}r^n, \tag{37}$$

with

$$\begin{aligned}\mu&=1+\gamma\Delta t+\Delta t^2A,\\ \nu&=-(2+\gamma\Delta t-\Delta t^2P),\\ \sigma&=1+\Delta t^2Q.\end{aligned} \tag{38}$$

If we impose the following two conditions on μ, ν, and σ:

$$\gamma_{\text{eff}}(\Delta t)=\gamma,$$

and

$$\omega_{\text{eff}}(\Delta t)=\frac{1}{2}\sqrt{4\omega^2-\gamma^2},$$

using Eqs. (27) and (29), the forms of P and Q are obtained:

$$\begin{aligned}P&=\frac{1}{\Delta t^2}\left[2I+\gamma\Delta tI-2\exp\left(\frac{-\gamma\Delta t}{2}\right)\right.\\ &\quad\left.\times\cos\left(\Delta t\left[A-\frac{\gamma^2}{4}\right]^{1/2}\right)(I+\gamma\Delta tI+\Delta t^2A)\right],\\ Q&=\frac{1}{\Delta t^2}\left[\exp(-\gamma\Delta t)(I+\gamma\Delta tI+\Delta t^2A)-I\right].\end{aligned} \tag{39}$$

For molecular dynamics using the Backward-Euler method ($\gamma=0$ in LI), the above equations can be simplified considerably. If one further imposes nondamping condition with frequency correction up to order Δt^3 only, the delay function is simply

$$\Psi=-A(x^n-x^{n-1}). \tag{40}$$

This procedure can intuitively be interpreted as an addition of energy to the system at each timestep. To test this idea, we performed MD simulations with LI using Eqs. (35), (36), and (40) on a system of 50 coupled harmonic chains, each made of four carbon atoms. This type of system is chosen to improve statistics, which are poor for a small molecular model. Each oscillator is assigned four frequencies: 0, $2\sin(\pi/8)$, and $2\sin(\pi/4)$, and $2\sin(3\pi/8)$, and we set Δt to 0.2. The results are compared to those from the velocity-Verlet scheme in Fig. 3. From the total and kinetic energy plots generated by the two schemes, we see that the delay-function method conserves energy very well. The fluctuation in total energy in the case of MD is larger with LI. This may be due to the symplecticness of Verlet as opposed to the LI variant above and also because the velocity in Verlet scheme is correct to a higher order, in comparison with LI. The overall agreement between the two methods is satisfactory for the kinetic energy as well. Without the Ψ addition, the LI energy would be much lower, especially in the $\gamma=0$ case, where the energy would decay to zero rapidly with time.

In the nonzero γ case, formulation of Ψ is more difficult, requiring higher order corrections in frequencies to avoid instability. The extent of correction in P can be determined from the following matrix inequality:

$$\Delta t^2P\leqslant[(2+\gamma\Delta t)-2e^{-(\gamma\Delta t)/2}(1+\gamma\Delta t)]I. \tag{41}$$

However, in energetic terms, the Backward-Euler scheme above with simple Ψ structure [Eq. (40)] appears to be a viable approach. Indeed, it is possible to generalize this method to a general nonlinear potential in a straightforward way: replace the term Ax by $M^{-1}g_E(x)$ in the above derivations. The resulting solution to the implicit difference equation can be found by transforming the nonlinear equation into a local optimization problem,[21] as done for LI. However, fluctuations in energy might be large if the eigenstructure changes rapidly. Preliminary experiments on the nucleic acid component deoxycytidine suggest that the fluctuations are quite large in practice, and that the delay-function method requires some modifications. Such applications of the delay function and suitable variations for nonlinear potentials are in progress.

G. Illustration of the five schemes

In Figs. 1 and 2, we illustrate the expected behavior of the five integration algorithms analyzed here with respect to the effective frequency and the effective damping constant. In Fig. 1, we show for two sets of γ and ω the products $\omega_{\text{eff}}\Delta t$ and $\gamma_{\text{eff}}\Delta t$, where ω_{eff} and γ_{eff} are defined by Eqs. (27) and (29), respectively, and expressed in Table II. To make the illustrations relevant to MD simulations, we show

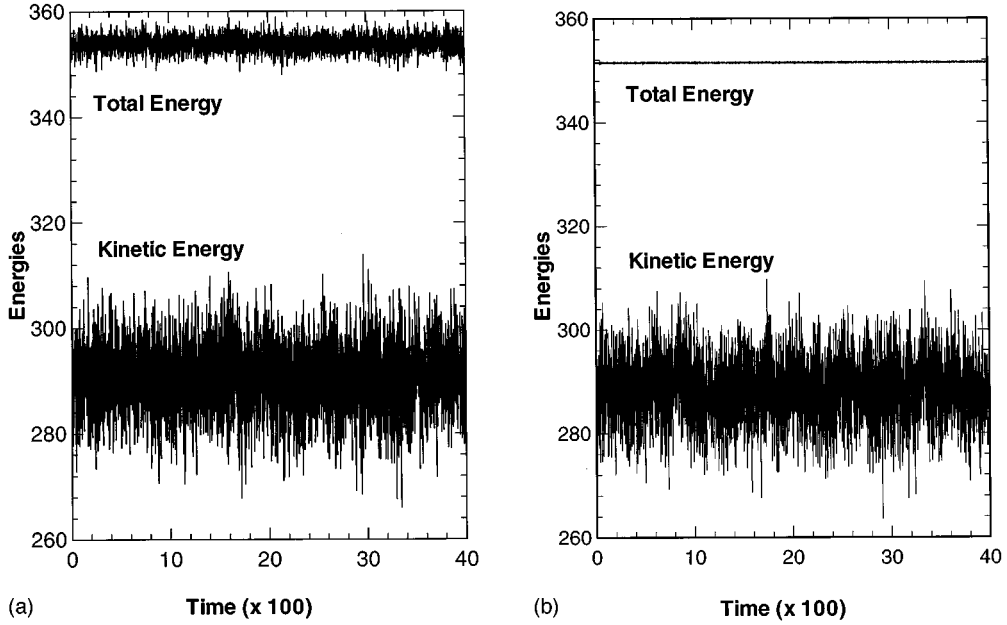

FIG. 3. MD simulations of a molecular model composed of 50 four-carbon systems using the backward-Euler scheme with the delay-function modification (LI with $\gamma=0$) (a) and velocity-Verlet schemes (b). The timestep used is 0.2, and the four frequencies assigned to each four-atom system are: 0, 0.76, 1.41, and 1.85. The lower curve shows the kinetic energy, and the upper curve gives the total energy (see Sec. II F for details).

the timestep in units of femtoseconds (1 fs$=10^{-15}$ s) and choose ω [in Eq. (4)] to correspond to the fastest oscillation present in biomolecular system, such as an O–H stretch. This vibrational mode has a characteristic period of 9 fs, which corresponds to $\omega=(2\pi)/T=0.7$ fs^{-1}. For the damping constant γ [Eq. (4)] we use two values: 0.05 fs^{-1} (Fig. 1), a typical setting in all-atom Langevin simulations, and $\gamma=0.5$ fs^{-1} (Fig. 2), closer to the diffusive (i.e., random-force driven) regime. Note that for the former case, the critical damping for the oscillator is at $\gamma=2\omega=1.4$ fs^{-1}.

From Fig. 1, we see that all algorithms give good agreement with $\omega\Delta t$ (straight line) but depart from the expected value at larger timesteps. The implicit algorithms (dashed and dotted line patterns) produce a smaller value for larger timesteps, and the explicit schemes (circles) give larger values within the stability region (Table II). There is a limit associated with each algorithm. For the effective frequency, we see that

$$\begin{aligned}\lim_{\Delta t\to\infty}\omega_{\text{eff}}(\Delta t)\Delta t&=\cos^{-1}\frac{\gamma}{2\omega}\quad\text{for LI,}\\&=\frac{\pi}{2}\quad\text{for LIM2,}\\&=\pi\quad\text{for MID1,}\\&=\pm\pi\quad\text{for BBK,}\\&=\pm\pi\quad\text{for LE.}\end{aligned}\tag{42}$$

In particular, this behavior implies a troubling scenario for biomolecules since many disparate frequencies may be mapped onto one effective frequency at large discretization steps. In case of LI, this can happen only when $\gamma=0$ (a case when LI is not relevant).

Note also that the π limit for MID1 might explain the third-order resonance of the implicit-midpoint scheme.[23] A scheme like LIM2 or possibly another symplectic algorithm[30] might overcome this order of resonance [Mandziuk, Schlick, Skeel, and Srinivas (unpublished)].

With respect to the product $\gamma_{\text{eff}}\Delta t$, we see again better agreement at small timesteps. However, while LIM2 and MID1 reproduce $\gamma\Delta t$ closely for $\Delta t<1.0$ fs, LI diverges very rapidly with increasing Δt and exhibits a different trend. In particular, the first-order Euler algorithms (both implicit and explicit) show no finite upper bound for $\gamma_{\text{eff}}(\Delta t)\Delta t$, but the others have a zero limit:

$$\begin{aligned}\lim_{\Delta t\to\infty}\gamma_{\text{eff}}(\Delta t)\Delta t&=\infty\quad\text{for LI and LE,}\\&=0\quad\text{for LIM2, MID1, and BBK.}\end{aligned}\tag{43}$$

The behavior for the explicit algorithms is more complex and appreciated by comparing the two γ cases in Figs. 1 and 2. For larger γ, the bifurcation of the product curves at certain values of timesteps is particularly evident. For the γ and ω chosen here (underdamped case) for Fig. 1, the timestep limits for stability are $\Delta t<2.85$ fs for BBK and $\Delta t<2.96$ fs for LE. For Fig. 2 (larger γ), the corresponding limits are 2.46 and 3.88 fs. It is evident that for the underdamped case, the upper bound decreases with increasing γ for BBK whereas for LE, it increases (see also Table III). We see from the figure that there are two branches for $\omega_{\text{eff}}\Delta t$ for both BBK and LE at all timesteps and that one of them matches the expected value at small timesteps. BBK diverges from this product at smaller timesteps (e.g., $\Delta t=2$ fs as opposed to 4 fs for LE).

III. THE FRAMEWORK FOR STATIONARY CALCULATIONS: EXPECTED KINETIC AND POTENTIAL ENERGIES AND CROSS CORRELATION FUNCTIONS

Though valuable for many kinetic aspects, MD simulations can also be used to obtain thermodynamic averages from the generated configurational ensembles. Certainly, the integrator will affect dynamic properties of the system; for example, we have shown that the effective friction and frequency depart from the theoretical values at finite timesteps. But what about stationary properties? How will the integrator affect mean kinetic and potential energies, for example? How will those expected quantities depend on the model parameters (γ,ω) and the timestep? To examine these mean values, we develop in this section the necessary mathematical tools to answer those questions, and additional ones, such as expected cross correlation functions. The latter can be used, for example, to compute the expected translational diffusion constants and other physical properties of interest (next section).

According to classical statistical mechanics, equipartition among all vibrational modes is assumed at thermal equilibrium. Although valid only for quadratic Hamiltonians, the equipartition theorem gives a general reference for the distribution of energy among all vibrational modes. Clearly, the governing Newtonian equations are classical, but quantum effects may be important in many cases for biomolecules, certainly for processes involving electronic rearrangements. The stationary approximation computed below can thus be compared to the expected energies according to classical mechanics.

We now define what we mean by a statistically stationary process.[31] A real process $[x(t)]$ is a *statistically stationary process* in the wide sense if the following two conditions hold:

(i) $\langle x(t)\rangle = \text{constant}$,

(ii) $\langle [x(t+\tau)-\langle x(t+\tau)\rangle][x(t)-\langle x(t)\rangle]\rangle = f(\tau)$ only.

Using condition (i), we can simplify the expression in the left hand-side of (ii), since $\langle x(t+\tau)\rangle = \langle x(t)\rangle$, to obtain

$$\langle [x(t+\tau)-\langle x(t+\tau)\rangle][x(t)-\langle x(t)\rangle]\rangle$$

$$= \langle x(t+\tau)x(t)\rangle - \langle x(t+\tau)\langle x\rangle\rangle - \langle x\langle x(t+\tau)\rangle\rangle + \langle x\rangle^2$$

$$= \langle x(t+\tau)x(t)\rangle - 2\langle x\rangle^2 + \langle x\rangle^2 = \langle x(t+\tau)x(t)\rangle - \langle x\rangle^2.$$

Therefore, the autocorrelation function $\langle x(t+\tau)x(t)\rangle$ depends on τ only; this implies that the first two moments are invariant with respect to translation on the time axis (''second order stationarity''). In particular, $\langle x(t+\tau)x(t)\rangle$ is symmetric with respect to time.

Assuming that the discretized equation [Eq. (10)] governs a statistically stationary process, we now proceed to derive useful expressions for the mean average energies and cross correlation functions. We first restrict ourselves to the algorithms that describe a coordinate propagation only; MID1, which involves a dependent velocity propagation, is treated separately in Appendix A.

We multiply Eq. (10) by Q^n and Q^{n-1}, respectively, and then take the stationary average to obtain the following pair of equations:

$$\mu\langle Q^+Q\rangle + \nu\langle Q^2\rangle + \sigma\langle Q^+Q\rangle = 0, \tag{44}$$

$$\mu\langle Q^+Q^-\rangle + \nu\langle Q^+Q\rangle + \sigma\langle Q^2\rangle = 0. \tag{45}$$

Note here that we have used the stationarity assumption to write

$$\langle Q^nQ^n\rangle = \langle Q^2\rangle,$$

$$\langle Q^{n+1}Q^n\rangle = \langle Q^nQ^{n-1}\rangle = \langle Q^+Q\rangle,$$

$$\langle Q^{n+1}Q^{n-1}\rangle = \langle Q^+Q^-\rangle.$$

We have three unknown quantities in Eqs. (44) and (45): $\langle Q^+Q^-\rangle$, $\langle Q^2\rangle$, and $\langle Q^+Q\rangle$. To obtain a third relationship among those quantities, we square Eq. (10) to arrive at the relation

$$(\mu^2+\nu^2+\sigma^2)\langle Q^2\rangle + 2\nu(\mu+\sigma)\langle Q^+Q\rangle + 2\mu\sigma\langle Q^+Q^-\rangle = \Delta t^4\langle F^2\rangle. \tag{46}$$

Solution to the above system of equations [Eqs. (44), (45), and (46)] leads to

$$\langle Q^2\rangle = \frac{\Delta t^4\langle F^2\rangle(\mu+\sigma)}{(\mu-\sigma)(\mu+\sigma-\nu)(\mu+\sigma+\nu)}, \tag{47}$$

$$\langle Q^+Q\rangle = -\frac{\nu}{(\mu+\sigma)}\langle Q^2\rangle, \tag{48}$$

$$\langle Q^+Q^-\rangle = \frac{[\nu^2-\sigma(\mu+\sigma)]}{\mu(\mu+\sigma)}\langle Q^2\rangle. \tag{49}$$

The value of $\langle F^2\rangle$ above is known: $\langle F^2\rangle = 2\gamma k_BT/\Delta t$.

Many important quantities can now be computed from these expressions for each discretization scheme. For example, once $\langle Q^2\rangle$ is computed according to Eq. (47) from the corresponding $\{\mu,\nu,\sigma\}$ triplet (Table I), the expected potential energy for each mode can be obtained from the expression

$$\langle E_{\text{pot}}\rangle = \tfrac{1}{2}\omega^2\langle Q^2\rangle. \tag{50}$$

For the system of coupled harmonic oscillators, the mean total energy is obtained as a sum over all mean energies for each mode j [see Eq. (4)]. According to classical mechanics, $\langle E_{\text{pot}}\rangle$ for each mode should be equal to $\frac{1}{2}k_BT$. Table V shows in the first data column the resulting expressions in k_BT units for the potential energy. Note that when $\gamma=0$, all potential energies converge to $\frac{1}{2}k_BT$, as expected, as $\Delta t\to 0$ (see also middle data column in Table VI, which shows various limiting cases). Note also that the $\langle E_{\text{pot}}\rangle$ values for LIM2 and BBK are γ independent while the others are not (see also first data column in Table VI). Thus, only these two algorithms give the desired result for the Langevin equation. This is because γ describes the strength at which the system is coupled to the heat bath. Whatever the value of γ is, it should not affect the equilibrium reached, only the rate at which equilibrium is attained. For MID1, for example, the larger the γ the lower the mean potential energy.

TABLE V. Stationary values of the potential, kinetic, and the total energy. Energies are given in k_BT units. See Eqs. (50) and (51) with Eqs. (47), (48), and (49) of the text. See Tables I and VI for further information.

Algorithm	$\langle E_{\rm pot}\rangle$	$\langle E_{\rm kin}\rangle$	$\langle E_{\rm tot}\rangle$
LI	$\frac{\delta(2+\delta+\epsilon^2)}{(\delta+\epsilon^2)(4+2\delta+\epsilon^2)}$	$\frac{2\delta}{(\delta+\epsilon^2)(4+2\delta+\epsilon^2)}$	$\frac{\delta(4+\delta+\epsilon^2)}{(\delta+\epsilon^2)(4+2\delta+\epsilon^2)}$
LIM2	$\frac{2+\epsilon^2}{4+\epsilon^2}$	$\frac{1}{2+\delta+\epsilon^2}$	$\frac{2+\epsilon^2}{4+\epsilon^2}+\frac{1}{2+\delta+\epsilon^2}$
MID1	$\frac{8+2\epsilon^2}{(4+2\delta+\epsilon^2)^2}$	$\frac{2(4+4\delta+\delta^2+\epsilon^2)}{(4+2\delta+\epsilon^2)^2}$	$\frac{2(8+4\delta+\delta^2+2\epsilon^2)}{(4+2\delta+\epsilon^2)^2}$
BBK	$\frac{2}{4-\epsilon^2}$	$\frac{1}{2+\delta}$	$\frac{2}{4-\epsilon^2}+\frac{1}{2+\delta}$
LE	$\frac{2+\delta}{4+2\delta-\epsilon^2}$	$\frac{2}{4+2\delta-\epsilon^2}$	$\frac{4+\delta}{4+2\delta-\epsilon^2}$

To derive the expected kinetic energy for each algorithm, we use the difference formula for the velocity of each algorithm in terms of the positions Q^{n+1}, Q^n, and Q^{n-1} (as defined by $\dot{Q}$ in Table I). We then obtain expressions for $\langle V^2\rangle$ in terms of $\langle Q^2\rangle$, $\langle Q^+Q\rangle$, and $\langle Q^+Q^-\rangle$ and calculate the corresponding energy from the expression:

$$\langle E_{\rm kin}\rangle = \tfrac{1}{2}\langle V^2\rangle. \tag{51}$$

For LI and LE, $V^n=(Q^n-Q^{n-1})/\Delta t$, and we obtain

$$\langle V^nV^n\rangle = \frac{1}{\Delta t^2}\langle (Q^n-Q^{n-1})^2\rangle = \frac{2}{\Delta t^2}[\langle Q^2\rangle - \langle Q^+Q\rangle]. \tag{52}$$

For LIM2 and BBK, $V^n=(Q^{n+1}-Q^{n-1})/2\Delta t$, leading to

$$\langle V^nV^n\rangle = \frac{1}{2\Delta t^2}[\langle Q^2\rangle - \langle Q^+Q^-\rangle]. \tag{53}$$

The symmetric velocity definition for the second-order algorithms LIM2 and BBK implies a zero cross correlation for the position and velocity:

$$\langle Q^nV^n\rangle = \frac{1}{2\Delta t}\langle [Q^{n+1}-Q^{n-1}]Q^n\rangle$$

$$= \frac{1}{2\Delta t}[\langle Q^+Q\rangle - \langle Q^+Q\rangle] = 0,$$

in agreement with exact calculatons. For MID1, since V^n and Q^n are defined interdependently, we must formulate a different linear system and then solve for $\langle (V^n)^2\rangle$ and $\langle Q^nQ^n\rangle$. Details are collected in Appendix A.

From the values of the expected kinetic energy (Table V), we see again that the energy converges to $\frac{1}{2}k_BT$ for $\gamma=0$ as $\Delta t\to 0$ for all methods (see also Table VI). The kinetic energy in all cases depends on γ and ω, expect for BBK where the kinetic energy depends on γ only (i.e., not on ω).

In Figs. 4–6, we plot the resulting curves for the kinetic, potential, and total energies for the various algorithms for one vibrational mode. Two cases are illustrated: $\gamma=0$ [Figs. 4, 6(a) and 6(b)] and $\gamma=0.05$ fs^{-1} [Figs. 5, 6(c) and 6(d)].

For these examples, we again use $\omega=0.7$ fs^{-1}, corresponding to the fastest period expected for biomolecules (O–H stretch, around 9 fs). Recall that the timestep limits for stability of both the BBK and LE schemes are 2.857 fs for $\gamma=0$; and 2.855 fs (BBK) and 2.959 fs (LE) for $\gamma=0.05$ fs^{-1}. We plot the energy means for the implicit and explicit algorithms on separate curves to illustrate the patterns of the explicit algorithms; these exhibit large values at certain timesteps and sharp discontinuities and are plotted on different scales. The scale in the abscissa is in k_BT units.

TABLE VI. Limiting cases of the stationary potential and kinetic energy. Energies are given in k_BT units. See Table V caption. ''NR'' denotes not relevant (in a strict sense, the stability breaks down for explicit schemes).

	$\langle E_{\rm pot}\rangle$		
Algorithm	$\gamma\to 0$	$\Delta t\to 0$	$\Delta t\to\infty$
LI	$\frac{\delta(2+\epsilon^2)}{\epsilon^2(4+\epsilon^2)}$	$\frac{1}{2}\left[1-\left(\delta+\frac{\epsilon^2}{\delta}\right)\right]$	$\frac{\delta}{\epsilon^2}$
LIM2	$\frac{2+\epsilon^2}{4+\epsilon^2}$	$\frac{1}{2}\left[1+\frac{\epsilon^2}{4}\right]$	$\frac{1}{2}\left[1-\frac{2}{\epsilon^2}\right]$
MID1	$\frac{2}{4+\epsilon^2}\left[1-\frac{4\delta}{4+\epsilon^2}\right]$	$\frac{1-\delta}{2}$	$\frac{2}{\epsilon^2}$
BBK	$\frac{2}{4-\epsilon^2}$	$\frac{1}{2}\left[1+\frac{\epsilon^2}{2}\right]$	NR
LE	$\frac{2}{4-\epsilon^2}\left[1-\frac{2\delta}{4-\epsilon^2}\right]$	$\frac{1}{2}\left[1+\frac{1}{4}(\epsilon^2-\delta^2)\right]$	NR
	$\langle E_{\rm kin}\rangle$		
	$\gamma\to 0$	$\Delta t\to 0$	$\Delta t\to\infty$
LI	$\frac{2\delta}{\epsilon^2(4+\epsilon^2)}$	$\frac{1}{2}\left[1-\left(2\delta+\frac{\epsilon^2}{\delta}\right)\right]$	$\frac{2\delta}{\epsilon^4}$
LIM2	$\frac{1}{2+\epsilon^2}$	$\frac{1}{2}\left[1-\frac{\delta}{2}\right]$	$\frac{1}{\delta}$
MID1	$\frac{2}{4+\epsilon^2}\left[1+\frac{12\delta^2}{(4+\epsilon^2)^2}\right]$	$\frac{1}{2}\left[1-\frac{\epsilon^2}{4}\right]$	$\frac{2}{\epsilon^2}$
BBK	$\frac{1}{2}\left[1-\frac{\delta}{2}\right]$	$\frac{1}{2}\left[1-\frac{\delta}{2}\right]$	NR
LE	$\frac{2}{4-\epsilon^2}\left[1-\frac{2\delta}{4-\epsilon^2}\right]$	$\frac{1}{2}\left[1-\frac{\delta}{2}\right]$	NR

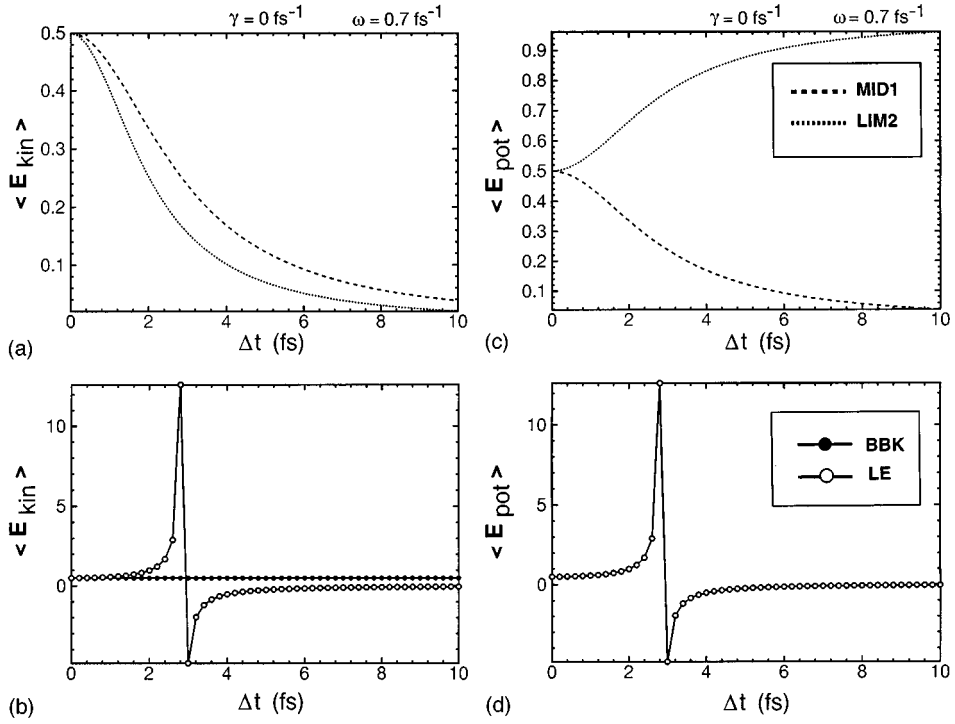

FIG. 4. Stationary kinetic (a),(b) and potential (c),(d) energy components for a harmonic oscillator of frequency ω=0.7 fs^{-1} and damping coefficient γ=0 for the five schemes. See Eqs. (50) and (51) of the text and Table V. The potential energy curves for BBK and LE coincide when γ=0.

We observe from the plots that the potential energy for the explicit algorithms rises sharply from the expected value (0.5 in k_BT units) with Δt, decreasing very sharply from the point where Δt is roughly one third the period (for γ=0, $\langle E_{\rm pot}\rangle$ for BBK and LE is the same). At that timestep, the effective frequency becomes purely imaginary and decreases in magnitude, so the potential energy approaches zero for large timesteps. The kinetic energy for LE exhibits a very

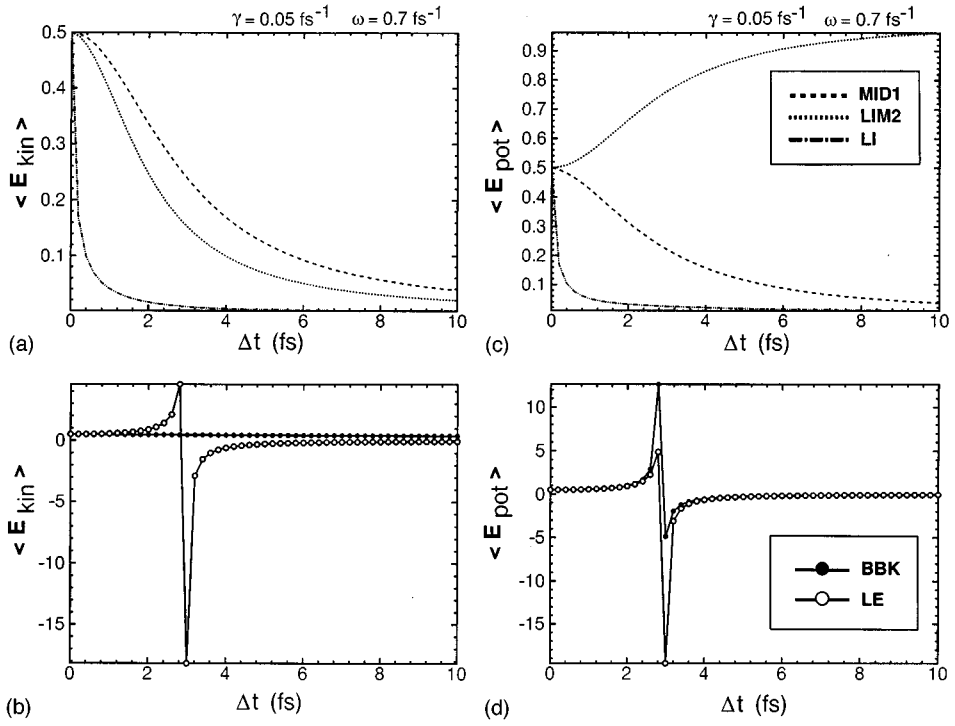

FIG. 5. Stationary kinetic (a),(b) and potential (c),(d) energy components for a harmonic oscillator of frequency ω=0.7 fs^{-1} and damping coefficient γ=0.05 fs^{-1} for the five schemes.

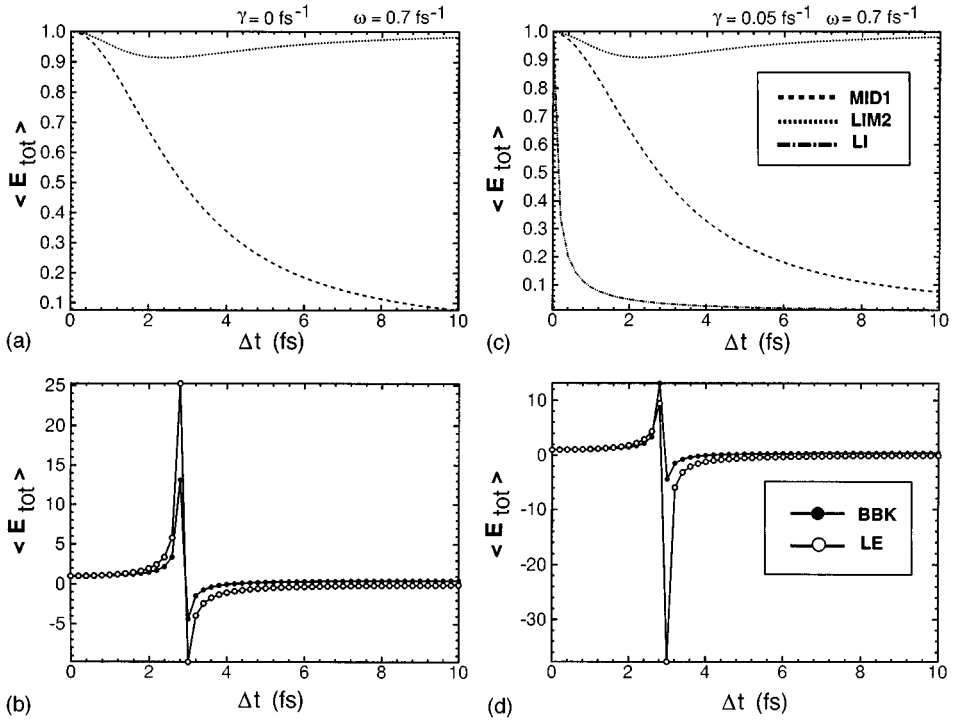

FIG. 6. Stationary total energy for a harmonic oscillator of frequency $\omega=0.7$ fs^{-1} and damping coefficient $\gamma=0$ (a),(b) and $\gamma=0.05$ fs^{-1} (c),(d) for the five schemes. See Table V also.

similar pattern for nonzero γ. In contrast, for BBK the *kinetic* energy decays exponentially (monotonically), whereas the *potential* energy departs from the 0.5 value rapidly. As a consequence, the total energy of both explicit algorithms diverges from the expected value of 1.0 rapidly [Figs. 6(b) and 6(d)], with the divergence much sharper for LE, and sharper for both algorithms when $\gamma\neq 0$.

The implicit algorithms exhibit very different and interesting trends. The kinetic energies decay to zero with increasing Δt for MID1, LIM2, and LI (LI is shown only for nonzero γ as intended). As expected from the asymptotic results for $\Delta t\rightarrow\infty$ (Table VI) the rate of decay is fastest for LI, followed by MID1, and then LIM2 (rates proportional to $1/\Delta t^3$, $1/\Delta t^2$, and $1/\Delta t$, respectively). However, the potential energy of LIM2 increases with Δt as opposed to the decreasing trend for MID1 and LI. As a result, the total energy for LIM2 stays much closer to the expected value, relatively speaking, than LI and MID1. Furthermore, the critical dip in the LIM2 total energy curve suggests a critical timestep to be tested.

The trends above are consistent with the findings of Zhang and Schlick[19] (see Fig. 3 of Ref. 19 in particular) and also extend beyond those results. Clearly, we see that: BBK is preferred over LE, implicit algorithms achieve stability at the cost of some damping, and total energy conservation may be an inadequate criterion for evaluation when dynamic properties are of interest. This fact is highlighted by the compensating trends of the kinetic and potential-energy components of LIM2. While the total energy may be reasonable, the dynamics may be very different (e.g., little momentum but overstretched bonds and angles). By comparison, it appears that for stationary mean square of coordinate calculations LIM2 and BBK are the only suitable candidates, as the averages calculated from them are independent of the damping constant. Also, for these two algorithms, since the velocities are defined symmetric, the cross-average $\langle QV\rangle$ has the true value zero.

Another point warrants emphasis. The stationary curves shown in the figures above serve only as reference, but most likely they have little relevance to practical MD calculations at a constant timestep for systems of many coupled vibrational modes. This is because equipartition is reached very slowly, namely at a rate inversely proportional to γ. In practice, a symplectic scheme at relatively small timesteps is likely to exhibit reasonable energy-conservation behavior if started at expected energy values. For harmonic oscillators, the conservation is exact even at large timesteps for such schemes. Thus, one might see far less severe energy deviations in practice.

For frequencies small compared to the damping coefficient (i.e., slower motions), the potential energy is very close to the equipartition value and the kinetic energy varies roughly as $(2+\gamma\Delta t)^{-1}$ (see Table V) for all five algorithms. Thus, at small timesteps (compared to the periods of the slow motions), the qualitative, as well as the quantitative features of the corresponding energy graphs for coupled normal modes are mostly governed by the highest frequencies.

IV. ZERO FREQUENCY LIMIT: FREE PARTICLE

We now show how to apply the constructs of the previous section, namely correlation functions, to analyze quanti-

ties of physical interest. In the limit of zero frequency, we have a free particle subject to Brownian motion. In this case, the translational diffusion constant, D_t, is related to the coordinate autocorrelation function in the large t limit. Specifically, for a Brownian particle of unit mass, we have

$$\lim_{t\to\infty}\langle Q^2\rangle = 2D_t t = 2\frac{k_BT}{\gamma}\,t. \tag{54}$$

From the discretized equation, we expect to see similar behavior for large n. We can test this by formulating the time-dependent coordinate-correlation function

$$\begin{aligned}\langle Q^nQ^n\rangle &= \frac{\Delta t^4}{(\nu^2-4\mu\sigma)}\sum_{p=1}^{n-1}\sum_{s=1}^{n-1}(\alpha_+^{n-p}-\alpha_-^{n-p})\\ &\quad\times(\alpha_+^{n-s}-\alpha_-^{n-s})\langle F^pF^s\rangle\\ &= \frac{2\gamma k_BT\Delta t^3}{(\nu^2-4\mu\sigma)}\sum_{p=1}^{n-1}(\alpha_+^p-\alpha_-^p)^2\end{aligned} \tag{55}$$

for each algorithm.

For the implicit and explicit Euler (first-order) algorithms, LI and LE, respectively, we get two solutions for α in the limit $\omega\to 0$:

$$\alpha_+ = 1, \quad \alpha_- = \frac{1}{1+\gamma\Delta t}.$$

The denominator, $\nu^2-4\mu\sigma$, for Eq. (55) is equal to $(\gamma\Delta t)^2$ for these schemes. Thus, in the large n limit (for finite Δt), ignoring terms independent of n and noting that $\alpha_-^n\to 0$, we have

$$\begin{aligned}\langle Q^nQ^n\rangle &= \frac{2\gamma k_BT\Delta t^3}{(\gamma\Delta t)^2}\sum_{p=1}^{n-1}(\alpha_+^p)^2\\ &= \frac{2k_BT\Delta t}{\gamma}\sum_{p=1}^{n-1}1 = 2\cdot\frac{k_BT}{\gamma}\cdot n\Delta t\\ &= 2t\cdot\frac{k_BT}{\gamma}.\end{aligned} \tag{56}$$

We therefore recover the exact diffusion coefficient.

Similarly, in the case of BBK, LIM2, and MID1, we find for $\omega=0$,

$$\alpha_+ = 1, \quad \alpha_- = \left(1-\frac{\gamma\Delta t}{2}\right)\Big/\left(1+\frac{\gamma\Delta t}{2}\right),$$

and

$$\nu^2-4\mu\sigma = (\gamma\Delta t)^2.$$

The same expression for the diffusion coefficient is obtained:

$$\lim_{n\Delta t\to\infty}\langle Q^nQ^n\rangle = 2\cdot\frac{k_BT}{\gamma}\cdot n\Delta t.$$

Thus, all five algorithms considered predict the exact diffusion coefficient. In practice, one observes diffusive motion when $(\gamma m)/(\omega^2\Delta t)$ is very large; the inertial effects become relatively small and the motion is mainly governed by the velocity dependent force and the random force. However, at large γ, the error associated with the stationary coordinate-autocorrelation function [see Eq. (47)] may be very large. We saw that LIM2 and BBK do not seem to have this problem, since for them, $\langle QQ\rangle$ is independent of γ. Indeed, it can be seen that for large γ, the difference equation for LIM2 [Eq. (6)] reduces to

$$\begin{aligned}\frac{x^{n+1}+x^{n-1}}{2} &= x^n-\frac{\Delta t}{\gamma}M^{-1}g_E\left[\frac{x^{n+1}+x^{n-1}}{2}\right]\\ &\quad+\frac{\Delta t}{\gamma}M^{-1}r^n,\end{aligned} \tag{57}$$

which is a Brownian algorithm if we assume $(x^{n+1}+x^{n-1})/2\approx x^{n+1}$. In case of LI and LE, we recover in the large γ limit the corresponding implicit and explicit versions of diffusive algorithms. For BBK, in the large γ limit, we find

$$x^{n+1} = x^{n-1}-\frac{2\Delta t^2}{\gamma}M^{-1}g_E(x^n)+\frac{2\Delta t^2}{\gamma}r^n. \tag{58}$$

Note, however, that BBK was derived in the low γ limit. This equation could be viewed as a propagation scheme with twice the timestep, since $g_E(x^n)$ can be considered an intermediate point in the $[x^{n+1},x^{n-1}]$ range. However, the above iteration is not a recommended procedure in practice for numerical reasons.

V. TRAJECTORY ERROR ESTIMATION

In stochastic dynamics, there is no unique trajectory but rather an ensemble of trajectories. Even in MD, different trajectories result from using: various algorithms, same algorithms but different timesteps, different starting points,[11] and different pseudorandom-number generators or seeds. An important question that must be addressed is how to evaluate a generated trajectory. In the absence of direct experimental data for comparison, the convention to date has been to compare results to a simulation at a small timesteps. But is this comparison adequate? How can it be supplemented? Are there rigorous mathematical tests that can be performed? These are difficult questions, and it is clear that the answers must be based on statistical considerations. However, both ways to evaluate static means as well as dynamic functions must be constructed.

Below, we begin to formulate certain statistical tests that could be performed for a given simulation. We then investigate theoretically the behavior of our five integrators in two relevant limits: large Δt and large γ. For these limits, theoretical references exist.

The precise question we wish first to address is: given a certain algorithm, set of parameters, etc., how does one particular resulting trajectory deviate from an exact, or mean, trajectory? Loosely speaking, this is a notion of convergence. However, there is a more rigorous way to quantify this time-dependent statistical behavior of such a deviation.

Let us define the quantity

$$\delta Q^n = Q^n - Q(n\Delta t),$$

where Q^n and $Q(n\Delta t)$ are the positions at time $n\Delta t$ for the discrete and the continuous case, respectively. By definition, the probability of having $\delta Q^n = x$ at time $n\Delta t$ is

$$P(x=\delta Q^n;\ \ t=n\Delta t)=\langle \delta(x-\delta Q^n)\rangle,$$

where the function $\delta(x)$ is defined by

$$\delta(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty} dk\ e^{ikx}.$$

We can now write the above probability as

$$P(x=\delta Q^n;\ \ t=n\Delta t)=\frac{1}{2\pi}\int_{-\infty}^{\infty} dk\langle e^{ik(x-\delta Q^n)}\rangle.$$

To separate the error $Q^n - Q(n\Delta t)$ into systematic and random components, we let $\overline{\delta Q^n}$ represent the former, namely the average value of $Q^n - Q(n\Delta t)$, and δR^n represent the latter, so that: $x - \delta Q^n = x - \overline{\delta Q^n} - \delta R^n$, i.e.,

$$\langle e^{ik(x-\overline{\delta Q^n}-\delta R^n)}\rangle = e^{ik(x-\overline{\delta Q^n})}\langle e^{-ik\cdot\delta R^n}\rangle.$$

Thus, our probability of having the error $Q^n - Q(n\Delta t)$ of the specified value x becomes

$$P(x=\delta Q^n,\ \ t=n\Delta t)$$
$$=\frac{1}{2\pi}\int_{-\infty}^{\infty} dk\ e^{ik(x-\overline{\delta Q^n})}\langle e^{-ik\cdot\delta R^n}\rangle.$$

In general, the random component is a nonlinear function of Gaussians, and may have very complicated distribution. For harmonic oscillators, however, it is a linear combination of the independent Gaussian random variables with time-dependent coefficients. In that case, the expectation value in the above integral can be simplified to the expression

$$\langle e^{-ik\cdot\delta R^n}\rangle = e^{-(k^2/2)\langle(\delta R^n)^2\rangle}.$$

Then, our probability of having the error $Q^n - Q(n\Delta t)$ of the specified value x becomes

$$P(x=\delta Q^n,\ \ t=n\Delta t)$$
$$=\frac{1}{2\pi}\int_{-\infty}^{\infty} dk\ e^{ik(x-\overline{\delta Q^n})}e^{-(k^2/2)\langle(\delta R^n)^2\rangle}.$$

Upon integration, we obtain the result

$$P(x=\delta Q^n,\ \ t=n\Delta t)$$
$$=\frac{1}{\sqrt{4\pi\langle(\delta R^n)^2\rangle}}\exp\left(-\frac{[x-\overline{\delta Q^n}]^2}{4\langle(\delta R^n)^2\rangle}\right).$$

We clearly see that $x(t)$ too is a Gaussian random variable with mean and variance given, respectively, by

$$\langle x(t)\rangle = \overline{\delta Q^n},\quad \langle x(t)^2\rangle = \langle(\delta R^n)^2\rangle.$$

Thus, for any algorithm, if the following limit holds:

$$\lim_{\Delta t\to 0}\langle(\delta R^n)^2\rangle = 0,$$

the algorithm is said to be convergent in the mean square limit.[32]

For a nonlinear system, the above simple statistical description of $x(t)$ is generally not possible. For such a system, we can still use this criteria of trajectory error to determine the stochastic effects as a function of time, as follows.

Let A and B be two trajectories starting at time $t_0 = n_0\Delta t$ separated by a distance

$$|\delta Q_{AB}^{n_0}| = |Q_A^{n_0} - Q_B^{n_0}|,$$

where the norm used here corresponds to the standard Euclidean norm. This difference has an upper bound of

$$|\delta Q_A^{n_0} - \delta Q_B^{n_0}| \leqslant |\delta Q_A^{n_0}| + |\delta Q_B^{n_0}|$$

At any given time $t = n\Delta t$, the expression for δQ_{AB}^n is directly related to the maximum Lyapunov exponent for the dynamics, λ. If the time evolution is written as a linear combination of exponentials, for small $t - t_0$, we can approximate it with the fastest growing term:

$$\langle \delta Q_{AB}^n - \delta Q_{AB}^{n_0}\rangle \simeq e^{\lambda\Delta t(n-n_0)}. \tag{59}$$

The maximum Lyapunov exponent λ reflects the rate of the fastest event in the dynamics of a system making multiple jumps among energy wells. This exponent is found to be a function of the system size.[33] From Eq. (59), the rate of divergence $\Delta(\langle \delta Q_{AB}^n\rangle)/\Delta t$ can be defined as

$$\frac{\langle \delta Q_{AB}^{n+1} - \delta Q_{AB}^n\rangle}{\Delta t} \simeq e^{\lambda\Delta t}. \tag{60}$$

If two trajectories start from the same energy well, the rate at which they diverge depends upon the rate at which they jump across a barrier. The more rare an event is, the slower the divergence rate. So, roughly speaking, there is close relation between these two quantities, namely λ and the transition rate.

A. Application: Barrier crossing

Let us consider a one-dimensional concave potential surface [$g_E(Q) = -\frac{1}{2}\omega_b^2 Q^2$ in Eq. (4)]. Then it is possible to calculate the transmission coefficient associated with the transition state rate of barrier crossing using Kramer's rate theory. The transmission coefficient κ is defined in terms of the rate constant k and the corresponding transition-state-theory value, k_{TST}:

$$\kappa = k/k_{\mathrm{TST}}. \tag{61}$$

It is closely related to the way trajectories diverge.[34] If the trajectory starts at position $Q(0)=0$, the continuous Langevin equation can be used to derive the transmission coefficient in the low-to-intermediate γ regime:[35]

$$\kappa = \left[1 + \frac{1}{k_B T}\langle \tilde{F}(\lambda_r)^2\rangle\right]^{-1/2} = \frac{\lambda_r}{\omega_b}. \tag{62}$$

Here $\tilde{F}(\lambda_r)$ is the Laplace transform of the random force evaluated at the larger root of the quadratic equation:

$$\lambda_r^2 + \gamma\lambda_r - \omega_b^2 = 0,$$

and ω_b is the barrier frequency. The value λ_r can be identified with the Lyapunov coefficient for a nonlinear concave region. Correspondingly, we can derive the transmission coefficient for the discrete case following exactly the same line of analysis.

Recall that in Sec. I we have derived the solution to the discrete equation for Q^n [Eq. (B6)] for real positive frequency. If the initial velocity V^0 is specified instead of Q^1, we can evaluate Q^1 either exactly or numerically using an explicit method:

$$Q^1 = a_1(\Delta t)V^0 + a_2(\Delta t), \tag{63}$$

where $a_1(\Delta t)$ and $a_2(\Delta t)$ (a linear combination of Gaussians) are functions of the method chosen. Above, we assumed the initial position $Q^0=0$, as in the continuous case. The expression for Q^n with $\omega = i\omega_b$ can now be written as

$$Q^n = \frac{\mu}{\sqrt{\nu^2 - 4\mu\sigma}}\left[(\alpha_+^n - \alpha_-^n)Q^1 + \frac{\Delta t^2}{\mu}\sum_{p=1}^{n-1}(\alpha_+^{n-p} - \alpha_-^{n-p})F^p\right]. \tag{64}$$

The roots $\alpha_\pm$, and μ, ν, and σ here correspond to $\omega = i\omega_b$. Now, long-time behavior is governed by the divergent terms in the expression above, so we can write

$$Q^n = \frac{\mu\alpha_+^n}{\sqrt{\nu^2 - 4\mu\sigma}}\left[Q^1 + \frac{\Delta t^2}{\mu}\sum_{p=1}^{n-1}\alpha_+^{-p}F^p\right] + \text{Nondivergent terms}. \tag{65}$$

The summation term above can be simplified using the convolution theorem for the discrete Laplace transform; thus we have for the first part of Q^n:

$$Q^n \approx \frac{\mu\alpha_+^n}{\sqrt{\nu^2 - 4\mu\sigma}}\left[Q^1 + \frac{\Delta t}{\mu}\widetilde{F}(\beta_+)\right], \tag{66}$$

where $\overline{F(\beta_+)}$ is the discrete Laplace transform of the random force,

$$\widetilde{F}(s) = \sum_{m=0}^{\infty}\Delta t F^m \exp(-ms\Delta t),$$

evaluation at $s=\beta_+=(1/\Delta t)\ln\alpha_+$ (see Sec. I). There are two ways for Q^n to be positive (i.e., cross the barrier): $\mu/\sqrt{\nu^2-4\mu\sigma}$ can be positive or negative, but not zero. This means that Q^1 is required to satisfy either:

$$Q^1 > -\frac{\Delta t}{\mu}\widetilde{F}(\beta_+)$$

or

$$Q^1 < -\frac{\Delta t}{\mu}\widetilde{F}(\beta_+),$$

respectively. In both cases we nevertheless arrive at the same expression for κ, valid in the low-to-intermediate friction regime. To determine κ, integration over all the trajectories which cross the barrier and normalization [division by k_{TST}, which assumes $V^0>0$ in Eq. (63)] is required:

$$\kappa = \frac{\left\langle \int_{-(\Delta t/\mu)\widetilde{F}}^{\infty} dQ^1\ Q^1 \exp\left[-\frac{(Q^1)^2}{2\langle (Q^1)^2\rangle}\right]\right\rangle}{\left\langle \int_{a_2}^{\infty} dQ^1\ Q^1\ \exp\left[-\frac{(Q^1)^2}{2\langle (Q^1)^2\rangle}\right]\right\rangle}. \tag{67}$$

Solving the integral, we get

$$\kappa = \frac{\left\langle \exp\left(-\frac{\Delta t^2\widetilde{F}^2}{2\mu^2\langle (Q^1)^2\rangle}\right)\right\rangle}{\left\langle \exp\left(-\frac{\alpha_2^2}{2\langle (Q^1)^2\rangle}\right)\right\rangle}. \tag{68}$$

As the linear transform of a Gaussian distribution is Gaussian, $\widetilde{F}(\beta_+)$ is also a Gaussian, having zero mean, $\langle\widetilde{F}(\beta_+)\rangle=0$, and variance given by $\langle\widetilde{F}^2(\beta_+)\rangle = 2\gamma k_B T/\beta_+$. The latter is obtained using convolution theorem from the autocorrelation function of the random force. Also, a_2 in Eq. (63) is a linear combination of Gaussians. Thus, the transmission coefficient is

$$\kappa = \left[\frac{\langle (Q^1)^2\rangle + \langle a_2^2\rangle}{\langle (Q^1)^2\rangle + \frac{\gamma\Delta t^2}{\mu^2\beta_+}}\right]^{1/2}$$

or

$$\kappa = \left[\frac{1 + \frac{\langle a_2^2\rangle}{\langle (Q^1)^2\rangle}}{1 + \frac{\gamma}{\mu^2\beta_+}\frac{k_B T\Delta t^2}{\langle (Q^1)^2\rangle}}\right]^{1/2}, \tag{69}$$

where from Eq. (63), $\langle (Q^1)^2\rangle = a_1^2\langle (V^0)^2\rangle + \langle a_2^2\rangle$. In the limit of $\Delta t\to 0$, the expression above for κ becomes identical to the one obtained by Gertner *et al.*[35]

B. Transmission coefficient for the five schemes

In order to study the comparative performance of various integrators, we examine the term containing β_+ [denominator in Eq. (69)], since the terms in the numerator will stay approximately the same for all schemes. Recall that a_1 and a_2 define the initial conditions [Eq. (63)]. These initial conditions are the same for LI and LE, as well as for BBK and LIM2. For MID1, the initial velocity V^0 is specified instead of Q^1; therefore, the expression for κ for MID1 does not contain the μ^2 factor in the denominator. Both μ and β_+ in the denominator of Eq. (69) are scheme dependent (Table I). If we further assume $\langle (Q^1)^2\rangle \approx k_B T\Delta t^2$, we can roughly write

$$\kappa \approx \left[1 + \frac{\gamma}{\mu^2\beta_+}\right]^{-1/2}, \tag{70}$$

where we have disregarded the factor determined by the initial conditions.

In Fig. 7, we illustrate κ as a function of Δt and γ for our five schemes. The barrier frequency $\omega_b = 1.0\ \mathrm{fs}^{-1}$ and the damping-constant settings examined are $\gamma = 50$, 500, and

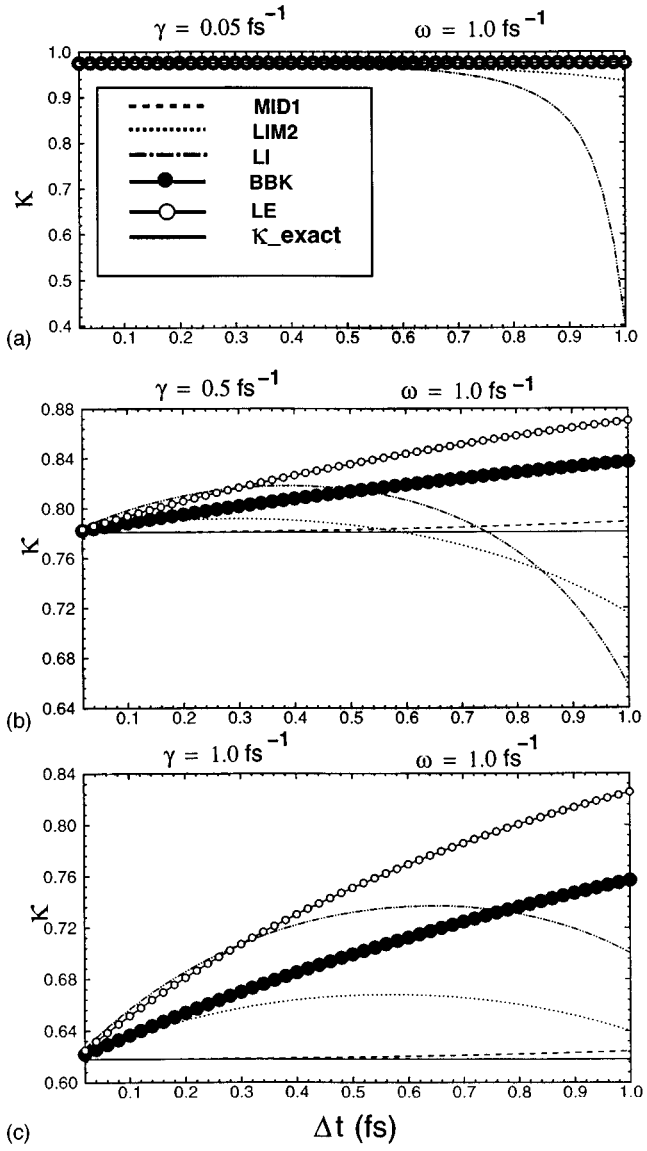

FIG. 7. Transmission coefficient as a function of timestep for a concave potential of frequency $\omega_b = 1.0\ \mathrm{fs}^{-1}$ and three values of γ: 0.05, 0.5, and 1.0 fs^{-1} for the five schemes. For the three γ values, the theoretical values of κ are 0.975, 0.781, 0.618, corresponding to (a), (b), and (c), respectively. See Eqs. (62) and (70) of the text.

1000 ps^{-1}. For larger Δt, unlike the exact case, β_+ can be complex for some schemes leading to complex κ. When γ is small, the behavior of κ is simply correlated to the loss in kinetic energy for all schemes. This is most prominent for LI, where at smaller γ there is little counteracting mechanism to numerical damping. BBK and LE show systematic divergent drifts with Δt for larger γ, whereas the implicit schemes show a batter overall balance in their behavior for the range of Δt examined. For example, the implicit midpoint scheme does quite well in the intermediate γ regime.

For an ideal scheme, the rate will be independent of the timestep used. However, this will not be the case in practice. The differences in transition rates can be considered as a global estimate of how the individual trajectories diverge. For biomolecules, the complex potential is composed of many concave and convex regions, and in such cases, the expression for the rate will be more intricate.

VI. CONCLUSION

This work was motivated by the problem of developing a framework for error analysis for practical simulations of macromolecules today by molecular and Langevin dynamics. This problem is a difficult one for chaotic, multiple-time-scale systems, such as biomolecules, typically simulated at relatively large timesteps for relatively short times. Trajectory assessment is likely to increase in importance in the coming years with improvements in computer hardware and software.

As a start, we have analyzed the behavior of five integration algorithms for the Langevin equation with respect to accuracy, stability, and statistical properties. Two concepts of practical and theoretical importance, namely perturbative damping and perturbative frequency, were introduced. The model considered is a very simple and exactly solvable one. Still, in absence of a general framework for analysis of nonlinear systems, this simple example already provides important theoretical insights with regards to choosing a particular algorithm for a given problem, appreciating algorithmic differences, and quantifying error.

If stochastic dynamics is used for searching the configurational landscape of a molecule for low energy conformations, large timesteps are preferred. In this case, the implicit algorithms are recommended due to their unconditional stability. These algorithms are feasible computationally with an efficient nonlinear minimizer.[30] The symplectic implicit algorithm LIM2 is an excellent candidate for moderate timesteps (with respect to the period of fastest oscillation) since it has many good statistical properties (mean energies) and small numerical damping. For larger timesteps, however, severe departures from the assigned damping coefficient and the original frequency are realized. Both quantities approach zero in the infinite Δt limit. In the case of MID1 (implicit midpoint), we observe similar behavior. However, numerical work is needed to assess these trends since the *asymptotic* limits may have little practical relevance to finite simulations of complex systems. In addition to this limiting behavior, distortion of the phase-space picture and occurrence of resonances have also been reported at timesteps that are relatively large with respect to the fastest period.[23] The perturbative damping and frequency functions introduced here can help in devising cures for these problems.[30]

For the implicit-Euler/Langevin algorithm, LI, theory suggested existence of timesteps for which there is no numerical damping nor shift in original frequency (e.g., Table IV) to overcome these difficulties. However, since the value of the corresponding damping constant is large, the use of such special timesteps in LI is probably not of practical value. Our delay-function approach, in contrast, appears more promising, as demonstrated on a system of linear oscillators. Extensions to nonlinear systems were outlined here and will be tested in further works.

When the detailed dynamic behavior of a molecular sys-

tem is of central interest, the implicit algorithms cannot gain advantage over the explicit algorithms because resolution cannot be as good with larger timesteps. The implicit algorithms are also computationally expensive.[19] As long as one chooses timesteps sufficiently smaller than the limiting value for linear stability (e.g., Table II), BBK appears to have the desirable properties on theoretical grounds with respect to stationary processes. The mean potential energy is independent of γ and the correlation between position and velocity is zero. Furthermore, the upper bound of $\pi/2$ for the effective rotation in space (effective frequency) might dampen low-order resonances in comparison to the implicit midpoint algorithm, which can have maximal rotation within a family of symplectic integrators.[23] In the limit $\gamma\to 0$, BBK becomes the favored Verlet algorithm.

Certainly, our conclusions regarding harmonic oscillators cannot be extrapolated automatically to more complex systems. Both extended analysis and practical simulations are required. In a subsequent paper, we plan to extend and apply the mathematical constructs developed here for analysis of simulations proper for nonlinear systems. The intriguing problems of resonance and perturbed effective Hamiltonians might also be investigated.

ACKNOWLEDGMENTS

We are grateful to Peter Wolynes for many stimulating discussions that led to this research. We thank Wei Xu for preparation of the figures and Heather Malcolm for typing assistance. The work is supported by the National Science Foundation, the National Institutes of Health, and an Alfred P. Sloan research fellowship. T.S. is an investigator of the Howard Hughes Medical Institute.

APPENDIX A: ANALYSIS FOR MID1

For the implicit-midpoint algorithm, MID1, the coordinate and velocity for each normal mode satisfy the following equations:

$$\frac{V^{n+1}-V^n}{\Delta t}+\frac{\gamma}{2}[V^{n+1}+V^n]+\frac{\omega^2}{2}[Q^{n+1}+Q^n]=F^n$$

and

$$\frac{V^{n+1}+V^n}{2}=\frac{Q^{n+1}-Q^n}{\Delta t}.$$

In a matrix form, this can be written as a single equation:

$$\begin{bmatrix} V^{n+1} \\ \omega Q^{n+1}\end{bmatrix}=U\begin{bmatrix} V^n \\ \omega Q^n\end{bmatrix}+\begin{bmatrix} \Delta t F^n \\ 0\end{bmatrix}, \tag{A1}$$

where

$$U=\frac{1}{\phi}\begin{bmatrix} 1-\frac{\delta}{2}-\frac{\epsilon^2}{4} & -\epsilon \\ \epsilon & 1+\frac{\delta}{2}-\frac{\epsilon^2}{4}\end{bmatrix},$$

$$\phi=1+\frac{\delta}{2}+\frac{\epsilon^2}{4},$$

with the definitions:

$$\delta=\gamma\Delta t,\quad \epsilon=\omega\Delta t.$$

The matrix U has two eigenvalues which satisfy a quadratic equation:

$$\phi\alpha^2-2\left[1-\frac{\epsilon^2}{4}\right]\alpha+\left[1-\frac{\delta}{2}+\frac{\epsilon^2}{4}\right]=0.$$

Comparing this form with the general expression of Eq. (11), we obtain the values of μ, ν, and σ (shown in Table I) and use local stability and accuracy analysis of Sec. II. The two eigenvalues of U are

$$\alpha_{\pm}=\frac{1-\frac{\epsilon^2}{4}\pm\sqrt{\frac{\delta^2}{4}-\epsilon^2}}{1+\frac{\delta}{2}+\frac{\epsilon^2}{4}}.$$

The homogeneous part of the solution can be written as

$$\begin{bmatrix} V^n \\ \omega Q^n\end{bmatrix}=\begin{bmatrix} c_1 & c_2 \\ c_3 & c_4\end{bmatrix}\begin{bmatrix} a_+^n \\ \alpha_-^n\end{bmatrix}.$$

For example, with the values of Q^n specified at $n=0$ and 1, the homogeneous solution for Q^n is

$$Q^n=\left[\frac{\alpha_+\alpha_-^n-\alpha_-\alpha_+^n}{\alpha_+-\alpha_-}\right]Q^0+\left[\frac{\alpha_+^n-\alpha_-^n}{\alpha_+-\alpha_-}\right]Q^1. \tag{A2}$$

Also, we have the following relation for $n=1$:

$$\begin{bmatrix} V^1 \\ \omega Q^1\end{bmatrix}=\begin{bmatrix} c_1 & c_2 \\ c_3 & c_4\end{bmatrix}\begin{bmatrix} \alpha_+ \\ \alpha_-\end{bmatrix}.$$

To calculate the stationary averages corresponding to MID1, we multiply on the right by the transpose of Eq. (A1) to obtain

$$\begin{bmatrix} \langle VV\rangle & \omega\langle VQ\rangle \\ \omega\langle QV\rangle & \omega^2\langle QQ\rangle\end{bmatrix}=U\begin{bmatrix} \langle VV\rangle & \omega\langle VQ\rangle \\ \omega\langle QV\rangle & \omega^2\langle QQ\rangle\end{bmatrix}U^T+\begin{bmatrix} \Delta t^2\langle F^2\rangle & 0 \\ 0 & 0\end{bmatrix}.$$

We then obtain three simultaneous linear equations in three unknowns. The solution for the three stationary averages are given by

$$\langle VV\rangle=4k_BT\frac{[4+4\delta+\delta^2+\epsilon^2]}{[4+2\delta+\epsilon^2]^2},$$

$$\omega^2\langle QQ\rangle=4k_BT\frac{[4+\epsilon^2]}{[4+2\delta\Delta t+\epsilon^2]^2},$$

$$\omega\langle QV\rangle=\frac{-4k_BT\delta\epsilon^2}{[4+2\delta+\epsilon^2]^2}.$$

Thus, the stationary values of energies can be written in terms of the above averages:

$$\langle E_{\rm pot}\rangle = \frac{1}{2}\,\omega^2\langle Q^2\rangle = 2k_BT\,\frac{[4+\epsilon^2]}{[4+2\delta+\epsilon^2]^2},$$

$$\langle E_{\rm kin}\rangle = \frac{1}{2}\,\langle VV\rangle = 2k_BT\,\frac{[4+4\delta+\delta^2+\epsilon^2]}{[4+2\delta+\epsilon^2]^2},$$

and

$$\langle E_{\rm tot}\rangle = \langle E_{\rm kin}\rangle + \langle E_{\rm pot}\rangle.$$

APPENDIX B: SOLUTION OF THE INHOMOGENEOUS EQUATION

Substitution of a solution of type

$$Q^n = A_n\alpha_+^n + B_n\alpha_-^n$$

in Eq. (10) of the text gives

$$\mu A_{n+1}\alpha_-^{n+1} + \nu A_n\alpha_+^n + \sigma A_{n-1}\alpha_-^{n-1} + \mu B_{n+1}\alpha_-^{n+1} + \nu B_n\alpha_+^n + \sigma B_{n-1}\alpha_-^{n-1} = (\Delta t)^2 F^n.$$

Since α_+^n and α_-^n satisfy the homogeneous part of the equation, the above equation simplifies to

$$-\nu\alpha_+^n(A_{n+1}-A_n) - \sigma\alpha_+^{n-1}(A_{n+1}-A_{n-1}) - \nu\alpha_-^n(B_{n+1}-B_n) - \sigma\alpha_-^{n-1}(B_{n+1}-B_{n-1}) = (\Delta t)^2 F^n.$$

By imposing the following constraint relating the functions A_n and B_n:

$$\alpha_+^n(A_{n+1}-A_n) + \alpha_-^n(B_{n+}-B_n) = 0, \tag{B1}$$

we can further simplify the above equation to

$$-\sigma\alpha_+^{n-1}(A_{n+1}-A_{n-1}) - \sigma\alpha_-^{n-1}(B_{n+1}-B_{n-1}) = (\Delta t)^2 F^n.$$

Using the constraint again, we obtain

$$-\sigma[\alpha_+^{n-1}(A_{n+1}-A_n) + \alpha_-^{n-1}(B_{n+1}-B_n)] = (\Delta t)^2 F^n. \tag{B2}$$

Thus, we arrive at two simultaneous equations in A_n and B_n, which can be solved to give

$$A_{n+1} = A_n - \frac{(\Delta t)^2 F^n \alpha_+^{-n}}{\sigma(\alpha_- - \alpha_+)},$$

$$B_{n+1} = B_n - \frac{(\Delta t)^2 F^n \alpha_-^{-n}}{\sigma(\alpha_- - \alpha_+)}. \tag{B3}$$

The above equations are first order and can be readily solved to obtain the functions A_n and B_n for $n \geqslant 1$:

$$A_n = A_0 - \frac{(\Delta t)^2}{\sigma(\alpha_- - \alpha_+)} \sum_{k=0}^{n-1} F^k \alpha_+^{-k},$$

$$B_n = B_0 - \frac{(\Delta t)^2}{\sigma(\alpha_- - \alpha_+)} \sum_{k=0}^{n-1} F^k \alpha_-^{-k}. \tag{B4}$$

The final solution is obtained by substituting these functions into the assumed solution to the inhomogeneous equation:

$$Q^n = A_n\alpha_+^n + B_n\alpha_-^n + \frac{(\Delta t)^2 F^n \alpha_-^{-n}}{\sigma(\alpha_- - \alpha_+)} \sum_{k=0}^{n-1} F^k(\alpha_-^{n-k} - \alpha_+^{n-k}). \tag{B5}$$

The constants can be determined by the given initial conditions for Q^0 and Q^1: $Q^0 = A_0 + B_0$ and $Q^1 = \alpha_+A_0 + \alpha_-B_0$. This leads to the final analytical expression for the solution:

$$Q^n = \frac{1}{\sqrt{\nu^2 - 4\mu\sigma}}\,[-\sigma(\alpha_+^{n-1} - \alpha_-^{n-1})Q^0 + \mu(\alpha_+^n - \alpha_-^n)Q^1] + \frac{\Delta t^2}{\sqrt{\nu^2 - 4\mu\sigma}} \sum_{p=1}^{n-1} (\alpha_+^{n-p} - \alpha_-^{n-p})F^p. \tag{B6}$$

[1] S. Nose, Prog. Theor. Phys. Supp. **103**, 1 (1991).
[2] W. Hoover, Phys. Rev. A **31**, 1695 (1985).
[3] T. Schneider and E. Stoll, Phys. Rev. B **17**, 1302 (1978).
[4] B. Dünweg, J. Chem. Phys. **99**, 6977 (1993).
[5] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Oxford University Press, New York, 1990).
[6] H. Mori, Prog. Theor. Phys. **34**, 399 (1965).
[7] R. Zwanzig, J. Stat. Phys. **9**, 215 (1973).
[8] T. C. Gard, *Introduction to Stochastic Differential Equations*, Vol. 114, *Monographs and Textbooks in Pure and Applied Mathematics* (Dekker, New York, 1988).
[9] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations* (Springer, Berlin, 1992).
[10] G. N. Milshtein, *Numerical Integration of Stochastic Differential Equations* (Kluwer Academic, Dordrecht, The Netherlands, 1995).
[11] P. Auffinger, S. Louis-May, and E. Westhof, J. Am. Chem. Soc. **117**, 6720 (1995).
[12] A. J. Lichtenberg and M. A. Lieberman, *Regular and Chaotic Dynamics*, Vol. 38 of *Applied Mathematical Sciences*, 2nd ed. (Springer, New York, 1992).
[13] A. Elofsson and L. Nilsson, J. Mol. Biol. **233**, 766 (1991).
[14] T. Schlick and C. S. Peskin, J. Chem. Phys. **103**, 9888 (1995).
[15] S. A. Adelman, J. Chem. Phys. **71**, 4471 (1979).
[16] W. F. van Gunsteren and H. J. C. Berendsen, Mol. Simul. **1**, 173 (1988).
[17] R. W. Pastor, in *The Molecular Dynamics of Liquid Cyrstals*, edited by G. R. Luckhurst and C. Veracini (Kluwer Academic, Dordrecht, The Netherlands, 1994), pp. 85–138.
[18] A. Brünger, C. B. Brooks, and M. Karplus, Chem. Phys. Lett. **105**, 495 (1982).
[19] G. Zhang and T. Schlick, Mol. Phys. **84**, 1077 (1995).
[20] G. Dahlquist and Å. Björk, *Numerical Methods* (Prentice Hall, Englewood Cliffs, NJ, 1974).
[21] C. S. Peskin and T. Schlick, Comm. Pure Appl. Math. **42**, 1001 (1989).
[22] P. G. Reinhall, T. K. Caughey, and D. W. Storti, Trans. ASME **56**, 162 (1989).
[23] M. Mandziuk and T. Schlick, Chem. Phys. Lett. **237**, 525 (1995).
[24] R. W. Pastor, B. R. Brooks, and A. Szabo, Mol. Phys. **65**, 1409 (1988).
[25] T. Schlick, in *Mathematical Applications to Biomolecular Structure and Dynamics*, edited by J. Mesirov, K. Schulten, and D. W. Sumners, Volume 82 of *IMA Volumes in Mathematics and Its Applications* (Springer, New York, 1996), pp. 219–247.
[26] G. Zhang and T. Schlick, J. Comp. Chem. **14**, 1212 (1993).
[27] J. M. Sanz-Serna and M. P. Calvo, *Numerical Hamiltonian Problems* (Chapman and Hall, London, 1994).
[28] C. M. Bender and S. A. Orszag, *Advanced Mathematical Methods for Scientist and Engineers* (McGraw-Hill, New York, 1978).
[29] P. Derreumaux and T. Schlick, Proteins, Structure, Function and Genetics **21**, 282 (1995).
[30] R. D. Skeel, G. Zhang, and T. Schlick, *A family of symplectic integrators: Stability, accuracy, and molecular dynamics applications* (SIAM J. Sci. Comp., in press).
[31] A. I. Khinchin, *Mathematical Foundations of Statistical Mechanics* (Dover, New York, 1948).
[32] A. Greiner, W. Strittmatter, and J. Honerkamp, J. Stat. Phys. **51**, 95 (1988).
[33] G. D. Quinlan and S. Tremaine, Mon. Not. R. Astron. Soc. **259**, 505 (1992).
[34] P. Hänggi, P. Talkner, and M. Borkovec, Rev. Mod. Phys. **62**, 251 (1990).
[35] B. J. Gertner, K. R. Wilson, and J. T. Hynes, J. Chem. Phys. **90**, 3537 (1989).