Report No.
UCB/SEMM-2009/01

Structural Engineering
Mechanics and Materials

Application of a Symplectic-Energy-Momentum Preserving Integrator to Molecular Dynamics

By

Wei He and Sanjay Govindjee

January 2009

Department of Civil and Environmental Engineering
University of California, Berkeley

Application of a Symplectic-Energy-Momentum Preserving Integrator to Molecular Dynamics

Wei He, Sanjay Govindjee

Structural Engineering, Mechanics and Materials
Department of Civil and Environmental Engineering
University of California, Berkeley,
Berkeley, CA 94720

Abstract: In molecular dynamics simulations, an integrator-induced resonance is observed for conservative molecular system subject to the classical equations of motion when a Verlet integrator or Implicit Midpoint scheme is used. In this report, an existing variational integrator with adaptive timesteps is introduced to handle resonance. This Symplectic-Energy-Momentum (SEM) preserving algorithm is first applied to a diatomic molecule governed by a Morse potential and then it is further applied to a 22-atom model system. Computational experiments indicate that the SEM algorithm can avoid energy resonances and produce more accurate sampling of phase space. Moreover, it can increase the feasible timestep and hence has the potential to improve the simulation times. These are the main advantages over other fixed timestep methods. Its main disadvantage, however, is that the algorithm is computationally more expensive since one needs to solve a complicated nonlinear system of equations during its use.

Keywords: Molecular Dynamics, Resonance, Symplectic-Energy-Momentum Preserving

1 Introduction

In molecular dynamics simulations, complex atomic motion is followed by numerically integrating Newton's classical equations [1]

$$\boldsymbol{M}\dot{\boldsymbol{V}}(t) = -\nabla E_p(\boldsymbol{X}) \tag{1}$$

$$\dot{\boldsymbol{X}}(t) = \boldsymbol{V}(t) \tag{2}$$

where $\boldsymbol{M}$ is a mass matrix, $\boldsymbol{X}$ is the vector of atomic positions, $\boldsymbol{V}$ is the corresponding velocity vector, and E_p is the potential energy. Because the multivariate potential energy landscape is usually highly non-convex and the resulting effective forces are nonlinear, molecular dynamics simulations are very costly in most cases.

Typically the Newtonian equations of motions are solved numerically using explicit schemes such as Verlet. Such schemes are simple to formulate and fast to solve, but they are known to induce resonance. Also, they impose a severe restriction on the integration time step size: Δt must be at least as small as the most rapid vibrational mode. This generally limits Δt to be in the femtosecond (10^{-15} s) range. This in fact is the typical step size used in molecular dynamics simulation programs such as CHARMM. Since key conformational changes in biomolecules occur on time scales of $10^{-12} - 10^{2}$ s, considerable effort has focused on increasing the integration time step.

Implicit numerical integrators with high stability have been introduced to molecular dynamics simulations since these integrators usually permit a larger step size than Verlet. However, implicit integrators are computationally more demanding since one has to solve a complicated system of equations which are usually nonlinear in every step. Moreover, most implicit integrators such as the midpoint scheme can only delay energy resonance. When the step size increases and hits specific values, resonance occurs and leads to incorrect phase diagrams.

Symplectic numerical integrators for Hamiltonian systems have been the focus of researchers in recent years. 'Symplectic' means the preservation of a specific two-form mathematically and the preservation of areas in phase space physically. Symplecticness is a desirable property which usually leads to long-time stability of the method since Hamiltonian systems are symplectic as well as energy and momentum preserving themselves [2, 9]. For problems in the linear regime, symplectic integrators are known to conserve energy and stability holds at large step sizes[18]. However, this is not the case for nonlinear systems. Implicit midpoint and Verlet, for instance, are both symplectic integrators but they still induce resonance when applied to Morse oscillators. Instability, or large energy fluctuation, occurs at timesteps that still satisfy the linear stability condition and these instabilities usually can only be avoided by reducing the timestep[21]. In other words, this sort of nonlinear instability sets a further timestep limitation. There has been a recent effort to find integrators which

preserve as many structural properties as possible to circumvent this limitation and achieve long simulation times. Unfortunately, a well-known theorem has limited the possibility that constant time stepping algorithms be symplectic and energy and momentum preserving[7]. As a consequence, a constant time stepping algorithm can only be energy-momentum preserving or symplectic-momentum preserving.

It should be pointed out that structure preservation and accuracy are two different concepts. Structure preserving integrators alone can not guarantee accurate trajectories which can only be obtained with high-order methods and small time steps[10]. High-order methods, on the other hand, may not preserve structural properties such as energy and momentum. Second order Newmark methods and fourth order Runge-Kutta methods have long been applied to conservative systems [11, 13] and their energy-momentum behaviors indicate that the Newmark family of methods has a fluctuating energy and momentum which is typical of symplectic methods as opposed to the divergent behaviors of classical Runge-Kutta, see Fig.1 (due to Kane *et al*, [11]). For systems with friction, say, simple Rayleigh type dissipation, Newmark methods also accurately simulate energy decay, unlike standard methods such as Runge-Kutta schemes. This is a clear demonstration of the fact that traditional measures of integrator accuracy, such as truncation error, are not necessarily appropriate when discussing these kinds of structure preserving schemes, as they often perform far better than expected. Some high-order classical numerical algorithms show deficiency in long-time performance. This also implies that one should consider a trade-off between accuracy and structure preserving properties when choosing integrators. Particularly, for those systems with complicated, unstable, or chaotic trajectories, such as biomolecules, one should probably concentrate on statistical properties and approach the true solution by preserving as much of the structure as feasible[10].

The nice structure preserving properties of Newmark methods are due to their variational nature, which is believed to be one of the primary reasons why this class of algorithms performs well[11, 12, 23]. The classical Newmark family with $\gamma = 1/2$ as well as related integration algorithms are variational in the sense of the Veselov formulation of discrete mechanics. Such variational algorithms are symplectic and momentum preserving and have excellent global energy behavior for linear systems. The formulation of these algorithms is based on the discretization of Hamilton's principle rather than the equations of motion directly, and this leads in a natural way to symplectic-momentum

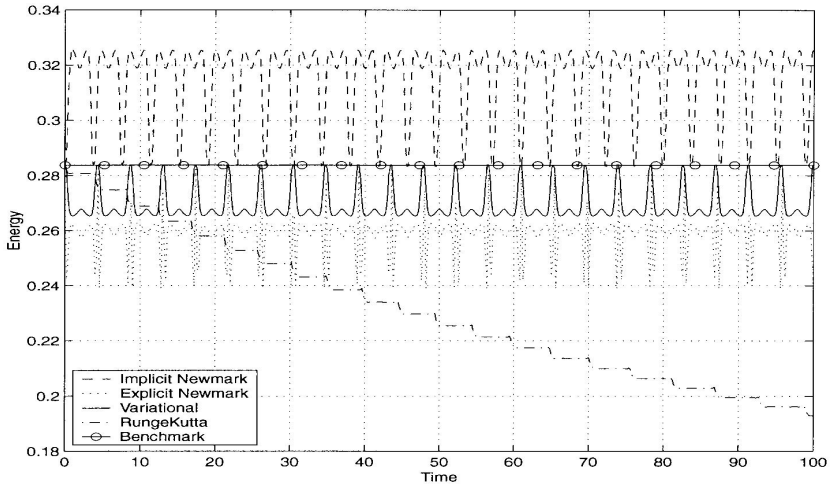

Figure 1: Long-time energy behaviors of different algorithms (Implicit Newmark $\beta = 1/4$, $\gamma = 1/2$, Explicit Newmark $\beta = 0$, $\gamma = 1/2$)[11].

preserving integrators. Such integrators, however, can not in general preserve the Hamiltonian exactly, though they usually have good energy performance for linear systems, assuming small time steps.

As mentioned above, variational integrators are naturally symplectic and momentum preserving but usually can not preserve energy, which has been shown theoretically in [7] and confirmed computationally in [1]. It should be emphasized that this only applies to integrators with constant timesteps. With adaptive timesteps, a symplectic integrator can be adjusted to conserve energy as well[10], for example, the SEM algorithm studied in this report. Loosely speaking, the algorithm introduces a constraint that energy be preserved and a new timestep is accordingly computed in every step to satisfy this constraint. Because of its energy preserving property, this SEM integrator can be applied to molecular dynamics simulations to handle the resonance issue mentioned earlier. Numerical experiments will be shown to demonstrate that it can avoid resonance at large time steps and produce more accurate trajectories in phase space than an implicit midpoint method. However, it will also be shown to be computationally more expensive because a nonlinear system needs to be solved in every step. Some optimization techniques are used to systematize this procedure, but a clear price must be paid.

2 Brief Description of SEM Algorithm

2.1 Governing Equations

The SEM algorithm used in this report was developed by Kane, Marsden and Ortiz[10]. This time-stepping algorithm approximates a flow of a system of ODEs for a mechanical system by discretizing Hamilton's principle rather than the equations of motion. Thereby, a framework is easily developed to conserve invariants of the motion.

For completeness, we start with the following quote from Arnold[2] to describe the terminology commonly used in Lagrangian mechanics:

"$L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) = T - V$ is the *Lagrange function* or *Lagrangian*, q_i are the *generalized coordinates*, $\dot{q}_i$ are *generalized velocities*, $\partial L/\partial \dot{q}_i = p_i$ are *generalized momenta*, $\partial L/\partial q_i$ are *generalized forces*. $\int_{t_0}^{t_1} L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t)dt$ is the *action*."

The equation

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\boldsymbol{q}}}\right) - \frac{\partial L}{\partial \boldsymbol{q}} = 0 \tag{3}$$

is the *Euler-Lagrange equation* for the functional

$$\Phi = \int_{t_0}^{t_1} L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t)dt. \tag{4}$$

Hamilton's principle of least action states that [2]"motions of Lagrangian mechanical systems coincide with extremals of the functional Φ". Furthermore, "the curve γ is an extremal of the function Φ on the space of curves joining $(t_0, \boldsymbol{q}_0)$ and $(t_1, \boldsymbol{q}_1)$, if and only if the *Euler-Lagrange equation* is satisfied along γ."

For our problems, the Lagrangian is of the standard form

$$L(\boldsymbol{q}, \dot{\boldsymbol{q}}) = \frac{1}{2}\dot{\boldsymbol{q}}^T \boldsymbol{M} \dot{\boldsymbol{q}} - V(\boldsymbol{q}) \tag{5}$$

where $\boldsymbol{M}$ is a symmetric positive-definite matrix and $\boldsymbol{V}$ the system potential. The associated *discrete Lagrangian* is defined by a 2-point map over the configuration space Q:

$$L_d : \ Q \times Q \to \mathbb{R} \tag{6}$$

where

$$L_d(\boldsymbol{q}_k, \boldsymbol{q}_{k+1}, h_k) = \frac{1}{2}\left(\frac{\boldsymbol{q}_{k+1}-\boldsymbol{q}_k}{h_k}\right)^T \boldsymbol{M}\left(\frac{\boldsymbol{q}_{k+1}-\boldsymbol{q}_k}{h_k}\right) - V\left(\frac{\boldsymbol{q}_k+\boldsymbol{q}_{k+1}}{2}\right). \quad (7)$$

Here $\boldsymbol{q}_k$ denotes the position coordinates at time step k and $h_k > 0$ is the time step. L_d is regarded as a function of two nearby position coordinates $\left(\boldsymbol{q}_k, \boldsymbol{q}_{k+1}\right)$. It is worth noting that the *discrete Lagrangian* can also have other definitions. Our choice is not unique.

For a positive integer N, one can further define the *action sum*, a discrete analog of the action integral, as

$$S_d = \sum_{k=0}^{N-1} h_k L_d\left(\boldsymbol{q}_k, \boldsymbol{q}_{k+1}, h_k\right) \quad (8)$$

The *discrete variational principle* states that the discrete evolution equations extremize the *action sum* given fixed end points, $\boldsymbol{q}_0$ and $\boldsymbol{q}_N$. The resulting *discrete Euler-Lagrange (DEL) equations* are as follows:

$$h_k D_1 L_d\left(\boldsymbol{q}_k, \boldsymbol{q}_{k+1}, h_k\right) + h_{k-1} D_2 L_d\left(\boldsymbol{q}_{k-1}, \boldsymbol{q}_k, h_{k-1}\right) = 0 \quad (9)$$

for all $k = 1, ..., N-1$, This is the first governing equation of the SEM algorithm, in which D_1 denotes the derivative with respect to the first slot and D_2 the derivative with respect to the second slot[11].

With the discrete energy defined as

$$E_d\left(\boldsymbol{q}_{k-1}, \boldsymbol{q}_k, h_{k-1}\right) = \frac{1}{2}\left(\frac{\boldsymbol{q}_k-\boldsymbol{q}_{k-1}}{h_{k-1}}\right)^T \boldsymbol{M}\left(\frac{\boldsymbol{q}_k-\boldsymbol{q}_{k-1}}{h_{k-1}}\right) + V\left(\frac{\boldsymbol{q}_{k-1}+\boldsymbol{q}_k}{2}\right) \quad (10)$$

the second governing equation defining the algorithm is

$$E_d\left(\boldsymbol{q}_{k-1}, \boldsymbol{q}_k, h_{k-1}\right) - E_d\left(\boldsymbol{q}_k, \boldsymbol{q}_{k+1}, h_k\right) = 0 \quad (11)$$

which imposes the constraint that total energy should be conserved from one step to the next. By solving these two equations, one can pass from data $\left(\boldsymbol{q}_{k-1}, \boldsymbol{q}_k, h_{k-1}\right)$ to $\left(\boldsymbol{q}_k, \boldsymbol{q}_{k+1}, h_k\right)$.

As described in [10], the algorithm is symplectic, energy preserving and momentum preserving. The nonlinear system of equations can be solved by using optimization techniques. More details will be presented in Section 5.

3 Application to the H-Br System

3.1 Model and Discrete Equations

The first model system is the diatomic molecule, HBr, with the hydrogen and bromine atoms interacting via the Morse potential [22]

$$E_p(r) = D\left(1 - \exp[-S(r - r_0)]\right)^2, \tag{12}$$

where $r = |x_H - x_{Br}|$ represents the interatomic distance. x_H and x_{Br} denote position of hydrogen and bromine atoms, respectively, and r_0 is the equilibrium bond distance. D is the well depth, and S is a parameter controlling the width of the well. The two molecules are assumed to move in one dimension under the influence of the Morse potential.

The main discrete equations for the SEM algorithm are as follows:

$$\begin{aligned} &h_{k-1}\left[\boldsymbol{M}\frac{\boldsymbol{q}_k - \boldsymbol{q}_{k-1}}{h_{k-1}^2} - \frac{1}{2}E_p'\left(\frac{\boldsymbol{q}_{k-1} + \boldsymbol{q}_k}{2}\right)\right] \\ &+h_k\left[-\boldsymbol{M}\frac{\boldsymbol{q}_{k+1} - \boldsymbol{q}_k}{h_k^2} - \frac{1}{2}E_p'\left(\frac{\boldsymbol{q}_k + \boldsymbol{q}_{k+1}}{2}\right)\right] = 0 \end{aligned} \tag{13}$$

$$\begin{aligned} &\frac{1}{2}\left(\frac{\boldsymbol{q}_k - \boldsymbol{q}_{k-1}}{h_{k-1}}\right)^T \boldsymbol{M}\left(\frac{\boldsymbol{q}_k - \boldsymbol{q}_{k-1}}{h_{k-1}}\right) + E_p\left(\frac{\boldsymbol{q}_{k-1} + \boldsymbol{q}_k}{2}\right) \\ &-\frac{1}{2}\left(\frac{\boldsymbol{q}_{k+1} - \boldsymbol{q}_k}{h_k}\right)^T \boldsymbol{M}\left(\frac{\boldsymbol{q}_{k+1} - \boldsymbol{q}_k}{h_k}\right) - E_p\left(\frac{\boldsymbol{q}_k + \boldsymbol{q}_{k+1}}{2}\right) = 0 \end{aligned} \tag{14}$$

where $\boldsymbol{q}$ denotes the position vector and h the timestep.

With the parameters and initial conditions given in [1], we can solve this nonlinear system of equations and plot the conjugate pair of position and momentum for the Morse oscillator alone $(\boldsymbol{r}, \boldsymbol{p})$, where for a reduced mass μ

$$\boldsymbol{r} = \boldsymbol{x}_H - \boldsymbol{x}_{Br} \tag{15}$$

$$\boldsymbol{p} = \mu(\boldsymbol{v}_H - \boldsymbol{v}_{Br}). \tag{16}$$

In addition, we can check the total energy of the system at each step.

3.2 Numerical Results and Discussions

1. *Energy Behavior*. We can see from Fig. 2 and Fig. 3 that the total energy is well conserved no matter what initial timestep is used and resonance is completely avoided. The horizontal solid line represents the initial input energy to the system. Furthermore note that $h = 4.02fs$ and $h = 7.22fs$ are two step sizes leading to resonance when the implicit midpoint method is used[1].

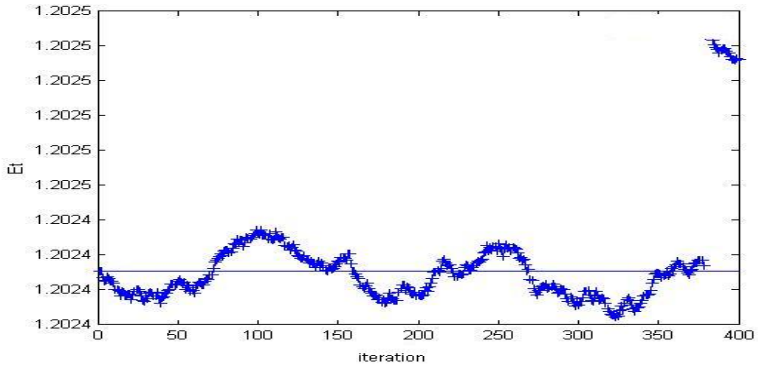

Figure 2: Energy vs. iteration number by SEM, $h_0 = 4fs$.

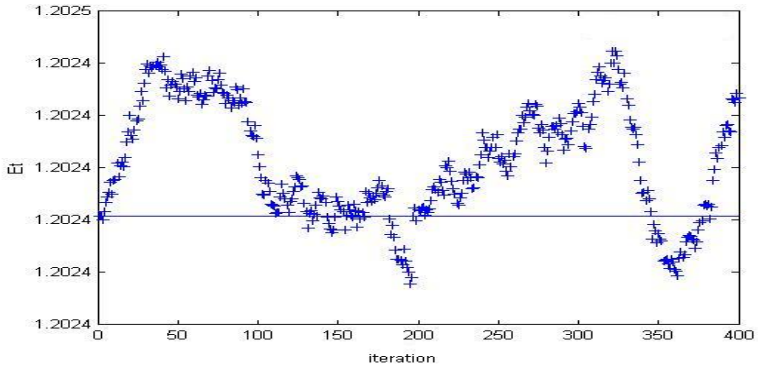

Figure 3: Energy vs. iteration number by SEM, $h_0 = 7.22fs$.

2. *Trajectories*. For this particular problem, exact trajectories are known analytically. If the energy E is less than D, then integration gives [22]:

$$Sx = \log[1 - \cos\theta \cos(2\pi\nu_0 t \sin\theta)/\sin^2\theta] \tag{17}$$

where $\cos^2\theta = E/D$, $\nu_0 = (S/2\pi)(2D/\mu)^{\frac{1}{2}}$, and x is the stretch of the interatomic bond. Thus the motion is periodic and may be described as the logarithm of a simple harmonic motion. The exact trajectory in phase space has an elliptical-like shape. Comparing Fig. 4 and Fig. 5 (by SEM) with Fig. 6 and Fig. 7 (by IM), we can see that SEM produces correct trajectories, while the IM method has trouble for large timesteps.

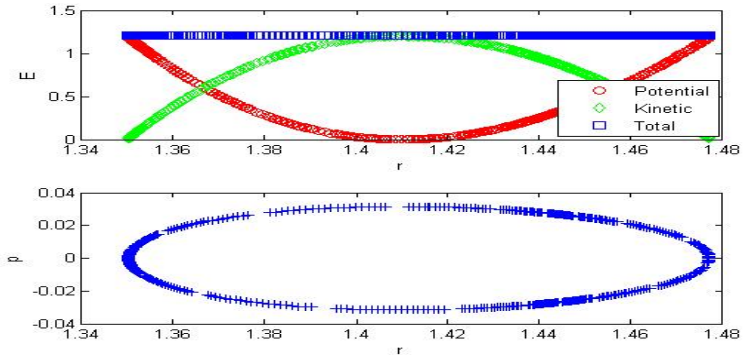

Figure 4: Energy and trajectory by SEM, $h_0 = 4fs$.

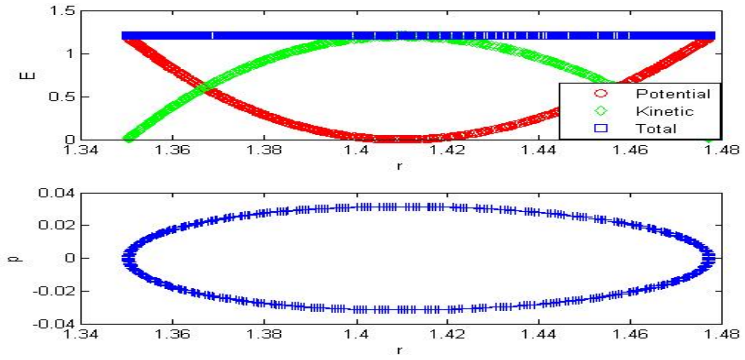

Figure 5: Energy and trajectory by SEM, $h_0 = 7.22fs$.

3. *Period*. The energy-dependent angular frequency ω_E is given by

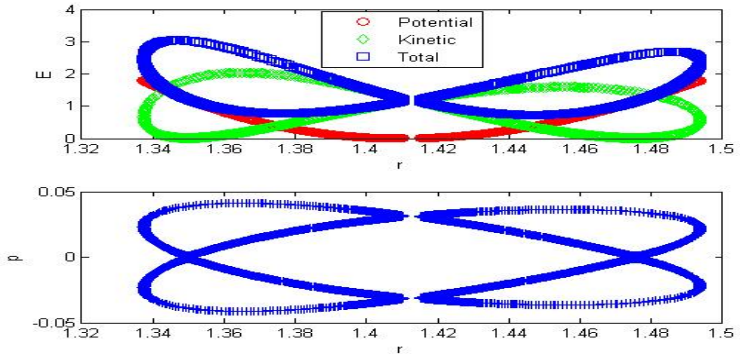

Figure 6: Energy and trajectory by IM, $h = 4fs$.

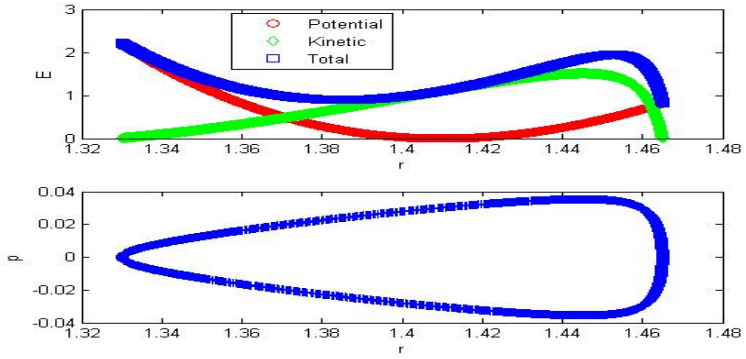

Figure 7: Energy and trajectory by IM, $h = 7.22fs$.

$$\omega_E = S\sqrt{\frac{2(D-E)}{\mu}}. \tag{18}$$

As noted in [1], an Implicit Midpoint (IM) method affects the frequency of the periodic motion of the system. The effective frequency, ω^{eff}, becomes timestep dependent:

$$\omega_E^{eff} = \frac{2}{h}\tan^{-1}(\frac{1}{2}\omega_E h) \tag{19}$$

where h is the timestep. As $h \to 0, \omega_E^{eff} \to \omega_E$, and as $h \to \infty, \omega_E^{eff} \to 0$.

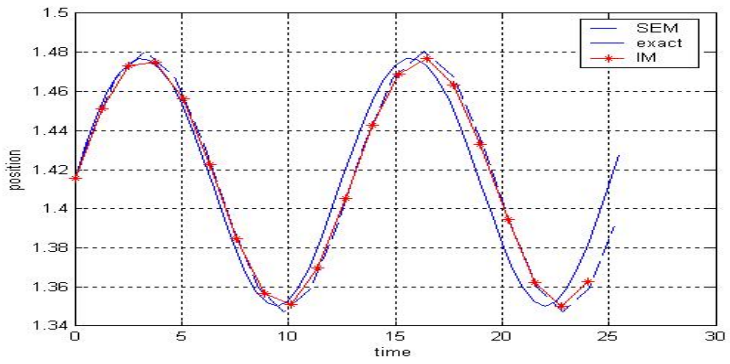

Figure 8: Position vs. time, $h_{av} = 1.266fs$.

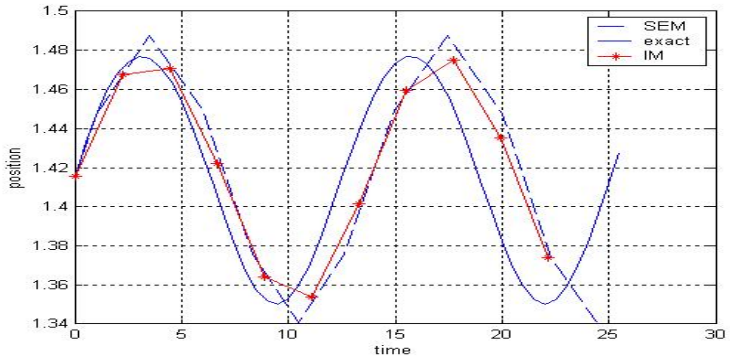

Figure 9: Position vs. time, $h_{av} = 2.2173fs$.

From Fig. 8 and Fig. 9 we can see that the SEM method has the same effect on the period as IM, where the smooth solid lines represent the exact solution. Note, if one fixes the timestep and uses only the discrete Euler-Lagrange equation to propagate the motion, the equation can be rewritten as:

$$\frac{\boldsymbol{V}_{k+1} - \boldsymbol{V}_{k-1}}{2h} = \boldsymbol{M}^{-1}\frac{1}{2}\left[\boldsymbol{f}\left(\frac{1}{2}(\boldsymbol{q}_k + \boldsymbol{q}_{k+1})\right) + \boldsymbol{f}\left(\frac{1}{2}(\boldsymbol{q}_{k-1} + \boldsymbol{q}_k)\right)\right] \tag{20}$$

$$\frac{\boldsymbol{q}_{k+1} - \boldsymbol{q}_k}{h} = \frac{1}{2}(\boldsymbol{V}_k + \boldsymbol{V}_{k+1}) \tag{21}$$

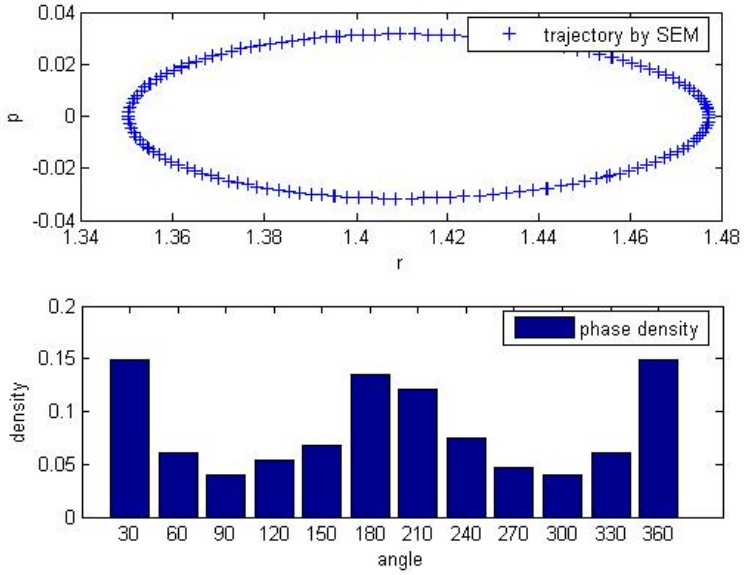

Figure 10: Phase density from the SEM algorithm, $h_{av} = 16.3fs$.

$$\frac{\boldsymbol{q}_k - \boldsymbol{q}_{k-1}}{h} = \frac{1}{2}(\boldsymbol{V}_{k-1} + \boldsymbol{V}_k). \tag{22}$$

This is equivalent to an Implicit Midpoint method.

4. *Phase Density*. The SEM algorithm can produce different phase densities than exact ones, see Fig.10 and 11. This can be explained as follows. When tracing trajectories in phase space, the SEM method always uses the position coordinates and velocities associated with the mid-points between two nearby grid points, i.e. $\boldsymbol{q} = (\boldsymbol{q}_k + \boldsymbol{q}_{k+1})/2, \boldsymbol{v} = (\boldsymbol{q}_{k+1} - \boldsymbol{q}_k)/h_k$. This leads to phase density errors unless timesteps are quite small; see Fig. 12. If we instead use position coordinates of each grid point then we would have trouble in evaluating the corresponding velocities because of nonuniform timesteps. Therefore, one can expect different phase densities even though the numerical results are actually satisfactory and a good trajectory has been obtained in phase space. The computation of phase density has to be done carefully to ensure good results.

5. *Timestep History*. From Fig. 13, we can see that timestep first increased and then oscillates around an average value. Timestep history depends on the

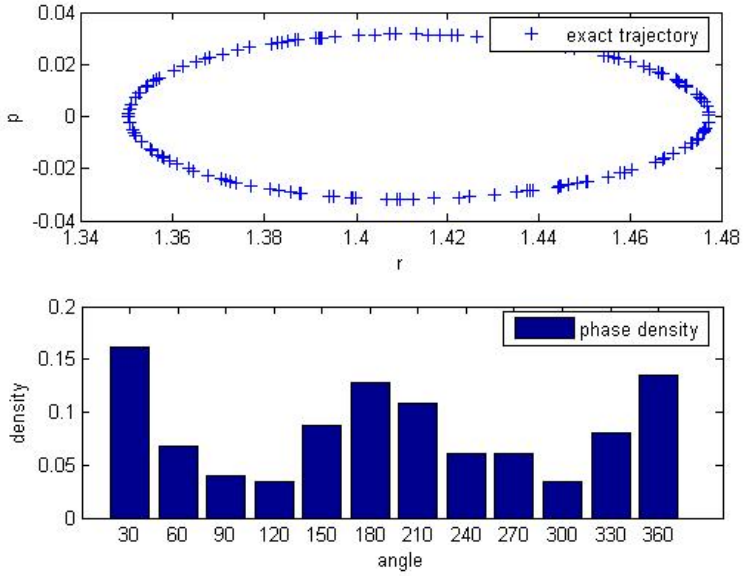

Figure 11: Phase density by exact solution, $h_{av} = 16.3fs$.

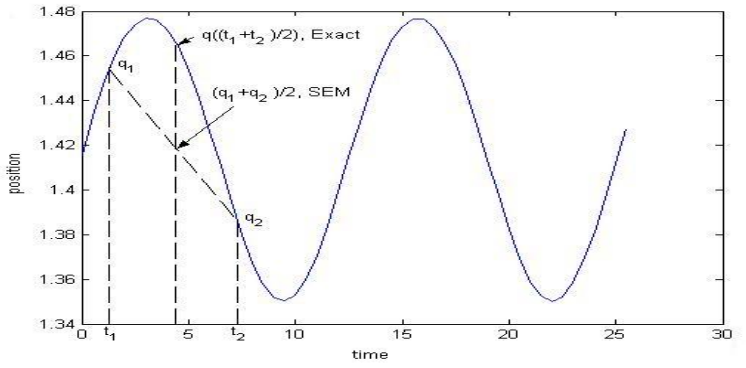

Figure 12: Phase density error caused by the SEM algorithm.

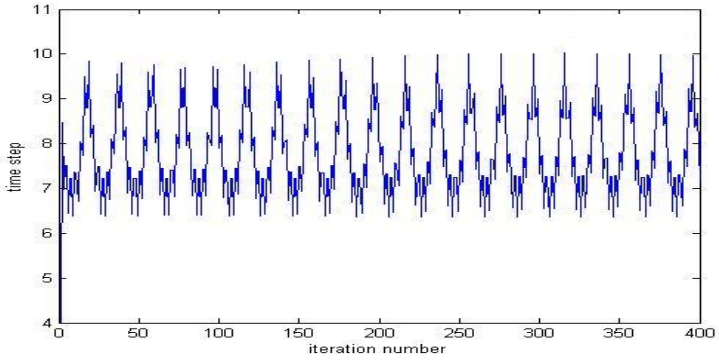

Figure 13: Timestep history of SEM, $h_0 = 4fs$.

exact system and initial conditions.

6. *Dependence on Optimization*. The accuracy of the numerical results by the SEM method strongly depends on the tolerance parameter specified in the optimization package because of the complex landscape of the objective function. This will be described in section 6.

4 Application to a 22-atom System

4.1 Model Description[21]

The second model system is $N-acetylalanyl-N^{'}-methylamide$, see Fig.14. Its chemical composition is given by $CH_3-CO-NH-CHCH_3-CO-NH-CH_3$. This 22-atom system contains representative characteristics of polypeptides (such as main chain dihedral-angle motion) and is particularly flexible, making it a good test case.

Compared with the HBr system, the calculation of forces and potential energy for this system are much more complicated. Therefore, a commercially available molecular dynamics program, CHARMM (version c28b2)[4], is used and the SEM algorithm is integrated into it. The total potential energy can be expressed as

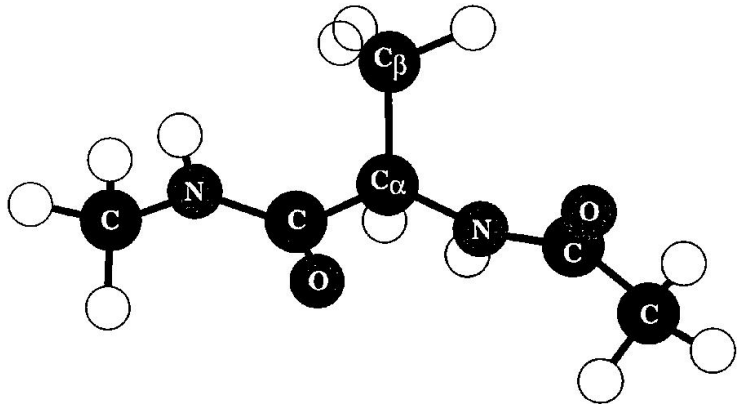

Figure 14: Blocked alanine ('dipeptide') model[21].

$$
\begin{aligned}
E = & \underbrace{E_b + E_\theta + E_\phi + E_\omega}_{internal\ potential} \\
& + \underbrace{E_{vdw} + E_{el} + E_{hb}}_{external\ potential} \\
& + \underbrace{E_{cr} + E_{c\phi}}_{constraints}
\end{aligned} \tag{23}
$$

In this equation, the first line represents internal potentials with E_b denoting bond potentials, E_θ bond angle potentials, E_ϕ torsion potentials and E_ω improper torsions. The second line represents external potentials with E_{vdw} denoting van der Waals interactions, E_{el} electrostatic potentials and E_{hb} hydrogen bonding. The third line represents constraints with E_{cr} denoting atom harmonics and $E_{c\phi}$ dihedral constraints. The mathematical expression for each of these potential terms can be found in [4].

The initial positions are usually based on known X-ray structure (followed by energy minimization to relieve local strain due to non-bonded overlaps and distortions). The initial velocities are typically assigned based on a Maxwellian distribution at some relatively low temperature and then the system is slowly heated to the desired simulation temperature. The actual dynamics simulations then start from that point.

Additionally, CHARMM has its own Verlet time stepping algorithm and can

be used for comparisons. Note that this problem has no exact solution, and small step size (say, $h = 0.5fs$ or $h = 1fs$) simulations of Verlet in CHARMM were employed for pseudo accuracy checks.

The discrete equations have the same forms as (13) and (14). The only difference is that the potential and forces have no explicit expressions and have to be computed by CHARMM in every step. As a consequence, the subsequent optimization will be executed on an objective function without explicit expressions.

4.2 Numerical Results and Discussions

1. *Potential Energy and Total Energy*. Since the exact solutions for trajectories are not available, we use CHARMM (Verlet) with a small stepsize to study the properties of the SEM method.

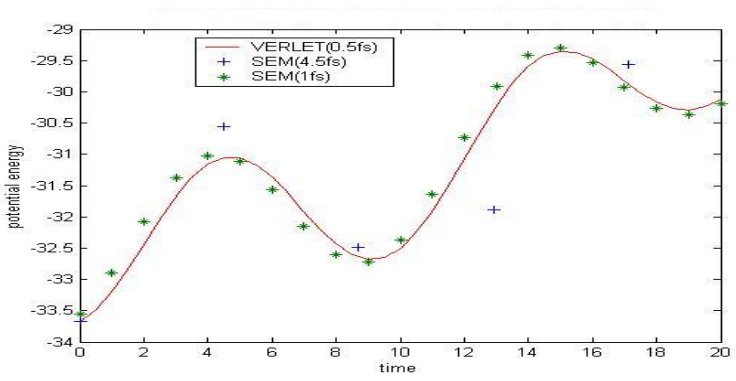

Figure 15: Potential energy history by Verlet and SEM.

It can be seen from Fig.15 that the SEM method is accurate when a small average timestep is used. The potential energy history is consistent with that computed by CHARMM. When the timestep increases, we observed some discrepancy in the potential energy (up to 10 percent in terms of relative errors). This, however, may result from different ways in which energy and velocities are computed.

Verlet:

$$\boldsymbol{V}_n = \frac{\boldsymbol{q}_{n+1} - \boldsymbol{q}_{n-1}}{2h_n} \tag{24}$$

$$\mathbf{E}_p = \mathbf{E}_p(\boldsymbol{q}_n) \tag{25}$$

SEM:

$$\boldsymbol{V}_n = \frac{\boldsymbol{q}_{n+1} - \boldsymbol{q}_n}{h_n} \tag{26}$$

$$\mathbf{E}_p = \mathbf{E}_p\left(\frac{\boldsymbol{q}_{n+1} + \boldsymbol{q}_n}{2}\right) \tag{27}$$

It should be noted that when the average timestep is very small this does not make much difference.

The consistency of potential energy history implies that the computed trajectories are reliable though it is not a very strict test. It can also be seen from Fig.16 that total energy is preserved and no instability occurs.

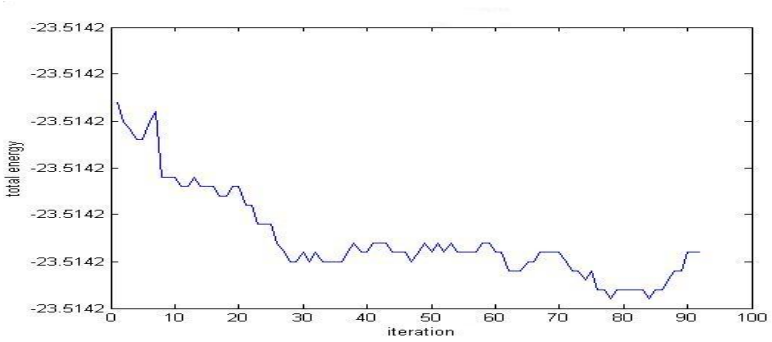

Figure 16: Total energy history by SEM with a timestep history as shown in Fig.17.

2. *Timestep History*. Fig.17 shows the timestep history. Again, the history depends on initial conditions. Fig. 18 shows a different sequence of timesteps caused by different initial conditions.

5 Optimization Techniques

The performance of the algorithm depends highly on how well we can solve the two governing discrete equations which are typically nonlinear. The technique used is the following optimization technique. Given $h_0, \boldsymbol{q}_0, \boldsymbol{q}_1$ we have to find $h_1, \boldsymbol{q}_2$ which are determined by the DEL equations

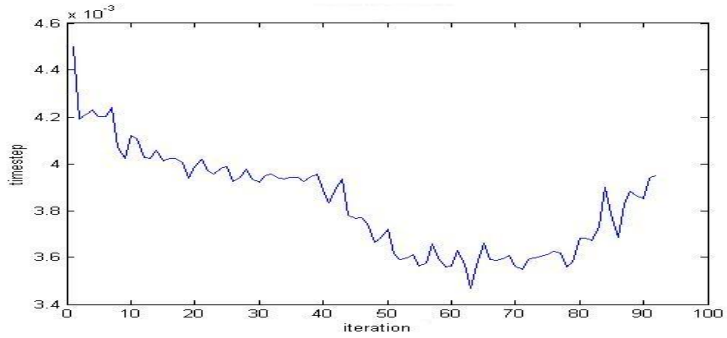

Figure 17: Timestep history of SEM (ps).

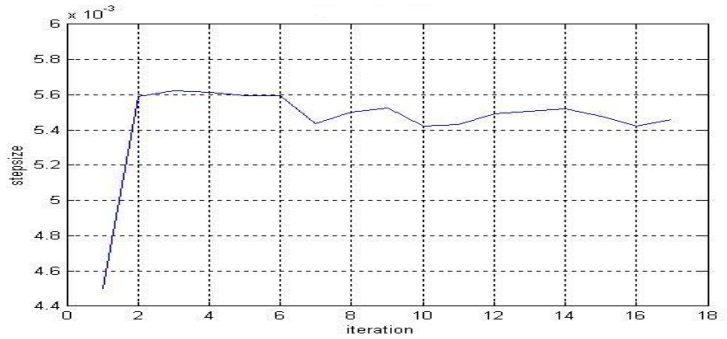

Figure 18: Timestep history of SEM (ps).

$$g\left(\boldsymbol{q}_0, \boldsymbol{q}_1, \boldsymbol{q}_2, h_0, h_1\right) := h_0 D_2 L\left(\boldsymbol{q}_0, \boldsymbol{q}_1, h_0\right) + h_1 D_1 L\left(\boldsymbol{q}_1, \boldsymbol{q}_2, h_1\right) = 0 \tag{28}$$

and the energy condition

$$f\left(\boldsymbol{q}_0, \boldsymbol{q}_1, \boldsymbol{q}_2, h_0, h_1\right) := E\left(\boldsymbol{q}_1, \boldsymbol{q}_2, h_1\right) - E\left(\boldsymbol{q}_0, \boldsymbol{q}_1, h_0\right) = 0. \tag{29}$$

The technique is to minimize the quantity[10]

$$\Psi = \left[f\left(\boldsymbol{q}_0, \boldsymbol{q}_1, \boldsymbol{q}_2, h_0, h_1\right)\right]^2 + \left[g\left(\boldsymbol{q}_0, \boldsymbol{q}_1, \boldsymbol{q}_2, h_0, h_1\right)\right]^2 \tag{30}$$

with respect to the variables $h_1, \boldsymbol{q}_2$, with the other variables given, and subject to the constraint $h_1 > 0$. This is a non-convex optimization problem with a unilateral constraint.

For the objective function associated with this problem, many schemes such as Nelder-Mead Simplex, Truncated-Newton, Quasi-Newton, Genetic Algorithm and Stochastic global optimization were tried. This section will briefly summarize these methods within the context of the problem at hand.

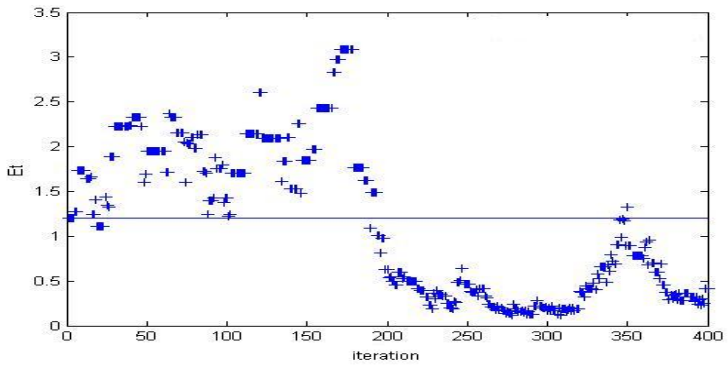

Figure 19: Energy vs. iteration number by SEM, $h_0 = 4fs, TOL = 1e-8$.

The Nelder-Mead Simplex method[15][5] was successfully employed to deal with the HBr system. The Matlab optimization toolbox function 'fminsearch' was used to search for the local minimizer of the objective function. A typical TOL parameter used was 1e-18. A larger TOL parameter such as 1e-8 can generate a bad energy history; see Fig. 19. However, the simplex method is not applicable to the 22-atom system, because without gradient information it becomes very time-consuming when the objective function has more variables. A simple test that one can try uses the extended rosenbrock function ('banana function') which is a part of the standard test suite for function 'fminsearch' in MATLAB

$$F(\boldsymbol{X}) = \sum_{j=1,3,5,\ldots,n-1} (1 - x_j)^2 + 100(x_{j+1} - x_j^2)^2. \tag{31}$$

The Nelder-Mead Simplex method shows an extremely low efficiency when n has a large value.

Truncated-Newton and Quasi-Newton methods demonstrate higher efficiency for minimizing multivariate functions. The main difference between these two is that the former allows a nonzero residual vector $\boldsymbol{r}_k = \boldsymbol{H}_k\boldsymbol{p} + \boldsymbol{g}_k$ in the solution for $\boldsymbol{p}_k$, where $\boldsymbol{H}$ and $\boldsymbol{g}$ denote the Hessian matrix and the gradient

vector, respectively, and $\boldsymbol{p}$ the search direction, while the latter approximates the Hessian $\boldsymbol{H}$ when it is difficult to obtain. A Truncated-Newton method leads to a doubly-nested iteration structure: for every outer Newton iteration k (associated with $\boldsymbol{X}_k$, the vector of unknowns), there corresponds an inner loop (for $\boldsymbol{p}_k$). For a Quasi-Newton method, the search direction is obtained by

$$\boldsymbol{p}_k = -\boldsymbol{D}^k \boldsymbol{g}_k \tag{32}$$

where $\boldsymbol{D}^k$ is a positive definite matrix approaching the inverse Hessian, which may be adjusted from one iteration to the next so that the direction $\boldsymbol{p}_k$ tends to approximate the Newton direction. In the most popular class of Quasi-Newton methods, the matrix $\boldsymbol{D}^{k+1}$ is obtained from $\boldsymbol{D}^k$, and the vector $\boldsymbol{X}_{k+1} - \boldsymbol{X}_k$ and $\boldsymbol{g}_{k+1} - \boldsymbol{g}_k$. Different ways to build curvature information into the matrix $\boldsymbol{D}^k$ lead to different methods. The BFGS method is one such method.

TNPACK is a Fortran package for unconstrained optimization which implements a Truncated-Newton algorithm [19, 20]. This package was applied to minimize the objective function associated with the IM method, HBr system [1], which is nearly quadratic for small timesteps. For the SEM algorithm, L-BFGS-B [24], another package which implements a Quasi-Newton algorithm, is a better choice because this package does constrained optimization such that we can enforce the constraint $h > 0$ conveniently and furthermore, the objective function associated with the SEM algorithm is highly nonlinear. It was applied on the HBr system and the same results were obtained as from the Simplex method. A sample output by L-BFGS-B is shown in Appendix 1, where it can be seen that the optimized function value is 10 orders of magnitude less than the initial function value (6.277e-13 versus 7.641e-3); See Fig.20 for the objective function history. Both of these packages incorporate a line search algorithm [16] which guarantees a global convergence to the closest local minimizer. This step-length procedure, based on safeguarded cubic and quadratic interpolation, is used to search for λ that satisfies both the sufficient decrease condition

$$F(\boldsymbol{X}_k + \lambda \boldsymbol{p}_k) \leq F(\boldsymbol{X}_k) + F_{tol} \lambda \boldsymbol{g}_k^T \boldsymbol{p}_k \tag{33}$$

and the condition regarding sufficient reduction in the magnitude of the directional derivative at $\boldsymbol{X}_{k+1} = \boldsymbol{X}_k + \lambda \boldsymbol{p}_k$:

$$|\boldsymbol{g}_{k+1}^T \boldsymbol{p}_k| \leq \boldsymbol{G}_{tol} |\boldsymbol{g}_k^T \boldsymbol{p}_k|. \tag{34}$$

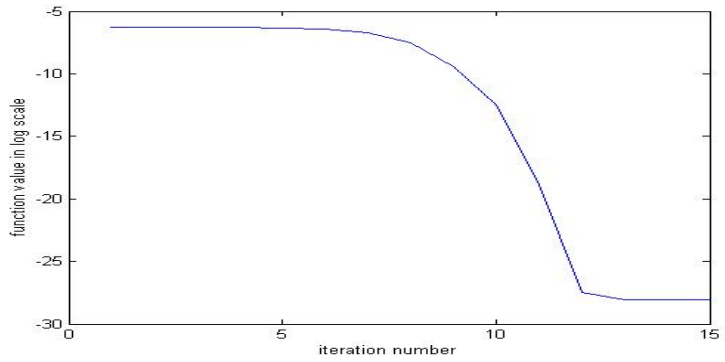

Figure 20: Objective function value history by L-BFGS-B.

In principle, we should look for the global minimizer of the objective function associated with the SEM algorithm. In practice, for the HBr system the closest local minimizer is satisfactory since it makes the function value small enough. However, this is not the case for the 22-atom system where the global minimizer is needed. A Genetic Algorithm is a possible choice for global optimization since it is a derivative-free method[8, 25, 26], but it does not work as well as expected. The main difficulty lies in the fact that the objective function is highly non-convex and has many 'deep valleys'. As a result, extremely narrow bounds have to be specified, otherwise the resulting global minimizer always returns to the starting point and the true solution is missed. Moreover, the bounds initially set for a Genetic algorithm impose a strict limit to the searching space. However, in numerical applications one may need to go slightly beyond this limit and accept a better solution though it falls outside the range. In this problem, for instance, it is not easy to initially choose an appropriate compact set for the search space as is required by these algorithms.

As an alternate derivative-free method, a stochastic method for global optimization is used instead [3]. The main advantage of this algorithm is that it incorporates a local search for local minima and provides a global minimum thereafter. Furthermore, the given limits are not treated as strict limits in the Fortran package implementing the algorithm[6]. The local search may step outside the search range. This is a nice feature and makes it more reliable than a Genetic algorithm for this particular problem. Its application is successful and the objective function value typically goes down from the order of 1e-1 to 1e-17 after iteration completes, while a Genetic algorithm only makes

it down to 1e-3. A sample output of a global optimization can be found in Appendix 2, which tabulates the optimized function values followed by the minimizers (37 variables in this example).

6 Conclusions

The SEM algorithm examined is a symplectic, momentum preserving and energy preserving variational integrator. It can be looked at as the implicit midpoint method combined with a constraint that energy be conserved from step to step. As a result, it has adaptive timesteps and can be expected to increase molecular dynamics simulation times to some extent. Furthermore, it can preserve energy and avoid resonance induced by constant timestep algorithms which are currently used. For a HBr system, we can see that it produces accurate trajectories by comparing with the exact ones in phase space. These are the main advantages over other algorithms. Its disadvantage lies in the fact that it is computationally more expensive, because it needs to solve a more complicated nonlinear system of equations, especially for large scale problems.

Optimization techniques are used to solve the nonlinear system mentioned above. Different algorithms (Simplex, Truncated-Newton, Quasi-Newton, Genetic algorithm, Stochastic global optimization) were tried and satisfactory results can be achieved. For local minimization, Quasi-Newton and Truncated-Newton methods are more efficient than the Simplex method if the gradient can be evaluated well. For global minimization, the stochastic method mentioned in the report shows much higher efficiency than a generic algorithm because it makes use of gradient information and incorporates a local search. Moreover, for the SEM algorithm, local optimization techniques suffice for small system, but large systems require global optimization.

7 Acknowlegement

The authors would like to gratefully thank Dr. Prashanth Vijalapura for his important suggestions for this work. His comments on earlier draft are also very helpful.

References

[1] M. Mandziuk, T. Schlick, *Chem. Phys. Lett.*, 237(1995), 525

[2] V. Arnold, *Mathematical Methods of Classical Mechanics*, Springer-Verlag

[3] C. Boender, A. Kan, G. Timmer, *Math. Program.*, 22(1982), 125

[4] B. Brooks, *et.al*, *J. Comp. Chem.*, 4,2(1983), 187

[5] D. Byatt, I. Coope, C. Price, *40 Years of the Nelder-Mead Algorithm*, Univ. of Canterbury, New Zealand, 2003

[6] T. Csendes, *Short Description of the Global Optimization Subroutine*, Hungary.

[7] Z. Ge, J. Marsden, *Phys. Lett. A* 133(1988), 134.

[8] D. Goldberg, K. Deb, *Comp. Meth. Appl. Mech. Engrg*, 186(2000), 121

[9] E. Hairer, C. Lubich, G. Wanner, *Geometric Numerical Integration*, Springer

[10] C. Kane, J. Marsden, M. Ortiz, *J. Math. Phys.*, 40(1999), 3353

[11] C. Kane, J. Marsden, M. Ortiz, M. West, *Int. J. Numer. Mech. Engng.*, 49(2000), 1295

[12] J. Marsden, G. Patrick, S. Shkoller, *Comm. Math. Phys.*, 199(1998), 351

[13] J. Marsden, W. West, *Acta Numerica*, 2001, 357

[14] W. Mascarenhas, *Math. Program.* Ser.A, 99(2004), 49

[15] J. Mathews, K. Fink, *Numerical methods using MATLAB*, Upper Saddle River, N.J, 4th edition

[16] J. More, D. Thuente, *ACM Trans. Math. Software*, 20,3(1994), 286

[17] S. Nash, J. Nocedal, *SIAM J. Optimization*, 1,3(1991), 358

[18] J. Sanz-Serna, M. Calvo, *Numerical Hamiltonian Problems*, Chapman and Hall, London, 1994

[19] T. Schlick, A. Fogelson, *ACM Trans. Math. Software*,18,1(1992), 46

[20] T. Schlick, A. Fogelson, *ACM Trans. Math. Software*,18,1(1992), 71

[21] T. Schlick, M. Mandziuk, R. Skeel, K. Srinivas, *J. Comp. Phys.*, 140(1998), 1

[22] N, Slater, *Nature*, 180(1957), 1352

[23] J. Wendlandt, J. Marsden, *Phys. D*, 106(1997), 223

[24] C. Zhu, *et.al*, *Tech. Report*, Northwestern Univ. 1994

[25] T. Zohdi, *Me287 Classnotes*, Univ. of California, Berkeley, Spring 2004

[26] T. Zohdi, P. Wriggers, *Computational Modeling and Design of Random Microheterogeneous Materials*, Springer, 2003

[27] X. Zou, *et.al*, *SIAM J. Optimization*, 3,3(1993), 582

Appendix 1

```
===============================================================
sample output of BFGS code (1 iteration) (HBr model, SEM algorithm):
===============================================================
RUNNING THE L-BFGS-B CODE
it    = iteration number
nf    = number of function evaluations
nint  = number of segments explored during the Cauchy search
nact  = number of active bounds at the generalized Cauchy point
sub   = manner in which the subspace minimization terminated:
        con = converged, bnd = a bound was reached
itls  = number of iterations performed in the line search
stepl = step length used
tstep = norm of the displacement (total step)
projg = norm of the projected gradient
f     = function value
Machine precision = 2.220E-16
 N =  3    M =  5

   it   nf  nint  nact  sub  itls  stepl    tstep     projg        f
    0    1     -     -   -     -     -        -     3.034E+00  7.641E-03
    1    3     1     0  ---    1  1.3E-03  3.8E-03  3.161E-02  1.880E-03
    2    4     1     0  con    0  1.0E+00  4.2E-05  3.167E-02  1.878E-03
    3    5     1     0  con    0  1.0E+00  6.1E-04  8.380E-02  1.864E-03
    4    6     1     0  con    0  1.0E+00  1.4E-03  1.783E-01  1.836E-03
    5    7     1     0  con    0  1.0E+00  4.4E-03  3.479E-01  1.757E-03
    6    8     1     0  con    0  1.0E+00  1.1E-02  5.748E-01  1.574E-03
    7    9     1     0  con    0  1.0E+00  2.5E-02  8.235E-01  1.175E-03
    8   10     1     0  con    0  1.0E+00  4.3E-02  8.429E-01  5.509E-04
    9   12     1     0  con    1  5.2E-02  3.5E-03  1.907E-02  8.257E-05
   10   13     1     0  con    0  1.0E+00  2.7E-02  5.974E-03  3.629E-06
   11   14     1     0  con    0  1.0E+00  4.8E-03  1.078E-03  6.390E-09
   12   15     1     0  con    0  1.0E+00  1.8E-04  7.393E-06  1.151E-12
   13   16     1     0  con    0  1.0E+00  1.7E-06  1.805E-09  6.277E-13
   14   17     1     0  con    0  1.0E+00  7.7E-10  3.549E-10  6.277E-13
   15   18     1     0  con    0  1.0E+00  4.0E-11  2.790E-09  6.277E-13

CONVERGENCE: REL_REDUCTION_OF_F <= FACTR*EPSMCH
```

Appendix 2

```
================================================================================
sample output of GLOBAL Minimization code (22-atom model, SEM algorithm):
================================================================================
5000 FUNCTION EVALUATIONS USED FOR SAMPLING
 *** THE LOCAL MINIMUM NO.  1: 0.61483705E-17, NFEV= 5839
 0.61483705E-17
    -36.875850     -28.091865      80.294103     -36.152512     -27.287214
     79.217775     -35.473275     -26.251896      79.502513     -36.166213
    -27.792084      78.027079     -36.573597     -28.645941      77.763815
    -35.665203     -27.123762      76.783571     -34.308545     -27.620819
     76.301469     -36.627671     -27.162202      75.613049     -37.387373
    -28.140131      75.435148     -36.818639     -26.086403      74.867283
    -36.623818     -25.141862      75.092929     -37.557768     -26.205655
     73.629116     0.41925717E-02
 NEW SEED POINT ADDED TO THE CLUSTER NO.  1, NFEV= 4732
 0.78162246E-16
    -36.913577     -28.090029      80.266953     -36.157653     -27.294592
     79.186917     -35.455510     -26.235450      79.503109     -36.137937
    -27.819550      78.013376     -36.619895     -28.637995      77.868729
    -35.687544     -27.154299      76.751197     -34.314566     -27.627738
     76.298890     -36.608112     -27.156941      75.579505     -37.376426
    -28.156712      75.438381     -36.858471     -26.085053      74.885161
    -36.553829     -25.185238      75.053539     -37.547058     -26.169085
     73.599399     0.49375820E-02
       ...            ...             ...            ...            ...
       ...            ...             ...            ...            ...
 LOCAL MINIMA FOUND:
 0.61483705E-17
    -36.875850     -28.091865      80.294103     -36.152512     -27.287214
     79.217775     -35.473275     -26.251896      79.502513     -36.166213
    -27.792084      78.027079     -36.573597     -28.645941      77.763815
    -35.665203     -27.123762      76.783571     -34.308545     -27.620819
     76.301469     -36.627671     -27.162202      75.613049     -37.387373
    -28.140131      75.435148     -36.818639     -26.086403      74.867283
    -36.623818     -25.141862      75.092929     -37.557768     -26.205655
     73.629116     0.41925717E-02
```