Molecular dynamics integration and molecular vibrational theory. II. Simulation of nonlinear molecules

Matej Praprotnik and Dušanka Janežič[a)]
National Institute of Chemistry, Hajdrihova 19, 1000 Ljubljana, Slovenia

(Received 17 November 2004; accepted 11 February 2005; published online 29 April 2005)

A series of molecular dynamics (MD) simulations of nonlinear molecules has been performed to test the efficiency of newly introduced semianalytical second-order symplectic time-reversible MD integrators that combine MD and the standard theory of molecular vibrations. The simulation results indicate that for the same level of accuracy, the new algorithms allow significantly longer integration time steps than the standard second-order symplectic leap-frog Verlet method. Since the computation cost per integration step using new MD integrators with longer time steps is approximately the same as for the standard method, a significant speed-up in MD simulation is achieved. © *2005 American Institute of Physics*. [DOI: 10.1063/1.1884608]

I. INTRODUCTION

In the preceding paper,[1] new semianalytical second-order symplectic integrators are presented, developed by combining the molecular dynamics (MD) integration[2] and the standard theory of molecular vibrations.[3–6] The unique feature of the new integrators is in that the standard theory of molecular vibrations, which is a very efficient tool to analyze the dynamics of the studied system from computed trajectories,[7–12] is used not to analyze, but to compute trajectories of molecular systems. Information about the energy distribution of normal modes and the energy transfer between them is thus obtained without additional computations.

The key property of a good MD integrator is the conservation of the system's total energy over a long time interval. Backward error analysis[13] has indicated that symplectic numerical integration methods approximately conserve the total energy of a system over time periods that are exponentially long in the size of the integration time step. Long-time conservation of the total energy by new integrators using long integration time steps is achieved by the analytical treatment of high-frequency molecular vibrations within the framework of the symplectic decomposition schemes.[13–16]

In this paper the new integration methods are employed to perform MD simulations of systems of nonlinear molecules with one equilibrium configuration and no internal rotation. The new integrators are superior to the standard leap-frog Verlet (LFV) method[17] because they allow longer integration time steps to be used for the same computational accuracy with nearly the same computational cost per integration step.

II. METHODS

The Hamilton equations, which are solved for each atom of the system in MD integration, can be written in terms of Lie operators as[13]

$$\frac{d\boldsymbol{\eta}}{dt} = \{\boldsymbol{\eta}, H\} = \hat{L}_H \boldsymbol{\eta}, \tag{1}$$

where $\hat{L}_H$ is the Lie operator {,} is the Poisson bracket,[18] and $\boldsymbol{\eta}=(\mathbf{q},\mathbf{p})$ is a vector of the coordinates and their conjugate momenta of all the particles.

The formal solution of the Hamiltonian system (1) is

$$\boldsymbol{\eta}|_{t_k+\Delta t} = \exp(\Delta t \hat{L}_H)\ \boldsymbol{\eta}|_{t_k} \tag{2}$$

and it represents the exact time evolution of a trajectory in phase space composed of coordinates and momenta of all the particles from t_k to $t_k+\Delta t$, where Δt is the integration time step.[18]

A. Split integration symplectic method (SISM)

In developing new MD integration method[1] we first decompose the Hamiltonian H of a system into two parts

$$H = H_0 + H_r, \tag{3}$$

where H_0 is the pure harmonic part of the Hamiltonian and H_r is the remaining part.[19]

Next, a second-order approximation for Eq. (2), known as the generalized leap-frog scheme,[14,15] is used

$$\boldsymbol{\eta}|_{t_{k+1}} = \exp\left(\frac{\Delta t}{2}\hat{L}_{H_0}\right)\exp(\Delta t \hat{L}_{H_r})\exp\left(\frac{\Delta t}{2}\hat{L}_{H_0}\right)\ \boldsymbol{\eta}|_{t_k} + O(\Delta t^3), \tag{4}$$

which defines the split integration symplectic method (SISM).[1,20–22] The propagation by $\exp((\Delta t/2)\hat{L}_{H_0})$ is solved analytically using the normal modes of an isolated molecule,[5] while the propagation by $\exp(\Delta t \hat{L}_{H_r})$ is solved numerically in the same way as in the standard LFV method.[17] The SISM differs from other decomposition MD integration methods in that it uses the standard theory of molecular vibrations, in particular, the concept of the Eckart frame, to define the translating and rotating internal coordinate system of a molecule for the time propagation. The

[a)]Author to whom correspondence should be addressed. Electronic mail: dusa@cmm.ki.si

method is described in full detail in the preceding paper.[1]

B. Multiple time stepping SISM (SISM-MTS)

First, we split the Hamiltonian H of the system as[1]

$$H = H_1 + H_2, \tag{5}$$

$$H_1 = V_{nb}, \tag{6}$$

$$H_2 = H_0 + V_{ah}, \tag{7}$$

where H_0 is the pure harmonic part of the Hamiltonian, V_{nb} is the sum of the Coulomb and Lennard-Jones potential, and V_{ah} is the anharmonic vibrational potential of higher terms (cubic, quartic, etc.) defined in terms of the displacements of atoms from their equilibrium positions.[1]

The propagator $\exp(\Delta t\hat{L}_H)$ is then approximated as

$$\exp(\Delta t\hat{L}_H) = \exp\left(\frac{\Delta t}{2}\hat{L}_{V_{nb}}\right) \times\left[\exp\left(\frac{\delta t}{2}\hat{L}_{H_0}\right)\exp(\delta t\hat{L}_{V_{ah}})\exp\left(\frac{\delta t}{2}\hat{L}_{H_0}\right)\right]^n \times\exp\left(\frac{\Delta t}{2}\hat{L}_{V_{nb}}\right) + O(\Delta t^3), \tag{8}$$

which is used to derive the multiple time stepping SISM (SISM-MTS). Here Δt is the integration time step and $\delta t = \Delta t/n$ is the smaller integration time step that corresponds to the time scale of high-frequency interactions defined by V_{ah}. The propagation by $\exp((\delta t/2)\hat{L}_{H_0})$ is performed analytically in the same way as in the SISM.[1]

C. Equilibrium SISM (SISM-EQ)

In the equilibrium SISM (SISM-EQ), the numerical scheme of the SISM given by Eq. (4) is used to propagate the coordinates and momenta of the atoms; the potential of the slow nonbonded forces is computed with the equilibrium positions of atoms[1,23–25]

$$V_{nb}(\mathbf{q}) \rightarrow V_{nb}[\mathbf{d}(\mathbf{q})], \tag{9}$$

$$\mathbf{F}_{nb}(\mathbf{q}) \rightarrow \boldsymbol{\mathcal{J}}^T \cdot \mathbf{F}_{nb}[\mathbf{d}(\mathbf{q})], \tag{10}$$

where V_{nb} is the sum of the Coulomb and van der Waals potentials, $\mathbf{F}_{nb} = -\partial V_{nb}/\partial\mathbf{q}$ is the corresponding force, $\partial/\partial\mathbf{q} = (\partial/\partial X_1, \partial/\partial Y_1, \partial/\partial Z_1, \ldots, \partial/\partial X_n, \partial/\partial Y_n, \partial/\partial Z_n)$, $\mathbf{q} = (q_1, \ldots, q_{3n}) = (X_1, Y_1, Z_1, \ldots, X_n, Y_n, Z_n)$ are the Cartesian coordinates of all atoms in the system and n is the number of atoms in the system, and $\mathbf{d}(\mathbf{q}) \in \mathbb{R}^{3n}$ are the equilibrium positions of atoms in all molecules of the system, given by the standard theory of molecular vibrations.[1,3,5,6]

When the scheme of the SISM-MTS, defined by Eq. (8), is used to propagate the coordinates and momenta of the atoms, it gives rise to the equilibrium SISM-MTS (SISM-MTS-EQ) method.[1] This method conserves the following quantity:

$$H = T(\mathbf{p}) + V_{vib}(\mathbf{q}) + V_{nb}[\mathbf{d}(\mathbf{q})]. \tag{11}$$

D. Leap-frog Verlet (LFV and LFV-EQ)

To demonstrate the effectiveness of the new methods, we compared the computational performances for the same level of accuracy with the standard second-order symplectic LFV algorithm[17] in which the Hamiltonian is split into the kinetic and potential energy,

$$H = T + V, \tag{12}$$

using second-order generalized leap-frog scheme[15,26]

$$\boldsymbol{\eta}|_{t_{k+1}} = \exp\left(\frac{\Delta t}{2}\hat{L}_T\right)\exp(\Delta t\hat{L}_V)\exp\left(\frac{\Delta t}{2}\hat{L}_T\right)\boldsymbol{\eta}|_{t_k} + O(\Delta t^3). \tag{13}$$

Equation (13) is explicitly written as

$$\mathbf{q}_k' = \mathbf{q}_k + \mathbf{M}^{-1}\cdot\mathbf{p}_k\frac{\Delta t}{2},$$

$$\mathbf{p}_{k+1} = \mathbf{p}_k - \Delta t\frac{\partial V}{\partial\mathbf{q}}(\mathbf{q}_k'),$$

$$\mathbf{q}_{k+1} = \mathbf{q}_k' + \mathbf{M}^{-1}\cdot\mathbf{p}_{k+1}\frac{\Delta t}{2}, \tag{14}$$

where $\mathbf{M} \in R^{3n\times 3n}$ is a diagonal mass matrix. The diagonal elements are $M_{11}=m_1$, $M_{22}=m_1$, $M_{33}=m_1, \ldots, M_{3n-2,3n-2}=m_n$, $M_{3n-1,3n-1}=m_n$, $M_{3n,3n}=m_n$, where m_i is the mass of the ith atom.

When the numerical scheme of the LFV defined by Eq. (13) is used to propagate the coordinates and momenta of the atoms, and the potential of the long-range electrostatic and van der Waals potential is calculated with the equilibrium positions of the atoms in each molecule in the same way as for the SISM-EQ and SISM-MTS-EQ,[1] then this gives rise to the equilibrium LFV (LFV-EQ) method.

III. COMPUTATIONAL DETAILS

The applicability of the SISM for MD integration is, at present, limited to systems of molecules with one equilibrium configuration and no internal rotation, and in which the displacements of atoms from their equilibrium positions are sufficiently small that we can use the dynamical molecular model[6] to describe molecular vibrations. A four-atom molecule, the hydrogen peroxide (H_2O_2), schematically shown in Fig. 1, has been chosen as an example of a nonlinear and nonplanar molecule.

A. Model potential development

For this class of molecules we first develop an appropriate model potential to be used in MD simulations of liquid H_2O_2 by the SISM. The H_2O_2 molecule, which has no center of symmetry, is one of the simplest molecules with a hindered internal rotation of the hydrogen atoms around the bond between the oxygen atoms. It has two equivalent stable equilibrium configurations (at $\pm\phi_0$) and two transition states, cis ($\phi_0 = 0°$) and trans ($\phi_0 = 180°$).[27,28] The experimental

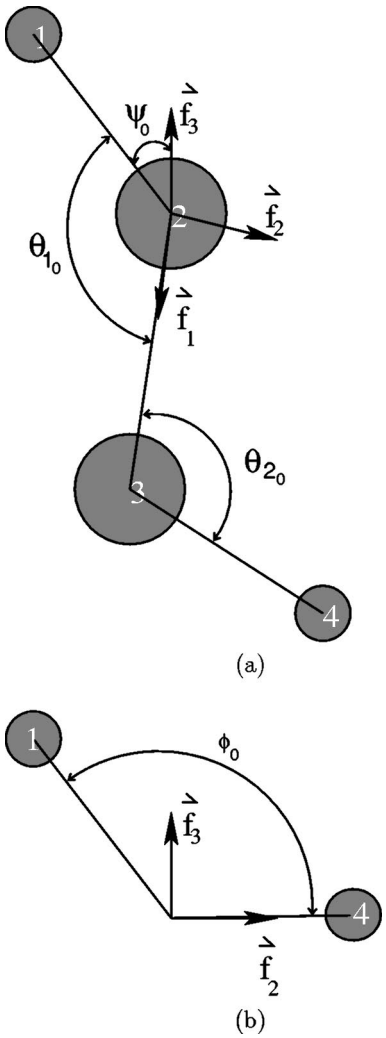


FIG. 1. Description of the positions of atoms in the equilibrium configuration of hydrogen peroxide. (a) Definition of angles. (b) View along the bond between atoms 2 and 3.

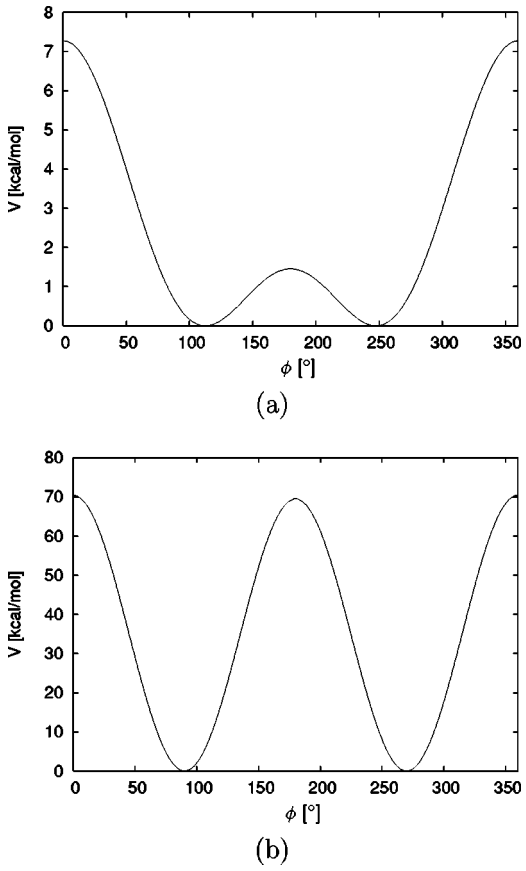


FIG. 2. (a) The torsional potential defined by Eq. (15) with $\phi_0=112.46°$, $V_0=2V_b/(1-\cos\phi_0)^2$, and $V_b=7.28$ kcal/mol corresponding to isolated molecule. (b) The torsional potential defined by Eq. (15) with $\phi_0=90.2°$ and $V_0=140$ kcal/mol defining a molecule with one equilibrium configuration and no internal rotation.

value of ϕ_0 is $119.8° \pm 3°$ for the gas phase.[29] The value for ϕ_0 determined from *ab initio* calculations of an isolated H_2O_2 molecule is 112.46°.[27] The torsional potential, also determined from *ab initio* calculations,[27] can be fitted with the function of the harmonic cosine form

$$V(\phi) = \tfrac{1}{2}V_0(\cos\phi - \cos\phi_0)^2, \tag{15}$$

with $V_0=2V_b/(1-\cos\phi_0)^2$, $V_b=7.28$ kcal/mol, $\phi_0=112.46^{\circ}$, and minima

$$\phi_{\min} = 2n\pi \pm \phi_0. \tag{16}$$

From the torsional potential, shown in Fig. 2(a), it can be observed that there are two potential barriers surrounding the minimum at 112.46°, which corresponds to the equilibrium configuration. The high 7.28 kcal/mol potential barrier corresponds to the cis transition state whereas the low 1.08 kcal/mol potential barrier corresponds to the trans transition state.

The equilibrium configuration of the H_2O_2 molecule in the gas state, however, does not correspond to the corresponding structure in the liquid state. The value of ϕ_0 in the liquid state, which is the most interesting physical system for testing the efficiency of different numerical integrators for MD simulation, is therefore different from the corresponding value in the gas state.[30] Because, at present state of development, the SISM is efficiently applicable only to systems of molecules with one equilibrium configuration and with no internal rotation, we have taken the experimentally determined structure in the solid state[30] with $\phi_0=90.2°$ for the equilibrium configuration of the H_2O_2 molecule instead of the corresponding structure in the liquid state and we have set the height of potential barriers, which surround the minimum at $\phi_0=90.2°$, artificially high at $V_0=140$ kcal/mol to ensure that the displacements of the hydrogens atoms are sufficiently small so that they can be considered as torsional vibrations. The corresponding torsional potential (15) is depicted in Fig. 2(b) from which it can be observed that the heights of the potential barrier for the trans and cis transition states are equal. An H_2O_2 molecule with the equilibrium configuration determined by these parameters can therefore be considered as a molecule with one equilibrium configuration and no internal rotation. Hence, we have used it as an example of the class of nonlinear and nonplanar molecules in MD simulation by the SISM.

The model Hamiltonian, which we have developed for MD simulation of the liquid H_2O_2, is

TABLE I. Parameters of the Hamiltonian defined by Eq. (17) for the H_2O_2 molecule. The quantity e_0 is the elementary charge.

Parameter	Value
$b_{OH_0}=b_{1_0}=b_{3_0}$	0.988 Å
$b_{OO_0}=b_{2_0}$	1.453 Å
θ_{1_0}	102.7°
θ_{2_0}	102.7°
ϕ_0	90.2°
$\psi_0=\phi_0-90^\circ$	0.2°
$k_{b_{OH}}=k_{b_1}=k_{b_3}$	900.0 kcal/mol/Å^2
$k_{b_{OO}}=k_{b_2}$	580.0 kcal/mol/Å^2
k_{θ_1}	140.0 kcal/mol/radian^2
k_{θ_2}	140.0 kcal/mol/radian^2
V_0	140.0 kcal/mol
e_H	$0.350\,53e_0$
e_O	$-0.350\,53e_0$
σ_{HH}	0.40 Å
σ_{OO}	3.1507 Å
ε_{HH}	0.045 98 kcal/mol
ε_{OO}	0.152 073 kcal/mol

$$H=\sum_i \frac{\mathbf{p}_i^2}{2m_i}+\frac{1}{2}\sum_{bonds} k_b(b-b_0)^2+\frac{1}{2}\sum_{angles} k_\theta(\theta-\theta_0)^2 + \frac{1}{2}\sum_{torsions} V_0(\cos\phi-\cos\phi_0)^2+\sum_{i>j}\frac{e_ie_j}{4\pi\epsilon_0 r_{ij}} + \sum_{i>j} 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12}-\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right], \tag{17}$$

where i and j run over all atoms, m_i is the mass of the ith atom, $\mathbf{p}_i$ is the linear momentum of the ith atom, b_0 and θ_0 are reference values for the bond lengths and angles, respectively, k_b and k_θ are the corresponding force constants, ϕ_0 is the reference value for the torsional angle, V_0 is the corresponding barrier height, e_i denotes the charge on the ith atom, ϵ_0 is the dielectric constant in vacuum, r_{ij} is the distance between the ith and jth atoms, and ε_{ij} and σ_{ij} are the corresponding constants of the Lennard-Jones potential, and the Lorentz–Berthelot mixing rules are used.[2] The van der Waals and Coulomb interactions between hydrogen atoms of the same molecule are explicitly taken into account. For the reference values of the bonds, angles, and torsional angles we have taken the experimental values from Ref. 30. For the constants of the Lennard-Jones potential we have taken the corresponding values of flexible TIP3P water model[31,32] and the partial charges of the atoms were calculated from the *ab initio* calculated dipole moment[33] and the corresponding structure of an isolated H_2O_2 molecule.[27] The force constants for bond stretching and angle bending were determined by fitting the normal mode frequencies calculated by normal mode analysis[1] to the experimental frequencies in the IR and Raman spectrum of liquid H_2O_2.[34–36] The parameters of the Hamiltonian (17) are reported in Table I. The experimental and calculated normal mode frequencies for the H_2O_2 molecule using these parameters are reported in Table II.

TABLE II. Experimental vibration frequencies (Refs. 34–36) of liquid H_2O_2 and normal mode frequencies of the H_2O_2 molecule determined by normal mode analysis using the parameters from Table I.

Normal mode	$1/\lambda$ (cm^{-1}) (experiment)[a]	$1/\lambda$ (cm^{-1}) (theory)[b]
Antisymmetric O–H stretch	3360	3358
Symmetric O–H stretch	3360	3357
Symmetric angle bending	1421	1410
Antisymmetric angle bending	1350	1386
O–O stretch	878	880
Torsional oscillation	635	1965

[a]Experimental vibration frequencies.
[b]Normal mode frequencies.

B. Vibrational potential energy and internal coordinate system

The vibrational potential energy is the sum of the vibrational potential energies of all of the molecules in the system,

$$V_{vib}=\sum_{k=1}^{m} V_{vib_k}=\frac{1}{2}\sum_{bonds} k_b(b-b_0)^2+\frac{1}{2}\sum_{angles} k_\theta(\theta-\theta_0)^2 + \frac{1}{2}\sum_{torsions} V_0(\cos\phi-\cos\phi_0)^2, \tag{18}$$

where V_{vib_k} is the vibrational potential energy of the kth molecule in the system and m is the number of molecules in the system.

In the SISM, the propagation by $\exp\mathbf{(}(\Delta t/2)\hat{L}_{H_0}\mathbf{)}$ is integrated analytically using the normal coordinates to describe the vibrational, rotational, and translational degrees of freedom of each molecule in the system. For the transformation of Cartesian coordinates and momenta into the normal coordinates, the relative Cartesian displacement coordinates are required. To determine the vibrational frequencies and normal modes of vibration of the kth molecule in the system, the mass-weighted Hessian $\mathbf{M}^{-1/2}\cdot\mathbf{H}_k\cdot\mathbf{M}^{-1/2}\in\mathbb{R}^{3N\times 3N}$ has to be diagonalized. The matrix $\mathbf{H}_k$ is a symmetric matrix with the elements

$$H_{k_{ij}}=H_{k_{ji}}=\left(\frac{\partial^2 V_{harm_k}}{\partial\Delta q_i\partial\Delta q_j}\right)_0 \tag{19}$$

and $\mathbf{M}$ is a diagonal mass matrix with the elements $M_{11}=m_1$, $M_{22}=m_1$, $M_{33}=m_1,\ldots,M_{3N-2,3N-2}=m_N$, $M_{3N-1,3N-1}=m_N$, $M_{3N,3N}=m_N$, where N is the number of all of the atoms in the kth molecule. The harmonic vibrational potential energy V_{harm_k} for the kth molecule is defined as

$$\begin{aligned} V_{harm_k}&=\frac{1}{2}\sum_{i,j=1}^{3N}\left(\frac{\partial^2 V_{vib_k}}{\partial\Delta q_i\partial\Delta q_j}\right)_0\Delta q_i\Delta q_j \\ &=\frac{1}{2}\sum_{i,j=1}^{3N}\left(\frac{\partial^2 V_{harm_k}}{\partial\Delta q_i\partial\Delta q_j}\right)_0\Delta q_i\Delta q_j \\ &=\frac{1}{2}\sum_{i,j=1}^{3N} H_{k_{ij}}\Delta q_i\Delta q_j=\frac{1}{2}\Delta\mathbf{q}\cdot\mathbf{H}_k\cdot\Delta\mathbf{q}, \end{aligned} \tag{20}$$

where $\Delta\mathbf{q}=(\Delta x_1,\Delta y_1,\Delta z_1,\ldots,\Delta x_N,\Delta y_N,\Delta z_N)$ is a vector of

the relative Cartesian displacement coordinates.[1] The Hessian $\mathbf{H}_k$ as well as the functional form of V_{harm_k} are equal for every molecule in the system and therefore the index k is omitted.

The harmonic potential V_{harm} is then expressed as

$$V_{harm} = V_{stretch} + V_{bend} + V_{torsion}, \tag{21}$$

where $V_{stretch}$, V_{bend}, and $V_{torsion}$ are the bond stretching, angle bending, and torsional potentials, respectively, expressed as a quadratic forms in terms of relative Cartesian displacement coordinates.

The equilibrium configuration of the H_2O_2 molecule is shown in Fig. 1 where atom 1 is the first hydrogen atom, atoms 2 and 3 are oxygen atoms, and atom 4 is the second hydrogen atom, respectively. The definitions of the angles θ_{1_0}, θ_{2_0}, and ψ_0 are evident from Fig. 1(a), and the distances are $b_{1_0}=b_{3_0}=b_{\mathrm{OH}_0}$ and $b_{2_0}=b_{\mathrm{OO}_0}$. The unit vector pointing from atom 2 to atom 1 is then $(\cos\theta_{1_0}, -\sin\psi_0 \sin\theta_{1_0}, \cos\psi_0 \sin\theta_{1_0})$, the unit vector pointing from atom 2 to atom 3 is (1, 0, 0), and the unit vector pointing from atom 3 to atom 4 is $(-\cos\theta_{2_0}, \sin\theta_{2_0}, 0)$. The expression for $V_{stretch}$ in terms of the relative Cartesian coordinates is obtained by projecting the difference of the displacements of the atoms onto the unit vectors, which point along the bonds between atoms.[37] Then

$$\begin{aligned}
V_{stretch} &= \tfrac{1}{2}k_{b_1}[(\Delta x_1 - \Delta x_2, \Delta y_1 - \Delta y_2, \Delta z_1 - \Delta z_2)\cdot(\cos\theta_{1_0}, -\sin\psi_0\sin\theta_{1_0}, \cos\psi_0\sin\theta_{1_0})^T]^2 + \tfrac{1}{2}k_{b_2}[(\Delta x_3 - \Delta x_2, \Delta y_3 \\
&\quad - \Delta y_2, \Delta z_3 - \Delta z_2)\cdot(1,0,0)^T]^2 + \tfrac{1}{2}k_{b_3}[(\Delta x_4 - \Delta x_3, \Delta y_4 - \Delta y_3, \Delta z_4 - \Delta z_3)\cdot(-\cos\theta_{2_0}, \sin\theta_{2_0}, 0)^T]^2 \\
&= \tfrac{1}{2}k_{b_1}[(\Delta x_1 - \Delta x_2)\cos\theta_{1_0} - (\Delta y_1 - \Delta y_2)\sin\psi_0\sin\theta_{1_0} + (\Delta z_1 - \Delta z_2)\cos\psi_0\sin\theta_{1_0}]^2 + \tfrac{1}{2}k_{b_2}(\Delta x_3 - \Delta x_2)^2 \\
&\quad + \tfrac{1}{2}k_{b_3}[-(\Delta x_4 - \Delta x_3)\cos\theta_{2_0} + (\Delta y_4 - \Delta y_3)\sin\theta_{2_0}]^2,
\end{aligned} \tag{22}$$

where $\cdot$ denotes the dot product of two vectors, $k_{b_1}=k_{b_3}=k_{b_{\mathrm{OH}}}$, and $k_{b_2}=k_{b_{\mathrm{OO}}}$.

Similarly, the expression for V_{bend} in terms of the relative Cartesian coordinates are obtained by taking the components of the difference of the displacements of the atoms perpendicular to the bonds between the atoms.[37] Therefore

$$\begin{aligned}
V_{bend} &= \frac{1}{2}k_{\theta_1}\left[\frac{1}{b_{1_0}}(\Delta x_1 - \Delta x_2, \Delta y_1 - \Delta y_2, \Delta z_1 - \Delta z_2)\cdot(-\sin\theta_{1_0}, -\sin\psi_0\cos\theta_{1_0}, \cos\psi_0\cos\theta_{1_0})^T + \frac{1}{b_{2_0}}(\Delta x_3 - \Delta x_2, \Delta y_3\right. \\
&\quad \left. - \Delta y_2, \Delta z_3 - \Delta z_2)\cdot(0, \sin\psi_0, -\cos\psi_0)^T\right]^2 + \frac{1}{2}k_{\theta_2}\left[\frac{1}{b_{2_0}}(\Delta x_2 - \Delta x_3, \Delta y_2 - \Delta y_3, \Delta z_2 - \Delta z_3)\cdot(0,-1,0)^T\right. \\
&\quad \left. + \frac{1}{b_{3_0}}(\Delta x_4 - \Delta x_3, \Delta y_4 - \Delta y_3, \Delta z_4 - \Delta z_3)\cdot(\sin\theta_{2_0}, \cos\theta_{2_0}, 0)^T\right]^2 \\
&= \frac{1}{2}k_{\theta_1}\left[\frac{1}{b_{1_0}}[-(\Delta x_1 - \Delta x_2)\sin\theta_{1_0} - (\Delta y_1 - \Delta y_2)\sin\psi_0\cos\theta_{1_0} + (\Delta z_1 - \Delta z_2)\cos\psi_0\cos\theta_{1_0}]\right. \\
&\quad \left. + \frac{1}{b_{2_0}}[(\Delta y_3 - \Delta y_2)\sin\psi_0 - (\Delta z_3 - \Delta z_2)\cos\psi_0]\right]^2 \\
&\quad + \frac{1}{2}k_{\theta_2}\left[-\frac{1}{b_{2_0}}(\Delta y_2 - \Delta y_3) + \frac{1}{b_{3_0}}[(\Delta x_4 - \Delta x_3)\sin\theta_{2_0} + (\Delta y_4 - \Delta y_3)\cos\theta_{2_0}]\right]^2,
\end{aligned} \tag{23}$$

where $k_{\theta_1}=k_{\theta_2}=k_{\theta_{\mathrm{OOH}}}$.

To express the torsional angle ϕ in terms of the relative Cartesian coordinates we define

$$\begin{aligned}
\mathbf{a}_{12} = -(&-b_{1_0}\cos\theta_{1_0} + \Delta x_2 - \Delta x_1, b_{1_0}\sin\psi_0\sin\theta_{1_0} \\
&+ \Delta y_2 - \Delta y_1, -b_{1_0}\cos\psi_0\sin\theta_{1_0} + \Delta z_2 - \Delta z_1),
\end{aligned} \tag{24}$$

$$\mathbf{a}_{23} = -(b_{2_0} + \Delta x_3 - \Delta x_2, \Delta y_3 - \Delta x_2, \Delta z_3 - \Delta z_2), \tag{25}$$

$$\begin{aligned}
\mathbf{a}_{34} = -(&-b_{3_0}\cos\theta_{2_0} + \Delta x_4 - \Delta x_3, b_{3_0}\sin\theta_{2_0} + \Delta y_4 \\
&- \Delta y_3, \Delta z_4 - \Delta z_3),
\end{aligned} \tag{26}$$

where $\mathbf{a}_{\alpha\beta}=\mathbf{a}_\alpha+\boldsymbol{\rho}_\alpha-(\mathbf{a}_\beta+\boldsymbol{\rho}_\beta)$, $\boldsymbol{\rho}_\alpha=(\Delta x_\alpha, \Delta y_\alpha, \Delta z_\alpha)$, $\alpha,\beta = 1,\ldots,4$, and

$$\mathbf{a}_1 = (b_{1_0}\cos\theta_{1_0}, -b_{1_0}\sin\psi_0\sin\theta_{1_0}, b_{1_0}\cos\psi_0\sin\theta_{1_0}), \tag{27}$$

$$\mathbf{a}_2 = (0,0,0), \tag{28}$$

$$\mathbf{a}_3 = (b_{2_0}, 0, 0), \tag{29}$$

$$\mathbf{a}_4 = (b_{2_0} - b_{3_0} \cos\theta_{2_0}, b_{3_0} \sin\theta_{2_0}, 0). \tag{30}$$

$\cos\phi$ is then obtained as[2]

$$\cos\phi = \frac{(\mathbf{a}_{12} \times \mathbf{a}_{23}) \cdot (\mathbf{a}_{23} \times \mathbf{a}_{34})}{|\mathbf{a}_{12} \times \mathbf{a}_{23}||\mathbf{a}_{23} \times \mathbf{a}_{34}|}, \tag{31}$$

where $|\cdot|$ denotes the vector norm. In Eq. (31) we only keep the linear and quadratic terms in relative Cartesian displacement coordinates.

Then

$$\begin{aligned} V_{torsion} = \frac{1}{2} V_0 \Bigg[& - \frac{1}{b_{1_0} \sin\theta_{1_0}} (\Delta y_2 - \Delta y_1) \\ & + \frac{1}{b_{2_0}} [-2 \sin\psi_0 (\Delta x_3 - \Delta x_2) - (\cot\theta_{1_0} \\ & + \sin\psi_0 \cot\theta_{2_0}) (\Delta y_3 - \Delta y_2) + \cos\psi_0 \\ & \times \cot\theta_{2_0} (\Delta z_3 - \Delta z_2)] \\ & + \frac{1}{b_{3_0} \sin\theta_{2_0}} [\cos\psi_0 (\Delta z_4 - \Delta z_3) \\ & - \sin\psi_0 (\Delta y_4 - \Delta y_3)] \Bigg]^2. \end{aligned} \tag{32}$$

The elements of the Hessian $\mathbf{H}$ are determined by Eq. (19). The dimension of the mass-weighted Hessian $\mathbf{M}^{-1/2} \cdot \mathbf{H} \cdot \mathbf{M}^{-1/2}$ is 12×12 in the case of the H_2O_2 molecule. The $3N-6=6$ vibrational normal mode frequencies for the H_2O_2 molecule are given in Table II. The mass-weighted Hessian $\mathbf{M}^{-1/2} \cdot \mathbf{H} \cdot \mathbf{M}^{-1/2}$ was diagonalized using subroutines TRED2 and TQLI taken from Ref. 38. These subroutines were also used to diagonalize the symmetric positive definite Gram matrix $\boldsymbol{\mathcal{F}}$.[1]

The translating and rotating internal coordinate system is defined by the right-handed triad of unit vectors $\mathbf{f}_i$, $i=1,2,3$, where $\mathbf{f}_j \cdot \mathbf{f}_k = \delta_{jk}$, with the origin in the center of mass of a molecule. The constant equilibrium distances of the atoms from the molecule's center of mass c_i^α, $i=1,2,3$.[1] which are required for defining of the internal coordinate system[1] are then obtained from vectors $\mathbf{c}_\alpha$, $\alpha=1,\ldots,4$, determined as

$$\mathbf{c}_\alpha = \mathbf{a}_\alpha - \mathbf{R} = \sum_{i=1}^{3} c_i^\alpha \mathbf{f}_i = (c_1^\alpha, c_2^\alpha, c_3^\alpha), \tag{33}$$

where $\mathbf{R}$ is

$$\begin{aligned} \mathbf{R} = \frac{\sum_\alpha m_\alpha \mathbf{a}_\alpha}{\sum_\alpha m_\alpha} = \frac{1}{m_1 + m_2 + m_3 + m_4} [& m_1 b_{1_0} \cos\theta_{1_0} \\ & + m_3 b_{2_0} + m_4 (b_{2_0} - b_{3_0} \cos\theta_{2_0}), \\ & - m_1 b_{1_0} \sin\psi_0 \sin\theta_{1_0} + m_4 b_{3_0} \sin\theta_{2_0}, \\ & m_1 b_{1_0} \cos\psi_0 \sin\theta_{1_0}]. \end{aligned} \tag{34}$$

The matrix $\boldsymbol{\mathcal{F}}^{-1/2}$, which is required to determine the unit vectors of the internal coordinate system of a molecule in the SISM[1], is calculated as

$$\boldsymbol{\mathcal{F}}^{-1/2} = \mathbf{P} \cdot \mathbf{D}^{-1/2} \cdot \mathbf{P}^T, \tag{35}$$

where $\mathbf{D}^{-1/2}$ is a diagonal matrix with the elements $D_{ii}^{-1/2} = 1/\sqrt{\lambda_i}$ and λ_i are the eigenvalues of the symmetric positive definite Gram matrix $\boldsymbol{\mathcal{F}}$. The columns of the transition matrix $\mathbf{P}$ are the eigenvectors of $\boldsymbol{\mathcal{F}}$ and $\mathbf{P}^T$ is the transpose of $\mathbf{P}$.

C. Simulation protocol

We have carried out MD simulation of a system of 256 H_2O_2 molecules with the density $\rho = 1.4425\ \mathrm{g/cm^3}$ at $T = 298$ K corresponding to the liquid state.[39] The corresponding size of the simulation box was $a = 21.6$ Å. Periodic boundary conditions were imposed to overcome the problem of surface effects; the minimum image convention was used.[2] The Coulomb interactions were truncated using the force-shifted potential[40] with a cutoff distance $r_{off} = 8.5$ Å.[41] The Lennard-Jones interactions were shifted by adding the term $C_{ij} r_{ij}^6 + D_{ij}$ to the potential, where C_{ij} and D_{ij} were chosen such that the potential and force are zero at $r_{ij} = r_{off}$.[32] The initial positions and velocities of the atoms were chosen at random. The system was then equilibrated for 50 ps where the velocities were scaled every 500 integration time steps, followed by an additional 50 ps of equilibration at constant energy of the system to ensure that the velocities assume the Maxwell distribution at $T = 298$ K. To obtain physically and numerically relevant initial conditions to perform the MD simulation of a system of flexible molecules, the equilibration was also monitored using the Vieillard–Baron rotational order parameter.[22,42]

IV. RESULTS AND DISCUSSION

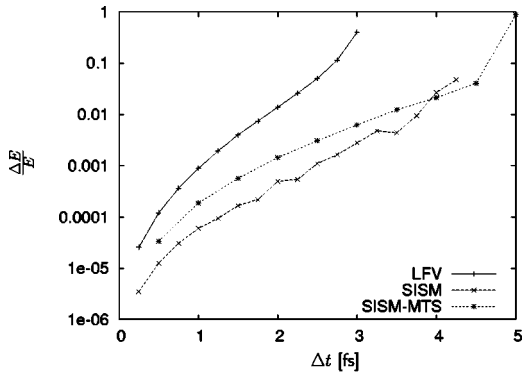
To demonstrate the effectiveness of the SISM, in all our numerical experiments we compared the computational performances for the same level of accuracy with the standard second-order LFV algorithm using an integration time step small enough to accurately describe the high-frequency molecular vibrations. In this way it is assured that the physical properties of the system determined from the trajectories computed by new integrators using long integration time steps are reliable.

For that purpose the error in total energy $\Delta E/E$ defined as

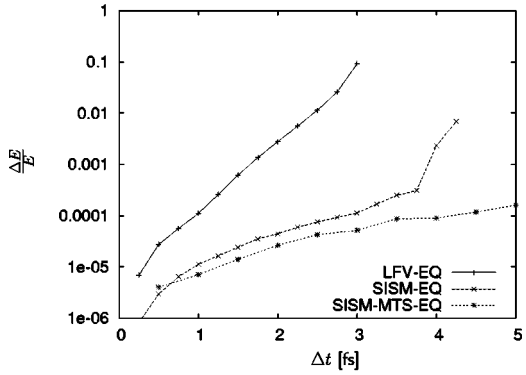
$$\frac{\Delta E}{E} = \frac{1}{M} \sum_{k=1}^{M} \left| \frac{E_0 - E_k}{E_0} \right|, \tag{36}$$

where E_0 is the initial energy, E_k is the total energy of the system at the integration step k, and M is the total number of integration steps, was monitored for all methods and was used as a measure of the efficiency and accuracy of numerical integrators for MD simulation.

The speed-up of the new methods due to a prolongation of the integration time step can be determined from the error in total energy, which is depicted in Fig. 3(a) for the system of 256 H_2O_2 molecules for the LFV, SISM, and SISM-MTS.



(a)



(b)

FIG. 3. (a) The error in the total energy of the system of $256H_2O_2$ molecules with $\rho=1.4425 g/cm^3$ at $T=298$ K using the LFV, SISM, and SISM-MTS for $M=1000$. (b) The error in the total energy of the system of $256H_2O_2$ molecules with $\rho=1.4425\ g/cm^3$ at $T=298$ K using the LFV-EQ, SISM-EQ, and SISM-MTS-EQ for $M=1000$.

The period for the antisymmetric stretching of the bond between the oxygen and the hydrogen atom in the H_2O_2 molecule is 9.9 fs (Table II). We estimate that the maximal acceptable size of the integration time step for the LFV to be 0.5 fs. From results in Fig. 3(a) we conclude that the error in total energy for a 1.25 fs integration time step in the case of the SISM corresponds to the error in the total energy using a 0.5 fs integration time step in the case of the LFV. This means that the SISM allows the use of an up to a two and a half times longer time step than the LFV for the same level of accuracy.

For the system of H_2O_2 molecules, a high amount of anharmonic forces derived from V_{ah} is expected due to the strong anharmonic potential describing the interactions in the system. Surprisingly enough the SISM-MTS, which employs $\delta t=0.5$ fs for the integration of motions generated by high-frequency anharmonic interactions is less accurate than the SISM [Fig. 3(a)]. This can be explained by the fact that the large amount of anharmonic forces stems from the strong electrostatic and van der Waals interactions and not from the high-frequency anharmonic interactions defined by V_{ah}. Due to strong intermolecular forces we expect good performance of the SISM-EQ and SISM-MTS-EQ integrators. This is confirmed by the results in Fig. 3(b) in which the error in total energy for the LFV-EQ, SISM-EQ, and SISM-MTS-EQ is shown. The total energy in this case equals expression (11). Since the vibrational potential V_{vib} taken into account in the LFV-EQ is the same as in the LFV, the estimated value of the maximal acceptable time step for the LFV-EQ is 0.5 fs.

From Fig. 3(b) we can determine that the error in total energy for the LFV-EQ with a 0.5 fs integration time step corresponds to the error in the case of the SISM-EQ using a 1.5 fs integration time step or 2.0 fs in the case of the SISM-MTS-EQ. The SISM-EQ therefore allows the use of a three times longer integration time step than the LFV-EQ for the same level of accuracy, whereas the SISM-MTS-EQ allows the use of even up to a four times longer time step as the LFV-EQ. We can also conclude that the SISM becomes unstable for integration time steps longer than 3.75 fs whereas no drift occurs in total energy using the SISM-MTS-EQ with integration time steps shorter than 5.0 fs. Using longer time steps results in a drift in the total energy, which is consistent with the conclusions in Ref. 25 where a linear numerical instability is predicted for the integration time step size corresponding to around half of the period of the fastest normal mode.

The prolongation of the maximal acceptable integration time step by the SISM-EQ and SISM-MTS-EQ in comparison to the LFV-EQ comes from the fact that the maximal acceptable integration time step is limited by the atoms' motion generated by the intermolecular forces in the case of the SISM-EQ and SISM-MTS-EQ.[22] The maximal integration time step allowed by the LFV-EQ method, however, is limited by the intramolecular high-frequency vibrations that are on the considerably smaller time scale.

We can also report resonances occurring in the error in total energy when the size of the integration time step corresponds to a multiple of the period of the fastest normal mode of a molecule (not shown in Figs. 3(a) and 3(b) where only physically meaningful integration time step sizes are considered) as in the case of the Verlet-I/r-RESPA method.[26,43] The Lennard-Jones and electrostatic interactions represent the external driving forces on the internal motion of the molecules and resonance always occurs if the frequency of the driving force corresponds to a multiple of the oscillator frequency, regardless of whether the high-frequency molecular vibrations are integrated numerically, as in the case of the Verlet-I/r-RESPA method, or analytically, as in the case of the SISM and its derivatives, which are also not stable in these resonances. Since the period of the fastest motion is 9.9 fs, we can draw the conclusion that the size of the maximal integration time step of the new methods is also limited in this case by the nonlinear instabilities at a third or fourth of the period of the fastest motion and by linear instabilities at half of the period of the fastest motion, as found for the Verlet-I/r-RESPA method.[25,44]

The actual speed-up of the new methods is confirmed by measuring the CPU time spent by the methods per integration step. The CPU times for the three methods (the SISM, SISM-MTS, and LFV) for 1000 MD steps measured on an AMD Athlon XP 1600+ processor for different system sizes (m) and equal time steps (1 fs) are given in Table III.

The results indicate that the computation cost per integration step is slightly larger for the SISM and SISM-MTS than for the LFV. However, for larger systems consisting of

TABLE III. CPU time (s) for SISM and LFV for 1000 MD steps measured on an AMD Athlon XP 1600+ processor for different system sizes (m) and equal integration time steps (1 fs).

m	t(LFV) (s)	t(SISM) (s)	t(SISM-MTS) (s)	$\frac{t(\text{SISM})}{t(\text{LFV})}$	$\frac{t(\text{SISM-MTS})}{t(\text{LFV})}$
64	14.21	16.39	18.72	1.15	1.32
128	49.57	54.51	58.51	1.10	1.18
256	170.52	182.65	188.64	1.07	1.11

more than 100 molecules the cost of computation per integration step becomes approximately the same for all methods because the computation of long-range forces, which is the same for all methods, prevails over the computation of extra transformations required by the SISM and SISM-MTS. Therefore the significant speed-up of the SISM and SISM-MTS over the LFV is due to the larger integration time step allowed by the SISM and SISM-MTS.

The applied rotational barrier of 140 kcal/mol is much higher than the estimated experimental value of $\approx$4 kcal/mol for the potential barrier in the trans transition state.[34] For efficient performance of the SISM in the case of the realistic rotational barrier, the normal mode associated with the torsional oscillation should be treated as the internal rotation and should be therefore excluded from the description by the normal coordinates.[4,5] In addition, an additional internal coordinate system should be introduced to describe the internal rotation of the molecule.[45–49]

Further verification that the new integrators yield correct trajectories using long integration time step can be obtained by computing various statistical properties of a molecular system, e.g., radial distribution functions, diffusion coefficients, orientational correlation times, vibrational spectra, for different integration time step sizes. We have performed this for a system of liquid water at ambient conditions. The results are presented in the next paper of this series[50] where we give further evidence that correct dynamics is achieved by the new integrators.

V. CONCLUSIONS

In the present paper we have tested the efficiency of newly introduced semianalytical symplectic integration methods for MD simulation on systems of nonlinear and nonplanar hydrogen peroxide molecules. The new integration methods are also stable even when using integration time steps several times longer than the standard LFV method. The numerical results show that the new integration methods allow the use of up to four times longer integration time step than the standard LFV method for the same level of accuracy. The results also indicate that the computation cost per integration time step of newly developed MD integrators is basically the same as that of the standard method. Therefore the up to a fourfold simulation speed up of new symplectic integrators is due to the larger time step they allow to use. However, much work remains to be done in the development of this approach to explore its advantages and limitations.

ACKNOWLEDGMENTS

The authors thank Dr. M. Hodošček, and Dr. F. Merzel for helpful discussions, and U. Borštnik for careful reading of the manuscript. This work was supported by the Ministry of Education, Science and Sports of Slovenia under Grant Nos. P1-0002 and J1-6331.

[1] D. Janežič, M. Praprotnik, and F. Merzel, J. Chem. Phys. **122**, 174101 (2005), preceding paper.
[2] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Clarendon, Oxford, 1987).
[3] C. Eckart, Phys. Rev. **47**, 552 (1935).
[4] A. Sayvetz, J. Chem. Phys. **7**, 383 (1939).
[5] E. B. Wilson, J. C. Decius, and P. C. Cross, *Molecular Vibrations* (McGraw-Hill, New York, 1955).
[6] J. D. Louck and H. W. Galbraith, Rev. Mod. Phys. **48**, 69 (1976).
[7] R. Rey, J. Chem. Phys. **104**, 2356 (1996).
[8] R. Rey, Chem. Phys. **229**, 217 (1998).
[9] R. Rey, J. Chem. Phys. **108**, 142 (1998).
[10] B. R. Brooks, D. Janežič, and M. Karplus, J. Comput. Chem. **16**, 1522 (1995).
[11] D. Janežič and B. R. Brooks, J. Comput. Chem. **16**, 1543 (1995).
[12] D. Janežič, R. M. Venable, and B. R. Brooks, J. Comput. Chem. **16**, 1554 (1995).
[13] J. M. Sanz-Serna and M. P. Calvo, *Numerical Hamiltonian Problems* (Chapman & Hall, London, 1994).
[14] H. F. Trotter, Proc. Am. Math. Soc. **10**, 545 (1959).
[15] G. Strang, SIAM (Soc. Ind. Appl. Math.) J. Numer. Anal. **5**, 506 (1968).
[16] H. Yoshida, Phys. Lett. A **150**, 262 (1990).
[17] L. Verlet, Phys. Rev. **159**, 98 (1967).
[18] H. Goldstein, *Classical Mechanics*, 2nd ed. (Addison-Wesley, New York, 1980).
[19] D. Janežič and F. Merzel, J. Chem. Inf. Comput. Sci. **35**, 321 (1995).
[20] D. Janežič and M. Praprotnik, Int. J. Quantum Chem. **84**, 2 (2001).
[21] M. Praprotnik and D. Janežič, Cell. Mol. Biol. Lett. **7**, 147 (2002).
[22] D. Janežič and M. Praprotnik, J. Chem. Inf. Comput. Sci. **43**, 1922 (2003).
[23] B. Garcia-Archilla, J. M. Sanz-Serna, and R. D. Skeel, SIAM J. Sci. Comput. (USA) **20**, 930 (1998).
[24] J. A. Izaguirre, S. Reich, and R. D. Skeel, J. Chem. Phys. **110**, 9853 (1999).
[25] R. D. Skeel and J. A. Izaguirre, in *Computational Molecular Dynamics: Challenges, Methods, Ideas*, Lecture Notes in Computational Science and Engineering Vol. 4, edited by P. Deuflhard, J. Hermans, B. Leimkuhler, A. E. Mark, S. Reich, R. D. Skeel (Springer, Berlin, 1999), pp. 318–331.
[26] H. Grubmüller, H. Heller, A. Windemuth, and K. Schulten, Mol. Simul. **6**, 121 (1991).
[27] J. Koput, Chem. Phys. Lett. **236**, 516 (1995).
[28] V. A. Benderskii, I. S. Irgibaeva, E. V. Vetoshkin, and H. P. Trommsdorff, Chem. Phys. **262**, 369 (2000).
[29] R. L. Redington, W. B. Olson, and P. C. Cross, J. Chem. Phys. **36**, 1311 (1962).
[30] W. R. Busing and H. A. Levy, J. Chem. Phys. **42**, 3054 (1965).
[31] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, J. Chem. Phys. **79**, 926 (1983).
[32] P. J. Steinbach and B. R. Brooks, J. Comput. Chem. **15**, 667 (1994).
[33] J. Koput, Chem. Phys. Lett. **257**, 36 (1996).
[34] P. A. Giguere, J. Chem. Phys. **18**, 88 (1950).
[35] R. C. Taylor, J. Chem. Phys. **18**, 898 (1950).

[36]O. Bain and P. A. Giguere, Can. J. Chem. **33**, 527 (1955).
[37]L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics: Mechanics*, 3rd ed. (Pergamon, Oxford, 1976), vol. 1.
[38]W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes: The Art of Scientific Computing* (Cambridge University Press, Cambridge, 1987).
[39]*CRC Handbook of Chemistry and Physics*, 66th ed., edited by R. C. Weast (CRC, Cleveland, 1986).
[40]C. L. Brooks III, B. M. Pettitt, and M. Karplus, J. Chem. Phys. **83**, 5897 (1985).
[41]M. Prevost, D. van Belle, G. Lippens, and S. Wodak, Mol. Phys. **71**, 587 (1990).
[42]J. E. Vieillard-Baron, J. Chem. Phys. **56**, 4729 (1967).
[43]M. E. Tuckerman, B. J. Berne, and G. J. Martyna, J. Chem. Phys. **97**, 1990 (1992).
[44]Q. Ma, J. A. Izaguirre, and R. D. Skeel, SIAM J. Sci. Comput. (USA) **24**, 1951 (2003).
[45]J. B. Howard, J. Chem. Phys. **5**, 442 (1937).
[46]J. B. Howard, J. Chem. Phys. **5**, 451 (1937).
[47]B. Kirtman, J. Chem. Phys. **37**, 2516 (1962).
[48]B. Kirtman, J. Chem. Phys. **41**, 775 (1964).
[49]B. Kirtman, J. Chem. Phys. **49**, 2257 (1968).
[50]M. Praprotnik and D. Janežič, J. Chem. Phys. **122**, 174103 (2005), following paper.