

# Symplectic algorithm for constant-pressure molecular dynamics using a Nosé–Poincaré thermostat

Jess B. Sturgeon and Brian B. Laird<sup>a)</sup>

*Department of Chemistry and Kansas Institute for Theoretical and Computational Science, University of Kansas, Lawrence, Kansas 66045*

(Received 20 July 1999; accepted 24 November 1999)

We present a new algorithm for isothermal–isobaric molecular-dynamics simulation. The method uses an extended Hamiltonian with an Andersen piston combined with the Nosé–Poincaré thermostat, recently developed by Bond, Leimkuhler, and Laird [J. Comp. Phys. **151**, 114 (1999)]. This Nosé–Poincaré–Andersen (NPA) formulation has advantages over the Nosé–Hoover–Andersen approach in that the NPA is Hamiltonian and can take advantage of symplectic integration schemes, which lead to enhanced stability for long-time simulations. The equations of motion are integrated using a generalized leapfrog algorithm (GLA) and the method is easy to implement, symplectic, explicit, and time reversible. To demonstrate the superior stability of the method we show results for test simulations using a model for aluminum and compare it to a recently developed time-reversible algorithm for Nosé–Hoover–Anderson. In addition, an extension of the NPA to multiple time steps is outlined and a symplectic and time-reversible integration algorithm, based on the GLA, is given.

© 2000 American Institute of Physics. [S0021-9606(00)51307-5]

## I. INTRODUCTION

Traditionally, molecular-dynamics simulations are performed using constant particle number  $N$ , volume  $V$ , and energy  $E$ . However, these are not usually the conditions under which experiments are done and there has been much attention to the development of simulation methods designed to sample from other, experimentally more relevant ensembles, such as constant temperature (canonical) and/or constant pressure.<sup>1–4</sup> Some of the most popular and useful of these are those based on so-called “extended” Hamiltonians, i.e., Hamiltonians in which extra degrees of freedom have been added to the system in order to ensure that the trajectory samples from the statistical distribution corresponding to the desired thermodynamic conditions.

For a constant pressure system, for example, Andersen<sup>5</sup> introduced the volume  $V$ , along with its corresponding conjugate momentum  $\pi_V$ , as extra variables. The new variables are coupled to the system in such a way as to guarantee that the trajectory (if ergodic) samples from an isobaric statistical distribution. Similarly, to generate a constant temperature distribution, Nosé<sup>6</sup> introduced a new mechanical variables  $s$  (with conjugate momentum  $\pi_s$ ) that couples into the system through the particle momenta and acts to effectively rescale time in such a way as to guarantee canonically distributed configurations. These two extensions can be combined to give a Hamiltonian whose trajectories can be shown to sample from an isothermal–isobaric ensemble.<sup>7</sup>

This combined Nosé–Andersen (NA) Hamiltonian is given by

$$\mathcal{H}_{\text{NA}} = V^{-2/3} \sum \frac{p_i^2}{2m_i s^2} + U(V^{1/3} \mathbf{q}) + \frac{\pi_V^2}{2Q_V} + \frac{\pi_s^2}{2Q_s} + gkT \ln s + P_{\text{ext}} V, \quad (1)$$

where  $p_i$  is the conjugate momentum to the scaled position  $q_i = V^{-1/3} r_i$ ,  $P_{\text{ext}}$  is the external pressure and  $g$  is given by  $N_f + 1$  where  $N_f$  is number of degrees of freedom of the original system. The quantities  $Q_V$  and  $Q_s$  are the masses of the Andersen “piston” and the Nosé thermostat, respectively. [Note that this is not the only formulation possible. For example, one can use the strain, proportional to  $\ln V$ , instead of the volume as an independent extended variable. This leads to the sampling by the trajectory of a slightly different distribution, but the difference between averages calculated using these distributions is generally small,  $\mathcal{O}(1/N)$ .<sup>8</sup>]

The equations of motion for this system are

$$\dot{p}_i = -V^{1/3} \nabla_i U(V^{1/3} \mathbf{q}), \quad (2a)$$

$$\dot{q}_i = \frac{p_i}{s^2 m_i V^{2/3}}, \quad (2b)$$

$$\dot{\pi}_V = \mathcal{P} - P_{\text{ext}}, \quad (2c)$$

$$\dot{V} = \pi_V / Q_V, \quad (2d)$$

$$\dot{\pi}_s = V^{-2/3} s^3 \sum \frac{p_i^2}{m_i} - \frac{gkT}{s}, \quad (2e)$$

$$\dot{s} = \pi_s / Q_s, \quad (2f)$$

where the instantaneous pressure  $\mathcal{P}$  is given by

$$\mathcal{P} = \frac{2}{3V} \sum \frac{p_i^2}{2m_i V^{2/3} s^2} - \frac{1}{3V} \sum \frac{\partial U}{\partial q_i} q_i. \quad (3)$$

<sup>a)</sup> Author to whom correspondence should be addressed; electronic mail: laird@pilsner.chem.ukans.edu

There are two major drawbacks to this approach: First, because of the time rescaling, the time variable in Nosé dynamics is not “real” time, so any discretized trajectory generated by numerically integrating the Nosé equations of motion must be transformed back into real time, leading to the configurations that are spaced at unequal real-time intervals. This is inconvenient for the construction of equilibrium averages, especially of dynamical quantities. Second, the Hamiltonian is not *separable*<sup>9</sup> (that is, the kinetic and potential terms in the Hamiltonian are not functions only of momenta and position variables, respectively), making standard Verlet–leapfrog approaches inapplicable.

By a change of variables and a time rescaling of the equations of motion. Hoover<sup>10</sup> derived new equations of motion that generate the same trajectories (for the exact solution) as the original Nosé Hamiltonian, but in real time. This Nosé–Hoover dynamics has become a standard method in molecular simulation. However, the change of variables that links the Nosé Hamiltonian to the Nosé–Hoover equations of motion is a noncanonical transformation—the total energy function of the system is still conserved, but it is no longer a Hamiltonian, since the equations of motion cannot be derived from it. Although a variety of very good time-reversible methods have been put forward,<sup>11,12</sup> the lack of Hamiltonian structure precludes the use of symplectic integration schemes, which have been shown to have superior stability over nonsymplectic methods.<sup>9</sup>

## II. THE NOSÉ–POINCARÉ–ANDERSEN (NPA) HAMILTONIAN

Recently, Bond, Leimkuhler, and Laird<sup>13</sup> have developed a new formulation of Nosé constant-temperature dynamics in which a Poincaré time transformation is applied directly to the Nosé Hamiltonian, instead of applying a time transformation to the equations of motion as in Nosé–Hoover. The result of this is a method that runs in real time, but is also Hamiltonian in structure. In this work we combine this new thermostat with the Andersen method for constant pressure to give an algorithm for isothermal–isobaric molecular dynamics. For a system with an Andersen piston, the new Nosé–Poincaré–Andersen (NPA) Hamiltonian is given by

$$\mathcal{H}_{\text{NPA}} = [\mathcal{H}_{\text{NA}} - \mathcal{H}_{\text{NA}}(t=0)]s, \quad (4)$$

where  $\mathcal{H}_{\text{NA}}$  is given in Eq. (1). As discussed in Ref. 11, the above form of the Hamiltonian (a specific case of a Poincaré time transformation) will generate the same trajectories as the original Nosé–Andersen Hamiltonian, except with time rescaled by  $s$  (which puts the trajectories back into real time). The resulting equations of motion (except for  $\pi_s$ ) for this constant pressure and temperature Nosé–Poincaré Hamiltonian are the same as those given above for the Nosé–Andersen system [Eqs. (2a)–(2f)], except that the right-hand-side is multiplied by the thermostat variable  $s$ . For  $\pi_s$  we have

$$\dot{\pi}_s = -s \frac{\partial \mathcal{H}}{\partial s} - \Delta H = V^{-2/3} \sum \frac{p_i^2}{m_i s^2} - gkT - \Delta H, \quad (5)$$

where  $\Delta H \equiv \mathcal{H}_{\text{NA}} - \mathcal{H}_{\text{NA}}(t=0)$  and, since the averages are taken over the real-time trajectory, the value of  $g$  is now  $N_f$  to give the correct isothermal–isobaric distribution. It can be shown (see Appendix A) that, assuming ergodicity, the NPA dynamics generates trajectories that sample from an isothermal–isobaric (NPT) statistical distribution. It is also relatively easy to show (see Appendix B) that the NPA equations of motion for  $\pi_v$ ,  $s$ , and  $V$  generate the appropriate virial relations<sup>11</sup> for the NPT distribution, namely,  $\langle \mathcal{P} \rangle = P_{\text{ext}}$  and  $\langle V(\mathcal{P} - P_{\text{ext}}) \rangle + kT = 0$ .

## III. INTEGRATING THE NPA EQUATIONS OF MOTION

The NPA Hamiltonian is nonseparable since the kinetic energy contains the extended “position” variables  $s$  and  $V$ . The equations of motion for a general time-independent, nonseparable Hamiltonian can be written (for general positions  $Q$  and conjugate momenta  $P$ )

$$\dot{Q} = G(P, Q), \quad \dot{P} = F(P, Q), \quad (6)$$

where  $G(P, Q) = \partial \mathcal{H} / \partial P$  and  $F(P, Q) = -\partial \mathcal{H} / \partial Q$ . (For a separable Hamiltonian,  $G$  is only a function of  $P$  and  $F$  is only a function of  $Q$ .) For such a nonseparable system, standard symplectic splitting methods, such as the Verlet–leapfrog algorithm, are not directly applicable. However, symplectic methods specifically for nonseparable systems have been developed.<sup>9</sup> One simple example that is second-order and time-reversible is the generalized leapfrog algorithm (GLA)

$$\begin{aligned} P_{n+1/2} &= P_n + hF(P_{n+1/2}, Q_n)/2, \\ Q_{n+1} &= Q_n + h[G(P_{n+1/2}, Q_n) + G(P_{n+1/2}, Q_{n+1})]/2, \end{aligned} \quad (7)$$

$$P_{n+1} = P_{n+1/2} + hF(P_{n+1/2}, Q_{n+1})/2,$$

where  $h$  is the time step and  $P_n$  and  $Q_n$  are the approximations to  $P(t)$  and  $Q(t)$  at  $t = t_n = nh$ . [This method can be obtained as the concatenation of the symplectic Euler method

$$\begin{aligned} P_{n+1} &= P_n + hF(P_{n+1}, Q_n), \\ Q_{n+1} &= Q_n + hG(P_{n+1}, Q_n). \end{aligned} \quad (8)$$

with its *adjoint*<sup>9</sup>

$$\begin{aligned} Q_{n+1} &= Q_n + hG(P_n, Q_{n+1}), \\ P_{n+1} &= P_n + hF(P_n, Q_{n+1}). \end{aligned} \quad (9)$$

The concatenation of an integrator with its adjoint guarantees a time-reversible method.] This method is a simple example of a class of symplectic integrators for nonseparable Hamiltonians.<sup>14–17</sup>

Applying the GLA to the NPA equations of motion gives

$$p_{i,n+1/2} = p_{i,n} - \frac{h}{2} s_n V_n^{1/3} \nabla_i U(V_n^{1/3} \mathbf{q}_n), \quad (10a)$$

$$\pi_{v,n+1/2} = \pi_{v,n} + \frac{h}{2} s_n [\mathcal{P}(\mathbf{q}_n, \mathbf{p}_{n+1/2}, V_n, s_n) - P_{\text{ext}}], \quad (10b)$$

$$\pi_{s,n+1/2} = \pi_{s,n} + \frac{h}{2} \left( \sum_{i=1}^N \frac{p_{i,n+1/2}^2}{m_i V_n^{2/3} s_n^2} - g k_B T \right) - \frac{h}{2} \Delta \mathcal{H}(\mathbf{q}_n, \mathbf{p}_{n+1/2}, V_n, \pi_{v,n+1/2}, s_n, \pi_{s,n+1/2}), \quad (10c)$$

$$s_{n+1} = s_n + \frac{h}{2} (s_n + s_{n+1}) \frac{\pi_{s,n+1/2}}{Q_s}, \quad (10d)$$

$$V_{n+1} = V_n + \frac{h}{2} (s_n + s_{n+1}) \frac{\pi_{v,n+1/2}}{Q_V}, \quad (10e)$$

$$q_{i,n+1} = q_{i,n} + \frac{h}{2} \left( \frac{1}{s_n V_n^{2/3}} + \frac{1}{s_{n+1} V_{n+1}^{2/3}} \right) \frac{p_{i,n+1/2}}{m_i}, \quad (10f)$$

$$\pi_{s,n+1} = \pi_{s,n+1/2} + \frac{h}{2} \left( \sum_{i=1}^N \frac{p_{i,n+1/2}^2}{m_i V_{n+1}^{2/3} s_{n+1}^2} - g k_B T \right) - \frac{h}{2} \Delta \mathcal{H}(\mathbf{q}_{n+1}, \mathbf{p}_{n+1/2}, V_{n+1}, \pi_{v,n+1/2}, s_{n+1}, \pi_{s,n+1/2}), \quad (10g)$$

$$\pi_{v,n+1} = \pi_{v,n+1/2} + \frac{h}{2} s_{n+1} [\mathcal{P}(\mathbf{q}_{n+1}, \mathbf{p}_{n+1/2}, V_{n+1}, s_{n+1}) - P_{\text{ext}}], \quad (10h)$$

$$p_{i,n+1} = p_{i,n+1/2} + \frac{h}{2} s_{n+1} V_{n+1}^{1/3} \nabla_i U(V_{n+1}^{1/3} \mathbf{q}_{n+1}). \quad (10i)$$

As in the case of the constant volume Nosé–Poincaré algorithm, the GLA for the NPA is explicit—this is not necessarily the case for a general nonseparable Hamiltonian. Note that Eq. (10c) requires the solution of a scalar quadratic equation for  $\pi_{s,n+1/2}$ . Details of how to solve this equation without involving subtractive cancellation can be found in Ref. 11.

Often, in specific practical applications, there is a sufficient separation of time scales between the fastest and slowest motions in the dynamics that a multiple time step approach is prudent. For constant particle number, volume and energy (NVE) simulations, this approach is well developed.<sup>18</sup> For NPT simulations of the type discussed here, multiple time step algorithms exist,<sup>12</sup> but are complicated (mainly due to the lack of Hamiltonian structure in the Nosé–Hoover–Anderson approach). In contrast, the multiple-time step extension of the NPA is straightforward and a modification of the generalized leapfrog algorithm to this system is presented in Appendix C.

#### IV. SIMULATION RESULTS

In order to evaluate this method, simulations were performed using an embedded atom potential for aluminum.<sup>19</sup> Unless otherwise specified all simulations were done on a system of 256 particles with periodic boundary conditions for an aluminum melt at  $T=1000$  K and  $P=0$ . For this model, mass is measured in amu, distance in Å, and energy

in eV. The natural time unit of the simulation is then 10.181 fs; that is, a simulation time step of 0.1 corresponds to an actual time step of 1.0181 fs.

In order to properly sample the isothermal–isobaric distribution, the masses for the extended variables need to be chosen with some care.<sup>20</sup> If the masses are too large or too small, the natural vibrational frequency of the extended variable will lie outside the density of states of vibrational frequencies of the system. This effectively decouples the extended variable from the motions of the system, destroying ergodicity. A useful method to monitor this is to examine the distributions of kinetic energy or instantaneous “temperature” (for the thermostat mass) and density (for the pressure piston). For a system with an isothermal–isobaric distribution, the distribution of the kinetic energy  $\mathcal{K}$  will be a Gaussian with variance  $2\langle\mathcal{K}\rangle^2/3N$ —the distribution of the instantaneous temperature  $\hat{T}$  is then a Gaussian with variance  $2\hat{T}^2/3N$  (Note that, for simplicity, and for the sake of this argument, we have ignored the corrections in these quantities due to the fact that the total momentum is conserved<sup>21</sup> as they are negligible with respect to the calculation of the distribution). For a system, in which the thermostat is decoupled from the system, the thermostat variable motion will be an undamped harmonic oscillation<sup>20</sup>—the resulting distribution is decidedly non-Gaussian, being peaked at the edges (the classical turning points) of the distribution and low around the average value.

In order to determine the proper masses for the thermostat we perform a series of simulations using various thermostat masses while holding the pressure piston mass constant with a value of  $Q_V=0.0001$ . (This initial value of  $Q_V$  was based on previous runs in the development phase of the algorithm.) Each run was initialized to an fcc (face-centered-cube) lattice of initial density  $0.06021 \text{ \AA}^{-3}$  with the individual velocity components chosen from a Maxwell–Boltzmann distribution at 1000 K. The system was equilibrated for 50 000 steps (50.91 ps) and then monitored for 50 000 steps (50.91 ps). [The fluctuations in instantaneous temperature as a function of time for the various thermostat masses are plotted in Fig. 1. The instantaneous temperature  $\hat{T}$  is given by

$$\hat{T} \equiv \frac{2\mathcal{K}}{2(N-1)k_B},$$

where  $\mathcal{K}$  is the instantaneous kinetic energy of the system, and  $k_B$  is Boltzmann’s constant—the factor  $(N-1)$  is used to correct for the fact that in a molecular-dynamics simulation, the total linear momentum is conserved.<sup>21</sup>] This figure shows that for a very small thermostat piston mass,  $Q_s=1.0$ , the system tightly oscillates around the input temperature. This system has a  $\langle(\delta\hat{T})^2\rangle=6.5 \text{ K}^2$ . Similar behavior is seen for  $Q_s=10.0$ . For a very large mass,  $Q_s=1\,000\,000.0$  the same harmonic oscillations are observed. Figure 2 shows the instantaneous temperature distributions. In both of these cases, from the non-Gaussian shape of the distributions, one can see that the system is not properly sampling the isothermal–isobaric distribution. From the formula given

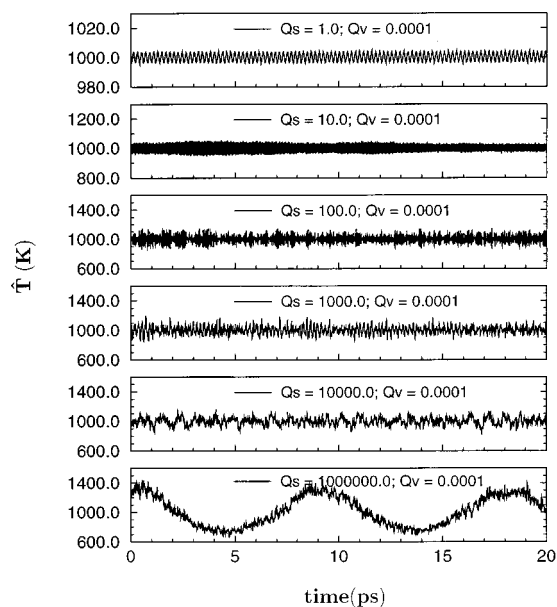


FIG. 1. The instantaneous temperature  $\hat{T}$  of the system plotted as a function of time for six different thermostat masses.

above for the expected distribution of  $\hat{T}$ , we expect for a system of 256 particles at 1000 K, the distribution to be a Gaussian with  $\langle(\delta\hat{T})^2\rangle = 2604.2 \text{ K}^2$ . From this value and an average temperature of 1000 K, a Gaussian was constructed and is plotted together with the distributions in Fig. 2. These curves match the observed distribution for  $Q_s$  in the range of 100.0 to 10 000.0; for subsequent simulations we choose a value of  $Q_s = 500.0$ . A similar series of runs were performed

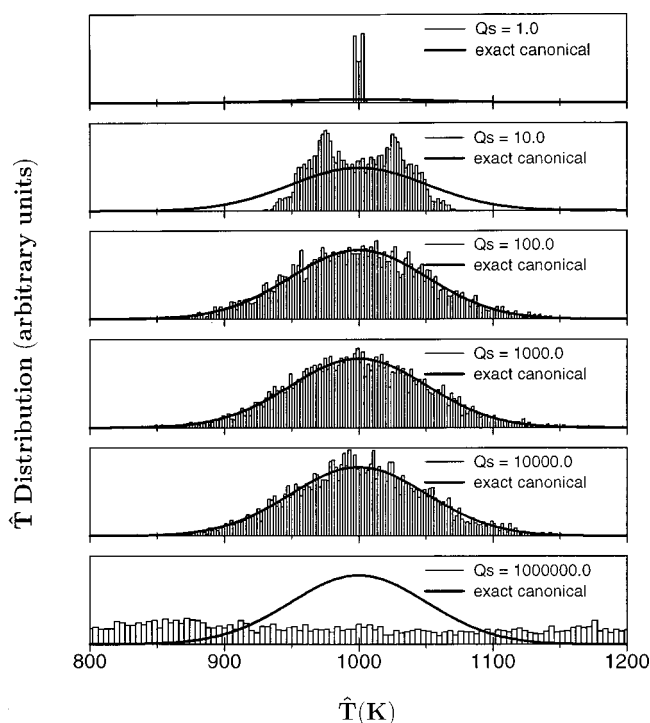


FIG. 2. These are the distributions in the instantaneous temperature  $\hat{T}$  corresponding to the temperature trajectories in Fig. 1. A Gaussian curve with variance  $2T^2/3N$  is shown in each plot for comparison.

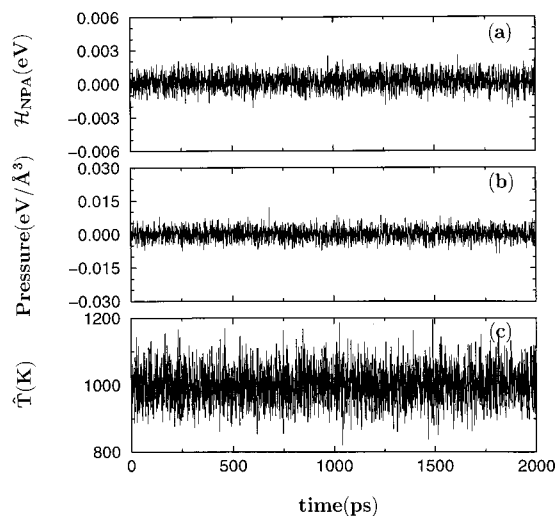


FIG. 3. (a) The value of the NPA Hamiltonian as a function of time for a long run (over 2 ns) on a 256 particle aluminum system. The starting configuration had been equilibrated at  $T = 1000 \text{ K}$  and  $P = 0$ . The values of  $Q_v$  and  $Q_s$  (in reduced units) are 500.0 and  $10^{-4}$ , respectively. The instantaneous pressure and temperature trajectories for this run are given in (b) and (c), respectively.

using  $Q_s = 500.0$  to determine the mass of the pressure piston. A value of  $Q_v = 0.0001$  was determined to be a suitable value for the mass of the pressure piston.

Using these values for the piston masses, the stability of the method was tested. Figure 3(a) shows the value of the NPA Hamiltonian (a conserved quantity) as a function of time for a long run using a time step of 0.1 (1.0181 fs). The trajectory shown here was begun after initial equilibration at 1000 K for  $2 \times 10^6$  time steps (2.036 ns). The stability of the method is excellent, giving no noticeable drift in  $\mathcal{H}_{\text{NPA}}$  over the course of a long trajectory. The pressure and temperature trajectories for this run are also shown in Figs. 3(b) and 3(c), respectively. In Figs. 4(a)–4(c), we examine the ability of the method to regulate the pressure, temperature and density, respectively. This system was initialized to an fcc lattice of initial density  $0.06021 \text{ \AA}^{-3}$  with the individual velocity components chosen from a Maxwell–Boltzmann distribution at 100 K. The simulation was then run with the NPA with  $T = 1000 \text{ K}$  and  $P = 0$ . In all cases, the instantaneous pressure, temperature, and density evolve quickly and stabilize about their desired values.

The GLA has a global error that is second order in the time step. To demonstrate that this also is true in our results, a series of simulations were performed using various values for the time step. The system was initialized in an identical manner to that described in the last paragraph and then run for a total time of 2.036 ps. Figure 5 shows a log–log plot of the energy error, as estimated by the standard deviation of  $\mathcal{H}_{\text{NPA}}$  [Eq. (5)], versus the time step.

In order to demonstrate that the method yields relevant dynamical quantities, the normalized velocity autocorrelation function,  $C(t) = \langle \mathbf{v}(t) \cdot \mathbf{v}(0) \rangle / \langle \mathbf{v}(0) \cdot \mathbf{v}(0) \rangle$ , was calculated using our constant NPT algorithm (with  $Q_v$  and  $Q_s$  as determined above) and compared to the same quantity calculated using standard constant NVE molecular dynamics (using a

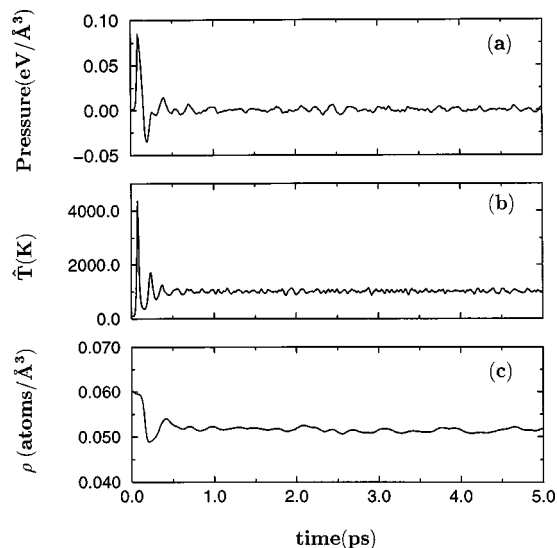


FIG. 4. Pressure (a), temperature (b), and density (c) trajectories for the 256 particle aluminum system using the NPA algorithm with  $T=1000$  K and  $P=0$  starting from an initial configuration in an fcc lattice with density  $\rho=0.06021 \text{ \AA}^{-3}$  and initial velocities chosen from a Maxwell-Boltzmann distribution at 100 K.

velocity-Verlet integrator<sup>22</sup>). Both simulations were run using the same time step of 0.1 (1.0181 fs). The NVE simulations were run at an energy and density corresponding to the average energy and density for the constant NPT simulations. This comparison is shown in Fig. 6. Both systems were first equilibrated at 1000 K for 200 000 steps (203.6 ps) and run for 20 000 steps (20.36 ps) to collect averages.  $C(t)$  for the Nosé-Poincaré-Andersen method for constant NPT molecular dynamics is seen to be nearly indistinguishable from that of the NVE simulation.

Finally, it is useful to compare the method with a state-of-the-art integrator for the Nosé-Hoover-Anderson (NHA) equations of motion, namely that of Martyna *et al.*<sup>13</sup> This method is a time-reversible, but nonsymplectic (due to the nonsymplectic nature of the NHA equations of motion)

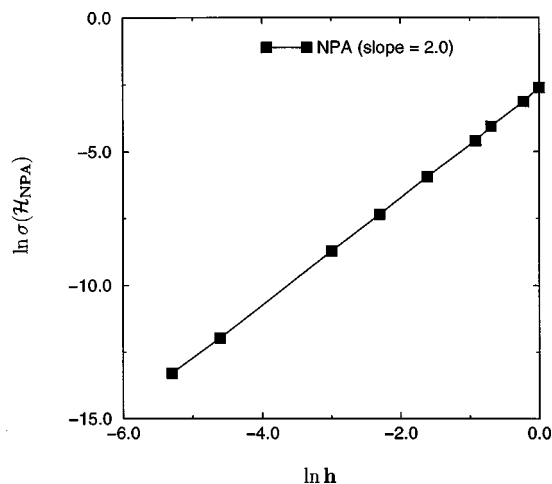


FIG. 5. Log-log plot of the energy error (standard deviation)  $\sigma(\mathcal{H}_{\text{NPA}})$  vs the time step for a variety of piston and thermostat masses. The order of the method is given by the slope of this line which is 2.0, indicating a second-order global error.

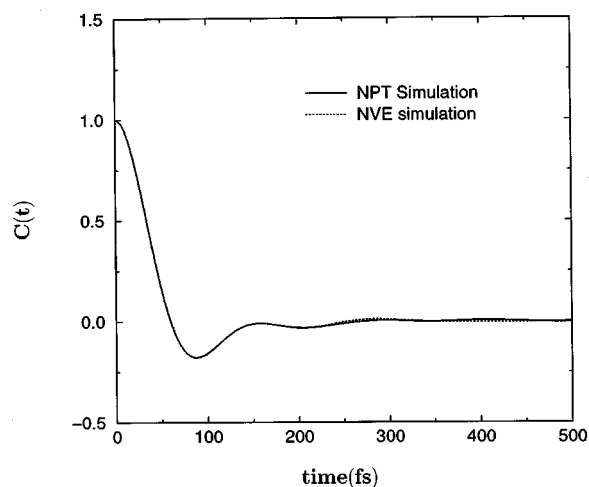


FIG. 6. The normalized velocity autocorrelation function,  $C(t)$ , for the aluminum model calculated for a 256 particle system at  $T=1000$  K and  $P=0$ . The solid line is calculated using the NPA algorithm described, herein. The dotted line is a constant NVE simulation under corresponding conditions.

method that was developed using splitting techniques on the Liouville operator corresponding to the NPA equations of motion, and it is somewhat more complicated to implement than the GLA. Like the GLA, it is a second-order method overall, but the integration of the thermostat variables employ a fourth-order symmetric splitting. To compare the methods we have performed simulations of both systems using identical starting configurations and extended variable masses at 1000 K and zero pressure. Each system was pre-equilibrated for 2.036 ns and then monitored for an additional 4.072 ns. Figure 7 shows the conservation of the NA energy  $\mathcal{H}_{\text{NA}} - \mathcal{H}_{\text{NA}}(t=0)$  for time steps of 0.2 (2.162 fs), 0.4 (4.324 fs), and 0.8 (8.648 fs) in Figs. 7(a)–7(c), respectively. Although both methods show good energy conservation, the GLA, due to its symplectic nature, does not exhibit the drift in energy that can be seen in the time-reversible, but nonsymplectic, method for the NHA.

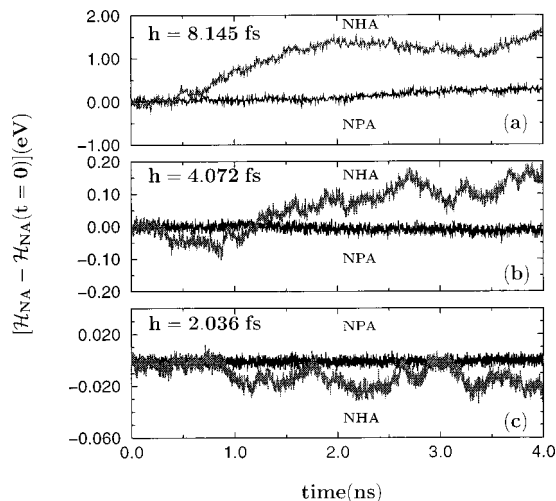


FIG. 7. Deviation of the Nosé-Anderson energy from its initial value for several time steps for the NPA (using GLA integrator) and for the method of Martyna *et al.* for the NHA equations of motion.

### V. SUMMARY

We have developed a real-time Hamiltonian formulation of isothermal–isobaric molecular-dynamics simulation based on a Poincaré time transformation of the Nosé–Anderson Hamiltonian. This new Nosé–Poincaré–Anderson (NPA) method is an extension of the recently developed Nosé–Poincaré method for isothermal molecular dynamics.<sup>13</sup> This new method has advantages over the standard Nosé–Hoover–Anderson (NHA)<sup>10</sup> reformulation of Nosé–Anderson dynamics, in that unlike the NHA, the NPA retains a Hamiltonian structure allowing for the use of symplectic integrators, which have recently proven to have superior stability over nonsymplectic methods. For integrating the equations of motion we use the generalized leapfrog algorithm (GLA), which is time-reversible, second-order, and symplectic. Using a model for aluminum, we have shown that the NPA is extremely stable over long runs and reproduces statistical mechanical properties of the system and distribution well. It is also shown to exhibit less energy drift at large time steps than a state-of-the-art method for integration of the NHA equations of motion. Finally, we have also demonstrated that a multiple-time step extension of the method is straightforward.

### ACKNOWLEDGMENTS

We gratefully acknowledge Steve Bond and Ben Lemkuhler for helpful conversations and invaluable advice, as well as the Kansas Center for Advanced Scientific Computing for the use of their computer facilities. We also would like to thank the National Science Foundation (under Grant Nos. CHE-9500211 and DMS-9627330) as well as the University of Kansas General Research Fund for their generous support of their research.

### APPENDIX A

In this Appendix we demonstrate that, assuming ergodicity, the Nosé–Poincaré–Anderson Hamiltonian generates trajectories in which the coordinates are distributed according to the isothermal–isobaric statistical distribution.

The statistical distribution for the reduced set of variables,  $f(\tilde{p}, \tilde{q}, V)$  is given by an integral over the extended variable distribution  $F_{\text{ext}}(p, q, s, \pi_s, V, \pi_V)$

$$d\tilde{q}d\tilde{p}dVf(\tilde{p}, \tilde{q}, V) = \int ds \int d\pi_s \int d\pi_V dpdq dV \times F_{\text{ext}}(p, q, s, \pi_s, V, \pi_V). \quad (\text{A1})$$

(Here the tildes are used to denote the real, unscaled dynamical variables.) Assuming that the dynamics is ergodic, the NPA dynamics will generate a microcanonical distribution based on the NPA zero-energy surface. Thus,

$$d\tilde{q}d\tilde{p}dVf(\tilde{p}, \tilde{q}, V) = \frac{1}{Z_{\text{NPA}}N!h^{N_f}} \int ds \int d\pi_s \int d\pi_V \times dpdq dV \delta\{s[H_{\text{NA}} - H_{\text{NA}}(0)]\}. \quad (\text{A2})$$

where  $N_f$  is the number of degrees of freedom,  $H_{\text{NA}}(0)$  is the value of the NA Hamiltonian at  $t=0$ , and  $Z_{\text{NPA}}$  is the partition function for the full extended system

$$Z_{\text{NPA}} = \frac{1}{N!h^{N_f}} \int ds \int d\pi_s \int d\pi_V \int dp \int dq \int dV \times \delta\{s[H_{\text{NA}} - H_{\text{NA}}(0)]\}. \quad (\text{A3})$$

Defining the reduced Hamiltonian as

$$H(\tilde{p}, \tilde{q}) = \sum_i \frac{\tilde{p}_i^2}{2m_i} + \Phi(\tilde{\mathbf{q}}), \quad (\text{A4})$$

we have

$$d\tilde{q}d\tilde{p}dVf(\tilde{p}, \tilde{q}, V) = \frac{1}{Z_{\text{NPA}}N!h^{N_f}} \int ds \int d\pi_s \int d\pi_V \times dpdq dV \delta\left\{s \left[ H(V^{-1/3}s^{-1}\mathbf{p}, V^{1/3}\mathbf{q}) + \frac{\pi_s^2}{2Q_s} + \frac{\pi_V^2}{2Q_V} + gkT \ln s + P_{\text{ext}}V - H_{\text{NA}}(0) \right] \right\}. \quad (\text{A5})$$

Changing variables to real quantities  $\tilde{q}_i = V^{1/3}q_i$  and  $\tilde{p}_i = p_i/(sV^{1/3})$  gives

$$d\tilde{q}d\tilde{p}dVf(\tilde{p}, \tilde{q}, V) = \frac{1}{Z_{\text{NPA}}N!h^{N_f}} \int ds \int d\pi_s \int d\pi_V \times d\tilde{\mathbf{p}}d\tilde{\mathbf{q}}dV s^{N_f} \delta\left\{s \left[ H(\tilde{\mathbf{p}}, \tilde{\mathbf{q}}) + \frac{\pi_s^2}{2Q_s} + \frac{\pi_V^2}{2Q_V} + gkT \ln s + P_{\text{ext}}V - H_{\text{NA}}(0) \right] \right\}. \quad (\text{A6})$$

For a function  $f(s)$  with a single pole at  $s=s_0$  we have

$$\delta(f(s)) = \frac{\delta(s-s_0)}{|f'(s_0)|},$$

which, for our case, gives

$$d\tilde{q}d\tilde{p}dVf(\tilde{p}, \tilde{q}, V) = \frac{1}{Z_{\text{NPA}}N!h^{N_f}} \int ds \int d\pi_s \times \int d\pi_V d\tilde{\mathbf{p}}d\tilde{\mathbf{q}}dV \frac{s^{N_f}}{gkT} \delta\left\{s - \exp\left[-\frac{1}{gkT} \left( H(\tilde{p}, \tilde{q}) + \frac{\pi_s^2}{2Q_s} + \frac{\pi_V^2}{2Q_V} + P_{\text{ext}}V - H_{\text{NA}}(0) \right) \right] \right\}. \quad (\text{A7})$$

Integration over  $s$  gives

$$d\tilde{q}d\tilde{p}dVf(\tilde{p},\tilde{q},V)=\frac{1}{gkTZ_{\text{NPA}}N!h^{N_f}}\int d\pi_s\int d\pi_v$$

$$\times d\tilde{p}d\tilde{q}dV\exp\left[-\frac{N_f}{gkT}\left(H(\tilde{p},\tilde{q})\right.\right.$$

$$\left.+\frac{\pi_s^2}{2Q_s}+\frac{\pi_v^2}{2Q_p}+P_{\text{ext}}V\right.$$

$$\left.-H_{\text{NP}}(o)\right]. \quad (\text{A8})$$

Setting  $g=N_f$  and integrating over  $\pi_s$  and  $\pi_v$ , together with the observation that when the above procedure is applied to  $Z_{\text{NPA}}$ , we obtain a result that is a constant times the normal isothermal–isobaric partition function

$$Z=\frac{1}{N!h^{N_f}}\int dV\int d\tilde{p}\int d\tilde{q}$$

$$\times\exp\{-\beta[H(\tilde{p},\tilde{q})+P_{\text{ext}}V]\}. \quad (\text{A9})$$

This gives (after canceling the differentials and common constants)

$$f(\tilde{p},\tilde{q},V)=\frac{1}{ZN!h^{N_f}}\exp\{-\beta[H(\tilde{p},\tilde{q})+P_{\text{ext}}V]\}. \quad (\text{A10})$$

which is the usual isothermal–isobaric distribution.

## APPENDIX B: VIRIAL RELATIONS AND THE NPA EQUATIONS OF MOTION

The equations of motion for any consistent extended Hamiltonian designed to generate the isothermal–isobaric distribution, should generate trajectory averages that obey the appropriate virial relations,<sup>11</sup> namely,  $\langle\mathcal{P}-P_{\text{ext}}\rangle_{\text{NPT}}=0$  and  $\langle V(\mathcal{P}-P_{\text{ext}})\rangle_{\text{NPT}}+kT=0$ . In this Appendix we show that these relations are indeed consistent with the NPA equations of motion.

To show the first relation, we start with the NPA equation of motion for  $\pi_v$

$$\pi_v=s(\mathcal{P}-P_{\text{ext}}). \quad (\text{B1})$$

Rearranging gives

$$\mathcal{P}-P_{\text{ext}}=\frac{\dot{\pi}_v}{s}=\frac{d}{dt}\left[\frac{\pi_v}{s}\right]+\frac{\dot{s}\pi_v}{s^2}. \quad (\text{B2})$$

Taking the equilibrium average within the NPA distribution gives

$$\langle(\mathcal{P}-P_{\text{ext}})\rangle_{\text{NPA}}=\left\langle\frac{\dot{s}\pi_v}{s^2}\right\rangle_{\text{NPA}}. \quad (\text{B3})$$

since the equilibrium average of the time derivative of  $\pi_v/s$  will be zero. Using the NPA equation of motion for  $s$  gives

$$\langle(\mathcal{P}-P_{\text{ext}})\rangle=\left\langle\frac{\pi_s\pi_v}{sQ_s}\right\rangle. \quad (\text{B4})$$

Since the distributions for both  $\pi_s$  and  $\pi_v$  are symmetric about zero, independent of  $s$  and each other, the term on the right-hand-side is zero giving

$$\langle(\mathcal{P}-P_{\text{ext}})\rangle_{\text{NPA}}=0. \quad (\text{B5})$$

Since  $\mathcal{P}-P_{\text{ext}}$  depends only upon the real positions and momenta (and not explicitly on  $s$ ,  $\pi_s$ , or  $\pi_v$ ), the average over the NPA distribution will be equal to that over the standard NPT distribution, as shown in Appendix A, giving

$$\langle(\mathcal{P}-P_{\text{ext}})\rangle_{\text{NPT}}=0, \quad (\text{B6})$$

as required.

Now for the second virial relation, we start by multiplying the NPA equation of motion for  $\pi_v$  by  $V$

$$V\dot{\pi}_v=s(\mathcal{P}-P_{\text{ext}})V. \quad (\text{B7})$$

Rearranging gives

$$(\mathcal{P}-P_{\text{ext}})V=\frac{V\dot{\pi}_v}{s}=\frac{d}{dt}\left[\frac{V\pi_v}{s}\right]-\frac{\dot{V}\pi_v}{s}+\frac{\dot{s}V\pi_v}{s^2}. \quad (\text{B8})$$

Taking an equilibrium average gives

$$\langle(\mathcal{P}-P_{\text{ext}})V\rangle_{\text{NPA}}=-\left\langle\frac{\dot{V}\pi_v}{s}\right\rangle_{\text{NPA}}+\left\langle\frac{\dot{s}V\pi_v}{s^2}\right\rangle_{\text{NPA}}. \quad (\text{B9})$$

where we have again used the fact that the equilibrium average of a time derivative will be zero at equilibrium. Using the NPA equations of motion for  $s$  and  $V$  gives

$$\langle(\mathcal{P}-P_{\text{ext}})V\rangle_{\text{NPA}}=-\left\langle\frac{\pi_v^2}{Q_v}\right\rangle_{\text{NPA}}+\left\langle\frac{\pi_s\pi_vV}{s}\right\rangle_{\text{NPA}}. \quad (\text{B10})$$

Now the last term is zero since the distributions for  $\pi_s$  and  $\pi_v$  are symmetric, independent of  $s$  and  $V$ . The next term can be evaluated by the equipartition theorem, since it does not involve  $s$ , to give

$$\left\langle\frac{\pi_v^2}{Q_v}\right\rangle_{\text{NPA}}=kT. \quad (\text{B11})$$

This gives

$$\langle(\mathcal{P}-P_{\text{ext}})V\rangle_{\text{NPA}}+kT=0, \quad (\text{B12})$$

which by the above reasoning yields

$$\langle(\mathcal{P}-P_{\text{ext}})V\rangle_{\text{NPT}}+kT=0. \quad (\text{B13})$$

We can see here that there is no inconsistency between the expected virial relations generated from the equations of motion and the isothermal–isobaric distribution.

## APPENDIX C: MULTIPLE TIME STEPPING WITH NPA

Suppose the potential energy can be separated into a quickly varying part  $U_1$  (e.g., stiff harmonic vibrations or short-ranged repulsive interactions) and a slowly varying part  $U_2$  (e.g., long-range potentials, nonbonded interactions). The Nosé–Anderson Hamiltonian can be written

$$H_{NA} = \frac{1}{V^{2/3}s^2} \sum \frac{p_i^2}{2m_i} + U_1(V^{1/3}\mathbf{q}) + U_2(V^{1/3}\mathbf{q}) + \frac{\pi_V^2}{2Q_V} + \frac{\pi_s^2}{2Q_s} + P_{\text{ext}}V + gkT \ln s \quad (C1)$$

$$= H_{NA}^{(1)} + H_{NA}^{(2)} \quad (C2)$$

where we split the Hamiltonian as follows:

$$H_{NA}^{(1)} \equiv \frac{1}{V^{2/3}s^2} \sum \frac{p_i^2}{2m_i} + \frac{\pi_V^2}{2Q_V} + \frac{\pi_s^2}{2Q_s} U_1(V^{1/3}\mathbf{q}) + gkT \ln s, \quad (C3)$$

$$H_{NA}^{(2)} \equiv U_2(V^{1/3}\mathbf{q}) + P_{\text{ext}}V. \quad (C4)$$

Consequently the Nosé–Poincaré–Anderson Hamiltonian is

$$H_{NPA} = s[H_{NA} - H_{NA}(0)] \quad (C5)$$

$$= H_{NPA}^{(1)} + H_{NPA}^{(2)}, \quad (C6)$$

where

$$H_{NPA}^{(1)} \equiv s[H_{NA}^{(1)} - H_{NA}^{(1)}(0)], \quad (C7)$$

$$H_{NPA}^{(2)} \equiv s[H_{NA}^{(2)} - H_{NA}^{(2)}(0)].$$

We define the flow map for a Hamiltonian  $H$  as the operator such that for any dynamical variable  $\mathbf{A}$  we have

$$\mathbf{A}(t) = \phi_{H,t} \mathbf{A}(t=0). \quad (C8)$$

Using the standard Trotter factorization used to develop multiple-time step molecular dynamics<sup>18</sup> we write

$$\phi_{H_{NPA},h} = \phi_{H_{NPA}^{(2)},h/2} (\phi_{H_{NPA}^{(1)},h/l})^l \phi_{H_{NPA}^{(2)},h/2}, \quad (C9)$$

where  $l$  is positive integer and the time step for the slow and fast variables is  $h$  and  $h/l$ , respectively. Using the symplectic Euler algorithm [Eq. (8)] and its adjoint [Eq. (9)], one can easily derive a multiple-time-step generalized leapfrog algorithm for this splitting

$$p_{i,n+1/2,0} = p_{i,n} - \frac{h}{2} s_n V_n^{1/3} \nabla_i U_2(V_n^{1/3} \mathbf{q}_n), \quad (C10a)$$

$$\pi_{V,n+1/2,0} = \pi_{V,n} + \frac{h}{2} s_n [\mathcal{P}_2(V_n, \mathbf{q}_n) - P_{\text{ext}}], \quad (C10b)$$

$$\pi_{s,n+1/2,0} = \pi_{2,n} - \frac{h}{2} s_n \Delta H_2(V_n, \mathbf{q}_n), \quad (C10c)$$

do  $j=1, l$

$$p'_{i,n+1/2,j} = p_{i,n+1/2,j-1} - \frac{h}{2l} s_{nj-1} V_{nj-1}^{-1/3} \nabla_i U_2(V_{nj-1}^{1/3} \mathbf{q}_{n,j-1}), \quad (C10d)$$

$$\pi'_{V,n+1/2,j} = \pi_{V,n+1/2,j-1} + \frac{h}{2l} s_{n,j-1} \mathcal{P}_1(s_{n,j-1}, V_{nj-1}, \mathbf{q}_{n,j-1}), \quad (C10e)$$

$$\pi'_{s,n+1/2,j} = \pi_{s,n+1/2,j-1} + \frac{h}{2l} \left\{ \frac{1}{V_{n,j-1}^{2/3} s_{n,j-1}^2} \sum \frac{p_{i,n+1/2,j}^2}{2m_i} - gkT - \Delta H_1(s_{n,j-1}, V_{n,j-1}, \mathbf{q}_{n,j-1}, \mathbf{p}'_{n+1/2,j}, \pi'_{s,n+1/2,j}, \pi'_{V,n+1/2,j}) \right\}, \quad (C10f)$$

$$s_{n,j} = s_{n,j-1} + \frac{h}{2l} (s_{n,j} + s_{n,j-1}) \frac{\pi_{s,n+1/2,j}}{Q_s}, \quad (C10g)$$

$$V_{n,j} = V_{n,j-1} + \frac{h}{2l} (s_{n,j} + s_{n,j-1}) \frac{\pi_{V,n+1/2,j}}{Q_V}, \quad (C10h)$$

$$q_{i,n,j} = q_{i,n,j-1} + \frac{h}{2l} \left( \frac{1}{s_{n,j-1} V_{n,j-1}^{2/3}} + \frac{1}{s_{n,j} V_{n,j}^{2/3}} \right) \frac{p_{i,n+1/2,j}}{m_i}, \quad (C10i)$$

$$\pi_{s,n+1/2,j} = \pi_{s,n+1/2,j} + \frac{h}{2l} \left\{ \frac{1}{V_{n,j}^{2/3} s_{n,j}^2} \sum \frac{(p'_{i,n+1/2,j})^2}{2m_i} - gkT - \Delta H_1(s_{n,j}, V_{n,j}, \mathbf{q}_{n,j}, \mathbf{p}'_{n+1/2,j}, \pi'_{s,n+1/2,j}, \pi'_{V,n+1/2,j}) \right\}, \quad (C10j)$$

$$\pi_{V,n+1/2,j} = \pi'_{V,n+1/2,j} + \frac{h}{2l} s_{n,j} \mathcal{P}_1(s_{n,j}, V_{n,j}, \mathbf{q}_{n,j}), \quad (C10k)$$

$$p_{i,n+1/2,j} = p'_{i,n+1/2,j} - \frac{h}{2l} s_{n,j} V_{n,j}^{1/3} \nabla_i U_2(V_{n,j}^{1/3} \mathbf{q}_{n,j}), \quad (C10l)$$

end do

$$q_{i,n+1} \equiv q_{i,n,l}$$

$$s_{n+1} \equiv s_{n,l}$$

$$V_{n+1} \equiv V_{n,l}$$

$$\pi_{s,n+1} \equiv \pi_{s,n+1/2,l} - \frac{h}{2} s_{n+1} \Delta H_2(V_{n+1}, \mathbf{q}_{n+1}), \quad (C10m)$$

$$\pi_{V,n+1} = \pi_{V,n+1/2,l} + \frac{h}{2} s_{n+1} [\mathcal{P}_2(V_{n+1}, \mathbf{q}_{n+1}) - P_{\text{ext}}], \quad (C10n)$$

$$p_{i,n+1} = p_{i,n+1/2,l} - \frac{h}{2} s_{n+1} V_{n+1}^{1/3} \nabla_i U_2(V_{n+1}^{1/3} \mathbf{q}_{n+1}), \quad (C10o)$$

where the partial instantaneous pressures  $\mathcal{P}_1$  and  $\mathcal{P}_2$  are given by

$$\mathcal{P}_1 = \frac{2}{3V} \sum_i \frac{p_i^2}{2m_i V^{2/3} s^2} - \frac{1}{3V} \sum_i \frac{\partial U_1}{\partial q_i} q_i, \quad (C11)$$

$$\mathcal{P}_2 = -\frac{1}{3V} \sum_i \frac{\partial U_2}{\partial q_i} q_i. \quad (C12)$$

This algorithm, being based on a symmetric concatenation of symplectic Euler steps, is both time reversible and



symplectic. It should be noted that other splittings are certainly possible here, but a detailed analysis and determination of the optimal splitting would be the subject of further work.

- <sup>1</sup>M. A. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Oxford Science Press, Oxford, 1987).
- <sup>2</sup>D. Frenkel and B. Smit, *Understanding Molecular Simulation* (Academic, New York, 1996).
- <sup>3</sup>D. Brown and J. H. R. Clarke, *Mol. Phys.* **51**, 1243 (1984).
- <sup>4</sup>D. J. Evans and G. P. Morriss, *Comput. Phys. Rep.* **1**, 297 (1984).
- <sup>5</sup>H. C. Anderson, *J. Chem. Phys.* **72**, 2384 (1980).
- <sup>6</sup>S. Nose, *Mol. Phys.* **52**, 255 (1984).
- <sup>7</sup>S. Nose, *J. Chem. Phys.* **81**, 511 (1984).
- <sup>8</sup>W. G. Hoover, *Phys. Rev. A* **34**, 2499 (1986).
- <sup>9</sup>J. M. Sanz-Serna and M. P. Calvo, *Numerical Hamiltonian Problems* (Chapman and Hall, New York, 1995).

- <sup>10</sup>W. G. Hoover, *Phys. Rev. A* **31**, 1695 (1985).
- <sup>11</sup>D. J. Tobias, G. J. Martyna, and M. L. Klein, *J. Chem. Phys.* **101**, 4177 (1994).
- <sup>12</sup>D. J. Tobias, G. J. Martyna, M. E. Tuckerman, and M. L. Klein, *Mol. Phys.* **87**, 1117 (1996).
- <sup>13</sup>S. D. Bond, B. J. Leimkuhler, and B. B. Laird, *J. Comput. Phys.* **151**, 114 (1999).
- <sup>14</sup>J. M. Sanz-Serna, *BIT* **28**, 877 (1991).
- <sup>15</sup>F. Lasagni, *Z. Angew. Math. Phys.* **39**, 952 (1988).
- <sup>16</sup>Y. B. Suris, *USSR Comput. Math. Math. Phys.* **29**, 138 (1989).
- <sup>17</sup>J. M. Sanz-Serna, *Acta Numer.* **1**, 243 (1991).
- <sup>18</sup>M. Tuckerman and B. J. Berne, *J. Chem. Phys.* **97**, 1990 (1992).
- <sup>19</sup>J. Mei and J. W. Davenport, *Phys. Rev. B* **46**, 21 (1992).
- <sup>20</sup>S. Nose, *Prog. Theor. Phys. Suppl.* **103**, 1 (1991).
- <sup>21</sup>T. Cagin and J. R. Ray, *Phys. Rev. A* **37**, 4510 (1988).
- <sup>22</sup>W. C. Swope, H. C. Anderson, P. H. Berens, and K. R. Wilson, *J. Chem. Phys.* **76**, 637 (1982).