Symplectic Molecular Dynamics Integration Using Normal Mode Analysis

DUŠANKA JANEŽIČ, MATEJ PRAPROTNIK

National Institute of Chemistry, Hajdrihova 19, 1000 Ljubljana, Slovenia

Received 3 September 2000; revised 13 February 2001; accepted 20 February 2001

ABSTRACT: The split integration symplectic method (SISM) for molecular dynamics (MD) integration using normal mode analysis based on a factorization of the Liouville propagator is presented. This approach is quite distinct from others that use fractional-step methods, owing to the analytical treatment of high-frequency motions. The method involves splitting the total Hamiltonian of the system into a harmonic part and the remaining part. Then the Hamilton equations are solved using a second-order generalized leapfrog integration scheme in which the purely harmonic Hamiltonian (which represents the main contribution of the chemical bonds and angles) is treated analytically, i.e., independent of the step size of integration, by a normal mode analysis that is carried out only once, at the beginning of calculation. The whole integration step combines analytical evolution of the harmonic part of the Hamiltonian with a correction arising from the remaining part. The proposed algorithm requires only one force evaluation per integration step. The algorithm was tested on a simple system of linear chain molecules. Results demonstrate the method makes possible the integration of the MD equations over larger time steps without loss of stability while being economical in computer time.

Key words: molecular dynamics simulation; normal mode analysis; symplectic integration methods; Hamiltonian systems; linear chain molecules

Introduction

Computer simulation methods, based on quantum and statistical mechanics theories [1], have already made significant contributions to the understanding of structure and dynamics of biological macromolecules, particularly proteins and nucleic acids, and thus to the connection with their functional properties. Many more fundamental problems of structure and dynamics are yet unsolved or, at least, are not sufficiently elaborated. The progress in this direction is slowed down by both the conceptual problems in the currently used methods and by the need for excessive amounts of computer time.

Molecular dynamics (MD) simulations have been applied to a wide variety of biological macromole-

Correspondence to: D. Janežič; e-mail: dusa@kihp5.cmm.ki.si.
Contract grant sponsor: Ministry of Science and Technology of Slovenia.
Contract grant numbers: P1-0104-503; J1-0104-7346; S41-104-001.

cules [2, 3]. The major obstacle that reduces the efficiency of MD simulation methods lays in the inability to sample sufficient portions of phase space. This is connected with the vast disparity in time scales between the fast vibrations of the individual atoms and the slow conformational changes [4]. A number of MD algorithms for solving Hamiltonian systems have been proposed [5]. Most of the methods presently used exhibit instability unless the time step is small enough [6]. The simplest and most commonly used algorithm is leapfrog-Verlet [7]. Its application to MD problems leads to severe restrictions on the step size. But, for studies of dynamics of large molecular systems, larger time steps in the MD integration procedure are needed. The problem of how to increase the time step in the MD integration can be overcome by the use of symplectic methods for the numerical solution of Hamilton equations. Hamiltonian systems possess an important property, in that the flow in the phase space is symplectic; hence the numerical methods for solving these systems are expected to reproduce this property [8].

There are several equivalent ways of endowing a phase space with a symplectic structure. A mapping is said to be symplectic or canonical if it preserves the differential form $d\mathbf{p} \wedge d\mathbf{q}$, which defines the symplectic structure in the phase space. Differential forms provide a geometric interpretation of symplecticness in terms of conservation of areas, which follows from Liouville's theorem. In one-degree-of-freedom example symplecticness is the preservation of oriented area. An example is the harmonic oscillator where the t-flow is just a rigid rotation and the area is preserved. The area-preserving character of the solution operator holds only for Hamiltonian systems. In more then one-degree-of-freedom examples the preservation of area is symplecticness rather than preservation of volume [8].

Symplectic integration methods replace the t-flow in the phase space by the symplectic transformation, which inherits the symplectic character of the Hamiltonian flow.

To perform MD simulation of a system with a finite number of degrees of freedom the Hamilton equations of motion

$$\frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i}, \qquad \frac{dq_i}{dt} = \frac{\partial H}{\partial p_i}, \qquad i = 1, \ldots, d, \quad (1)$$

where H is the Hamiltonian, q_i and p_i are the coordinate and momentum, respectively, and d is the number of degrees of freedom are to be solved.

Normal mode analysis is an alternative approach to MD. The basic assumption is that the potential energy function can be approximated by a sum of quadratic terms in displacements. To determine the vibrational motions of the system, the eigenvalues and eigenvectors of a mass-weighted matrix of the second derivatives of potential function has to be calculated. For a nonlinear molecule composed of N atoms, $3N - 6$ eigenvalues provide the normal or fundamental vibrational frequencies of the vibration and the associated eigenvectors, called normal modes, give the directions and relative amplitudes of the atomic displacements in each mode. Linear molecules are exceptional in that they have only $3N - 5$ normal modes of vibration.

Harmonic analysis (normal modes) at given temperature and curvature gives complete time behavior of the system in the harmonic limit [9–11]. Although the harmonic model may be incomplete because of the contribution of anharmonic terms to the potential energy, it is nevertheless of considerable importance because it serves as a first approximation for which the theory is highly developed. This model is also useful in the split integration symplectic method (SISM), which uses normal mode analysis [12–14].

In this article, we describe the further development of SISM for MD integration. The difference from our previous work is that we have extended the method to more general MD potential function. In the following section, the details of the method are given. Then, the leapfrog-Verlet algorithm is presented, and finally the numerical experiments are presented and discussed.

Split Integration Symplectic Method for MD Integration

The explicit symplectic integrator can be derived in terms of free Lie algebra in which Hamilton equations (1) are written in the form [8]

$$\frac{d\mathbf{x}}{dt} = \{\mathbf{x}, H\} = \hat{L}_H \mathbf{x}, \quad (2)$$

where $\{\mathbf{x}, H\}$ denotes the Poisson bracket [15], $\hat{L}_H$ is the Poisson bracket operator, and $\mathbf{x} = (\mathbf{q}, \mathbf{p})$ is a vector in the phase space composed of the coordinates and momenta of all particles. The formula

$$\mathbf{x}|_{t_0+\Delta t} = \exp\left(\Delta t \hat{L}_H\right)\mathbf{x}|_{t_0} \quad (3)$$

provides a way for integrating the Hamiltonian system in terms of Lie operators. It is the formal

solution of Hamilton equations or the exact time evolution of trajectories in phase space from t_0 to $t_0 + \Delta t$, and Δt is a time step. The trouble with it lays in the impossibility of evaluating $\exp(\Delta t \hat{L}_H)$. The Lie formalism used is the key in the development of symplectic integration. Symplectic integration consists in replacing $\exp(\Delta t \hat{L}_H)$ by a product of symplectic maps, which approximate $\exp(\Delta t \hat{L}_H)$ to a given order [16].

The construction of an efficient algorithm rests on the ability to separate the Hamiltonian into parts that are themselves integrable and also efficiently computable [17]. Suppose that the MD Hamiltonian H is split into two parts as

$$H = H_0 + H_r. \tag{4}$$

The following approximation for (3) can be used

$$\exp(\Delta t \hat{L}_H) \approx \exp\left(\frac{\Delta t}{2}\hat{L}_{H_0}\right)\exp(\Delta t \hat{L}_{H_r}) \times \exp\left(\frac{\Delta t}{2}\hat{L}_{H_0}\right). \tag{5}$$

Using the Baker–Campbell–Hausdorff formula for combining the exponents of two Lie group elements [8], it follows that the approximation in Eq. (5) is quadratic. Inserting Eq. (5) into Eq. (3) gives

$$\mathbf{x}|_{t_0+\Delta t} \approx \exp\left(\frac{\Delta t}{2}\hat{L}_{H_0}\right)\exp\left(\Delta t \hat{L}_{H_r}\right) \times \exp\left(\frac{\Delta t}{2}\hat{L}_{H_0}\right)\mathbf{x}|_{t_0}, \tag{6}$$

which prescribes how to propagate from one point in phase space to another. First, the system is propagated a half step evolution with H_0, then a whole step with H_r, and finally another half step with H_0. This scheme is called the generalized leapfrog [7, 19] and was widely used as a basis for development of multiple time-step MD integration algorithms [18, 19]. This integration scheme was used in the development of SISM, a second-order symplectic integration algorithm for MD integration.

This algorithm is quite distinct from other approaches using the fractional-step methods [18, 19] owing to the analytical treatment of high-frequency motions. Knowing the pure harmonic Hamiltonian H_0, which represents the main contribution of the strong chemical bonds and angles, we can treat the cumbersome high-frequency terms analytically, i.e., independent of the step size of integration. The whole integrating step combines analytical evolution of the harmonic part of the Hamiltonian with a correction arising from the remaining part.

In this integration scheme the partitioned part of the MD Hamiltonian H_0 describes the vibrational motion of the system. It represents the dynamically leading contribution whose potential depends only on constant parameters of the simulation. This separation allows us to calculate normal modes only once, at the beginning of the calculation. In this term are included all bonding and angle bending interactions within the harmonic approximation. The motion governed by H_0 is resolved by means of normal coordinates that rotate in phase space with the corresponding vibrational frequencies which are obtained by solving the secular equation at the outset of the calculation [9].

For a chosen model MD Hamiltonian

$$H = \sum_i \frac{\mathbf{p}_i^2}{2m_i} + \sum_{\text{bonds}} k_b(b - b_0)^2 + \sum_{\text{angles}} k_\phi(\phi - \phi_0)^2 + \sum_{i>j}\frac{e_i e_j}{r_{ij}} + \sum_{i>j} 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right], \tag{7}$$

where i and j run over all atoms, e_i denotes the charge on the ith atom, and r_{ij} is the distance between atoms i and j. We can define

$$H_0 = H^{\text{harm}}(m_i, b_0, \phi_0, k_b, k_\phi), \tag{8}$$

$$H_r = H - H_0, \tag{9}$$

where H^{harm} denotes the harmonic approximation; m_i is the mass of the ith atom, b_0 and ϕ_0 are reference values for bond lengths and angles, respectively, and k_b and k_ϕ are corresponding force constants.

Following the procedure defined by Eq. (6), the SISM for MD integration is written explicitly as follows:

Step 0: Perform normal mode calculation for the harmonic part (H_0) of the Hamiltonian to get frequencies and normal modes of vibration. Normal modes are represented by normal coordinates, hereafter denoted P and Q, obtained by means of the transformational matrix $\mathbf{A}$. The columns of matrix $\mathbf{A}$ are the normal mode vectors.

Step 1: Propagate by H_0 for the time $\Delta t/2$. This corresponds to the rotation of normal coordinates, P_i^0 and Q_i^0, in the phase space by the corresponding vibrational frequency ω_i:

$$\begin{bmatrix} P_i' \\ Q_i' \end{bmatrix} = \begin{bmatrix} \cos(\omega_i \frac{\Delta t}{2}) & -\omega_i \sin(\omega_i \frac{\Delta t}{2}) \\ (\frac{1}{\omega_i})\sin(\omega_i \frac{\Delta t}{2}) & \cos(\omega_i \frac{\Delta t}{2}) \end{bmatrix} \begin{bmatrix} P_i^0 \\ Q_i^0 \end{bmatrix}. \tag{10}$$

For $\omega_i \neq 0$ in Eq. (10) vibrations are taken into account, and for $\omega_i = 0$ translations and rotations. Since $\lim_{x\to 0} \sin x/x = 1$ for $\omega_i = 0$ from Eq. (10), follow the equations that describe translations and rotations in normal coordinates:

$$P_i' = P_i^0, \tag{11}$$

$$Q_i' = P_i^0 \frac{\Delta t}{2} + Q_i^0. \tag{12}$$

Step 2: Intermediate step in preparation for step 3. Transformation from normal coordinates P_k' and Q_k' to Cartesian displacement coordinates p_i' and q_i':

$$\begin{aligned} p_i' &= \sqrt{m_i} \sum_k \mathbf{A}_{ik} P_k', \\ q_i' &= \frac{1}{\sqrt{m_i}} \sum_k \mathbf{A}_{ik} Q_k'. \end{aligned} \tag{13}$$

Step 3: Evolve with H_r by means of numerical integration:

$$\begin{aligned} p_i'' &= p_i' - \Delta t \left(\frac{\partial H_r}{\partial q_i} \right)_{q_i = q_i'}, \\ q_i'' &= q_i' + \Delta t \left(\frac{\partial H_r}{\partial p_i} \right)_{p_i = p_i'}. \end{aligned} \tag{14}$$

Since $H_r = H_r(\mathbf{q})$,

$$\left[\left(\frac{\partial H_r}{\partial p_i} \right)_{p_i = p_i'} = 0 \right],$$

only the momentum changes at this step.

Step 4: Intermediate step in the preparation for step 5. Back transformation from Cartesian displacement coordinates p_k'' and q_k'' to normal coordinates P_i'' and Q_i'':

$$\begin{aligned} P_i'' &= \sum_k \frac{1}{\sqrt{m_k}} \mathbf{A}_{ik}^T p_k'', \\ Q_i'' &= \sum_k \sqrt{m_k}\, \mathbf{A}_{ik}^T q_k''. \end{aligned} \tag{15}$$

Step 5: Again the rotation of normal coordinates in the phase space:

$$\begin{bmatrix} P_i \\ Q_i \end{bmatrix} = \begin{bmatrix} \cos(\omega_i \frac{\Delta t}{2}) & -\omega_i \sin(\omega_i \frac{\Delta t}{2}) \\ (\frac{1}{\omega_i}) \sin(\omega_i \frac{\Delta t}{2}) & \cos(\omega_i \frac{\Delta t}{2}) \end{bmatrix} \begin{bmatrix} P_i'' \\ Q_i'' \end{bmatrix}. \tag{16}$$

This concludes one whole integration step. At this point, the transformation from normal coordinates to Cartesian displacement coordinates has to be performed whether the physical properties of the system, e.g., energy and displacements, are to be derived.

Step 6: Return to step 1 ($P_i \to P_i^0$, $Q_i \to Q_i^0$) and continue for the desired number of time steps.

This treatment of high-frequency terms in the Hamiltonian permits the SISM proposed in Eqs. (10)–(16) to use a much larger integration step size Δt than the standard methods [12–14].

Leapfrog-Verlet Algorithm

The simplest of the numerical techniques for the integration of equations of motion is the leapfrog-Verlet (LFV) algorithm, which is known to be symplectic and of second order. The name leapfrog steams from the fact that coordinates and velocities are calculated at different times.

The MD Hamiltonian H of the system is the sum of kinetic and potential energy:

$$H = T + U, \tag{17}$$

$$T = \sum_i \frac{\mathbf{p}_i^2}{2m_i}, \tag{18}$$

$$\begin{aligned} U = \sum_{\text{bonds}} k_b (b - b_0)^2 + \sum_{\text{angles}} k_\phi (\phi - \phi_0)^2 + \sum_{i>j} \frac{e_i e_j}{r_{ij}} \\ + \sum_{i>j} 4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right], \end{aligned} \tag{19}$$

where m_i is the mass of the ith atom, b_0 and ϕ_0 are reference values for bond lengths and angles, respectively, and k_b and k_ϕ are corresponding force constants; i and j run over all atoms, e_i denotes the charge on the ith atom and r_{ij} is the distance between atoms i and j, ε_{ij} and σ_{ij} are the corresponding constants of Lennard-Jones potential.

Using the approximation

$$\begin{aligned} \exp(\Delta t \hat{L}_H) \approx \exp\left(\frac{\Delta t}{2} \hat{L}_T \right) \exp(\Delta t \hat{L}_U) \\ \times \exp\left(\frac{\Delta t}{2} \hat{L}_T \right), \end{aligned} \tag{20}$$

the LFV integration method propagates coordinates and momenta on the basis of the equation of motion (1) by the following relations:

$$q_i' = q_i + \frac{p_i}{m} \frac{\Delta t}{2}, \tag{21}$$

$$p_{i+1} = p_i - \Delta t \left(\frac{\partial U}{\partial q} \right)_{q = q_i'}, \tag{22}$$

$$q_{i+1} = q_i' + \frac{p_{i+1}}{m} \frac{\Delta t}{2}, \tag{23}$$

where q_i is the coordinate, p_i is the momentum, $\dim(p_i, q_i) = d$, d is the number of degrees of freedom, Δt is the time step, and m is the mass of the corresponding atom.

Numerical Experiment

The SISM described above was evaluated on the model system composed of 20 linear butadiyne molecules, a six-atom molecule, of the form H—(—C≡C—)$_2$—H.

Atom positions in molecules were described in the Cartesian coordinate system. Relative displacements from equilibrium atom positions in molecules were described in the internal coordinate system, which is defined by three orthogonal unit vectors. A molecule with zero displacements was chosen as an equilibrium position for each molecule in a system. In accordance with the theory of small oscillations, that the atom displacements at the equilibrium are small, the molecule is not deformed. For this reason we have defined the unit vector $\mathbf{e}_{3,j}$, which defines the z direction of the internal coordinate system of the jth molecule directing along the molecule at the time when the molecule is still linear, as

$$\mathbf{e}_{3,j} = \frac{\mathbf{q}^0_{N,j} - \mathbf{r}_{\mathbf{T}j}}{|\mathbf{q}^0_{N,j} - \mathbf{r}_{\mathbf{T}j}|}, \tag{24}$$

where $|\cdot|$ denotes the norm of the vector, $\mathbf{q}^0_{N,j}$ is the radius vector to the equilibrium position of the last atom in the jth molecule in the Cartesian coordinate system, and $\mathbf{r}_{\mathbf{T}j}$ is the radius vector to the center of mass of the jth molecule:

$$\mathbf{r}_{\mathbf{T}j} = \frac{\sum_i m_i \mathbf{q}_{i,j}}{\sum_i m_i}, \tag{25}$$

where m_i is the mass of the ith atom in the jth molecule, and $\mathbf{q}_{i,j}$ is the radius vector; i runs over all atoms in the jth molecule, and j runs over all molecules. The remaining two unit vectors $\mathbf{e}_{1,j}$ and $\mathbf{e}_{2,j}$ were defined to be orthogonal to each other and to $\mathbf{e}_{3,j}$. The internal coordinate system of the jth molecule is by $\mathbf{e}_{1,j}$, $\mathbf{e}_{2,j}$, and $\mathbf{e}_{3,j}$ fully defined. It changes at each step in correspondence with normal modes with zero frequency, which does not deform the molecule.

Let $a^0_{i,j}$ be the equilibrium distance of the ith atom in the jth molecule from the center of mass of the jth molecule. Then the radius vector to the equilibrium position of the ith atom in the jth molecule in the Cartesian coordinate system is given by

$$\mathbf{q}^0_{i,j} = \mathbf{r}_{\mathbf{T}j} + a^0_{i,j}\mathbf{e}_{3,j}. \tag{26}$$

Displacement vector of the ith atom in the jth molecule from the equilibrium position in the Cartesian coordinate system is

$$\Delta\mathbf{q}_{i,j} = \mathbf{q}_{i,j} - \mathbf{q}^0_{i,j}, \tag{27}$$

where $\mathbf{q}_{i,j}$ is the radius vector to the ith atom.

In Figure 1 the displacements of the jth molecule in the internal and the Cartesian coordinate systems are shown.

Relative displacement coordinates of atoms in the jth molecule are obtained by projecting $\Delta\mathbf{q}_{i,j}$ onto the unit vectors $\mathbf{e}_{1,j}$, $\mathbf{e}_{2,j}$, and $\mathbf{e}_{3,j}$ as

$$\begin{aligned}\Delta\mathbf{q}_{i,j} &= (\Delta x_{i,j},\ \Delta y_{i,j},\ \Delta z_{i,j}) \\ &= (\Delta\mathbf{q}_{i,j}\cdot\mathbf{e}_{1,j},\ \Delta\mathbf{q}_{i,j}\cdot\mathbf{e}_{2,j},\ \Delta\mathbf{q}_{i,j}\cdot\mathbf{e}_{3,j}).\end{aligned} \tag{28}$$

Back transformation from relative displacement coordinates to the Cartesian coordinates is given by

$$\Delta\mathbf{q}_{i,j} = \Delta x_{i,j}\mathbf{e}_{1,j} + \Delta y_{i,j}\mathbf{e}_{2,j} + \Delta z_{i,j}\mathbf{e}_{3,j}, \tag{29}$$

$$\mathbf{q}_{i,j} = \Delta\mathbf{q}_{i,j} + \mathbf{q}^0_{i,j}. \tag{30}$$

The harmonic part of the potential energy for the jth molecule can be written as

$$U_0 = \frac{1}{2}\sum_{i=1}^{N-1} k_{b_i}\Delta r^2_{i,j} + \frac{1}{2}\sum_{i=2}^{N-1} k_{\phi_i}\phi^2_{i,j}, \tag{31}$$

where

$$\Delta r_{i,j} = \sqrt{(\mathbf{q}_{i+1,j} - \mathbf{q}_{i,j})^2} - \sqrt{(\mathbf{q}^0_{i+1,j} - \mathbf{q}^0_{i,j})^2}, \tag{32}$$

and N is the number of atoms in the molecule. The first term in Eq. (31) describes the stretching of the molecule. Because of small relative displacements in a transversal direction $\Delta r_{i,j}$ can be approximated as

$$\Delta r_{i,j} \approx \Delta z_{i+1,j} - \Delta z_{i,j}. \tag{33}$$

The second term in Eq. (31) describes the bending of the molecule. The reference angle $\phi^0_{i,j}$ is 0 because the molecule is linear. Using relative displacement coordinates $\phi_{i,j}$ can be expressed as a scalar product of the unit vector pointing from the $(i-1)$th atom to the ith atom, $\mathbf{u}_{i-1,j}$, and the unit vector pointing from the ith atom to the $(i+1)$th atom, $\mathbf{u}_{i,j}$. The distance between two sequential atoms, $d_{i,j}$, is approximately constant:

$$d_{i,j} \approx d_{i,j_0} = d_{i_0}, \tag{34}$$

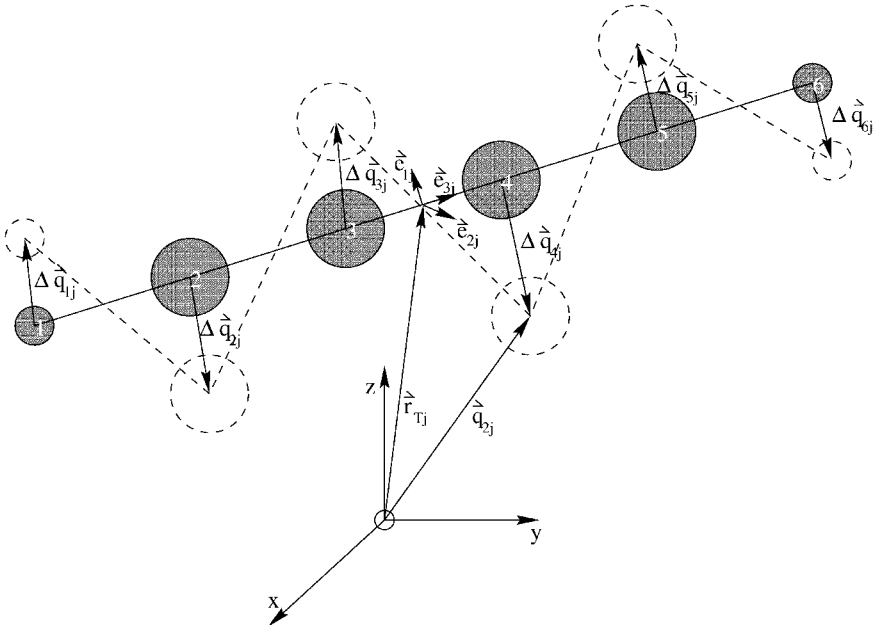

FIGURE 1. Displacements of the *j*th molecule in the internal and the Cartesian coordinate system.

where d_{i,j_0} is the equilibrium distance between $(i+1)$th and the ith atom in the jth molecule. Since the equilibrium distances among atoms in the molecule are equal for all molecules, the index j in Eq. (34) can be omitted. Therefore

$$\cos\phi_{i,j} = \mathbf{u}_{i-1,j} \cdot \mathbf{u}_{i,j}, \tag{35}$$

where

$$\mathbf{u}_{i-1,j} = \frac{1}{\sqrt{1+\left(\frac{\Delta x_{i,j}-\Delta x_{i-1,j}}{d_{i-1_0}}\right)^2+\left(\frac{\Delta y_{i,j}-\Delta y_{i-1,j}}{d_{i-1_0}}\right)^2}} \times \begin{bmatrix} \frac{\Delta x_{i,j}-\Delta x_{i-1,j}}{d_{i-1_0}} \\ \frac{\Delta y_{i,j}-\Delta y_{i-1,j}}{d_{i-1_0}} \\ 1 \end{bmatrix}, \tag{36}$$

$$\mathbf{u}_{i,j} = \frac{1}{\sqrt{1+\left(\frac{\Delta x_{i+1,j}-\Delta x_{i,j}}{d_{i_0}}\right)^2+\left(\frac{\Delta y_{i+1,j}-\Delta y_{i,j}}{d_{i_0}}\right)^2}} \times \begin{bmatrix} \frac{\Delta x_{i+1,j}-\Delta x_{i,j}}{d_{i_0}} \\ \frac{\Delta y_{i+1,j}-\Delta y_{i,j}}{d_{i_0}} \\ 1 \end{bmatrix}. \tag{37}$$

Because the first two components of vectors $\mathbf{u}_{i-1,j}$ and $\mathbf{u}_{i,j}$ are much smaller than 1, the square root can be expanded by the formula $1/\sqrt{1+x} = 1 - x/2$ where $x \ll 1$. Because $\phi_{i,j}$ is almost 0, after expanding $\cos\phi_{i,j} = 1 - \phi_{i,j}^2/2$, we get

$$\begin{aligned} 1 - \frac{\phi_{i,j}^2}{2} = &\left[1 - \frac{1}{2}\left(\left(\frac{\Delta x_{i,j}-\Delta x_{i-1,j}}{d_{i-1_0}}\right)^2 + \left(\frac{\Delta y_{i,j}-\Delta y_{i-1,j}}{d_{i-1_0}}\right)^2\right)\right] \\ &\times \left[1 - \frac{1}{2}\left(\left(\frac{\Delta x_{i+1,j}-\Delta x_{i,j}}{d_{i_0}}\right)^2 + \left(\frac{\Delta y_{i+1,j}-\Delta y_{i,j}}{d_{i_0}}\right)^2\right)\right] \\ &\times \begin{bmatrix} \frac{\Delta x_{i,j}-\Delta x_{i-1,j}}{d_{i-1_0}} \\ \frac{\Delta y_{i,j}-\Delta y_{i-1,j}}{d_{i-1_0}} \\ 1 \end{bmatrix}^T \begin{bmatrix} \frac{\Delta x_{i+1,j}-\Delta x_{i,j}}{d_{i_0}} \\ \frac{\Delta y_{i+1,j}-\Delta y_{i,j}}{d_{i_0}} \\ 1 \end{bmatrix}. \end{aligned} \tag{38}$$

Due to harmonic approximation all higher than second-order terms in the Eq. (38) are neglected. Using relative displacement coordinates $\phi_{i,j}$ can be expressed as

$$\frac{\phi_{i,j}^2}{2} = \frac{1}{2}\left[\left(\frac{\Delta x_{i,j}-\Delta x_{i-1,j}}{d_{i-1_0}}\right)^2 + \left(\frac{\Delta x_{i+1,j}-\Delta x_{i,j}}{d_{i_0}}\right)^2 + \left(\frac{\Delta y_{i,j}-\Delta y_{i-1,j}}{d_{i-1_0}}\right)^2 + \left(\frac{\Delta y_{i+1,j}-\Delta y_{i,j}}{d_{i_0}}\right)^2\right]$$

$$-\frac{(\Delta x_{i,j}-\Delta x_{i-1,j})(\Delta x_{i+1,j}-\Delta x_{i,j})}{d_{i-1_0}d_{i_0}}$$
$$-\frac{(\Delta y_{i,j}-\Delta y_{i-1,j})(\Delta y_{i+1,i}-\Delta y_{i,j})}{d_{i-1_0}d_{i_0}}. \quad (39)$$

Displacements in x and y directions in the Eq. (39) are equivalent because of the symmetry around the z axis.

The matrix $\mathbf{T}^{-1/2}\mathbf{V}\mathbf{T}^{-1/2}$, where $\mathbf{V} = \nabla^2 U_0$ and U_0 is the harmonic part of the potential energy and $\mathbf{T} = \mathrm{diag}\,(m_i)$ a diagonal mass matrix, is symmetric with positive eigenvalues. In the example of the six-atom molecule, it is a 18×18 matrix.

Because relative displacements in the x, y, and z directions are uncoupled, the matrix is composed of three 6×6 block diagonal matrices only, and due to z axis symmetry the first two block matrices are equal. Then the matrix $\mathbf{T}^{-1/2}\mathbf{V}\mathbf{T}^{-1/2}$ can be written as

$$\begin{bmatrix} \mathbf{X} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{X} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{Z} \end{bmatrix}. \quad (40)$$

Diagonal elements of the matrix $\mathbf{X}$ are

$$\begin{aligned} X_{1,1} &= \frac{1}{m_1}\frac{k_{\phi_2}}{d_{1_0}^2}, \\ X_{2,2} &= \frac{1}{m_2}\left(\frac{2k_{\phi_2}}{d_{1_0}d_{2_0}} + \frac{k_{\phi_2}}{d_{1_0}^2} + \frac{k_{\phi_2}}{d_{2_0}^2} + \frac{k_{\phi_3}}{d_{2_0}^2}\right), \\ X_{6,6} &= \frac{1}{m_6}\frac{k_{\phi_5}}{d_{5_0}^2}, \\ X_{5,5} &= \frac{1}{m_5}\left(\frac{2k_{\phi_5}}{d_{4_0}d_{5_0}} + \frac{k_{\phi_5}}{d_{4_0}^2} + \frac{k_{\phi_5}}{d_{5_0}^2} + \frac{k_{\phi_4}}{d_{4_0}^2}\right), \end{aligned} \quad (41)$$

$$X_{i,i} = \frac{1}{m_i}\left(\frac{2k_{\phi_i}}{d_{i-1_0}d_{i_0}} + \frac{k_{\phi_i}}{d_{i-1_0}^2} + \frac{k_{\phi_i}}{d_{i_0}^2} + \frac{k_{\phi_{i-1}}}{d_{i-1_0}^2} + \frac{k_{\phi_{i+1}}}{d_{i_0}^2}\right), \quad 3 \le i \le 4. \quad (42)$$

Off-diagonal elements are

$$X_{1,2} = X_{2,1} = -\frac{1}{\sqrt{m_1 m_2}}\left(\frac{k_{\phi_2}}{d_{1_0}d_{2_0}} + \frac{k_{\phi_2}}{d_{1_0}^2}\right), \quad (43)$$

$$X_{5,6} = X_{6,5} = -\frac{1}{\sqrt{m_5 m_6}}\left(\frac{k_{\phi_5}}{d_{4_0}d_{5_0}} + \frac{k_{\phi_5}}{d_{5_0}^2}\right), \quad (44)$$

$$X_{i-1,i} = X_{i,i-1} = -\frac{1}{\sqrt{m_{i-1}m_i}}\left(\frac{k_{\phi_i}}{d_{i-1_0}d_{i_0}} + \frac{k_{\phi_i}}{d_{i-1_0}^2} + \frac{k_{\phi_{i-1}}}{d_{i-2_0}d_{i-1_0}} + \frac{k_{\phi_{i-1}}}{d_{i-1_0}^2}\right), \quad 3 \le i \le 5, \quad (45)$$

$$X_{i+1,i-1} = X_{i-1,i+1} = \frac{1}{\sqrt{m_{i-1}m_{i+1}}}\frac{k_{\phi_i}}{d_{i-1_0}d_{i_0}}, \quad 2 \le i \le 5. \quad (46)$$

Diagonal elements of matrix $\mathbf{Z}$ are

$$Z_{1,1} = \frac{1}{m_1}k_{b_1}, \qquad Z_{6,6} = \frac{1}{m_6}k_{b_5}, \quad (47)$$

$$Z_{i,i} = \frac{1}{m_i}(k_{b_i} + k_{b_{i-1}}), \qquad 2 \le i \le 5. \quad (48)$$

Off-diagonal elements are

$$Z_{i+1,i} = Z_{i,i+1} = -\frac{1}{\sqrt{m_i m_{i+1}}}k_{b_i}, \qquad 1 \le i \le 5. \quad (49)$$

Solving the eigenvalue problem for a matrix $\mathbf{T}^{-1/2}\mathbf{V}\mathbf{T}^{-1/2}$, we obtain 18 normal frequencies out of which five equal 0. One zero frequency describes the translation of the molecule in the z direction of the internal coordinate system and is determined by the matrix $\mathbf{Z}$. The remaining four are double degenerated and are determined by matrix $\mathbf{X}$. Two frequencies describe translation of the molecule along the x and y direction of the internal coordinate system, and the other two describe the rotation of the molecule; 13 vibrational frequencies are nonzero. Five of them are singlets and represent the stretching motion of the molecule. They are determined by the matrix $\mathbf{Z}$. The remaining eight are double degenerated. They describe the bending motion of the molecule. Degeneration occurs because of the symmetry around the z axis.

In the remaining part of the Hamiltonian, H_r in Eq. (7), also unharmonic contributions due to the difference between the true quadratic potential and the harmonic approximation are included. The true distance $\Delta\mathbf{r}_{i,j}$ is calculated by means of Eq. (32). The true angle $\phi_{i,j}$ is calculated as

$$\phi_{i,j} = \arccos\left(\frac{(\mathbf{q}_{i,j} - \mathbf{q}_{i-1,j}) \cdot (\mathbf{q}_{i+1,j} - \mathbf{q}_{i,j})}{|\mathbf{q}_{i,j} - \mathbf{q}_{i-1,j}||\mathbf{q}_{i+1,j} - \mathbf{q}_{i,j}|}\right). \quad (50)$$

The initial atomic positions for these molecules were taken at the equilibrium, and the initial velocities were distributed according to the Maxwell distribution at 300 K, and the periodic boundary conditions were imposed. Potential parameters were taken the same as in Ref. [12].

In order to test the importance of including bending term $\sum_{\text{angles}} k_\phi(\phi - \phi_0)^2$ for the linear molecule into the potential function, we first performed the MD simulation of the linear molecule with no bending term included in Eq. (7). In Figure 2 the butadiyne molecule after 100 steps of MD simulation using SISM with time step of 1 fs is presented. At

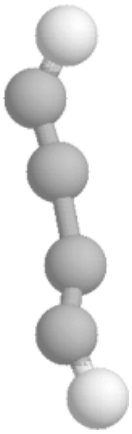

FIGURE 2. Butadiyne molecule after 100 steps of MD simulation using SISM with time step of 1 fs with no bending term included in the potential function.

the beginning of the simulation the molecule was linear and during the simulation it has deformed. Therefore the term $\sum_{\text{angles}} k_\phi(\phi - \phi_0)^2$ has to be included in the Hamiltonian to prevent the molecule from deformation.

Figure 3 presents a butadiyne molecule after 100 steps of MD simulation using SISM with time step of 1 fs with bending term included in the potential function. It can be observed that the molecule remains close to linear.

The primary metric used to quantify the accuracy of the results of simulation is conservation of total energy. The long-term stability for a variety of time steps is evident from results shown in Figure 4, which depicts the time evolution of the total energy

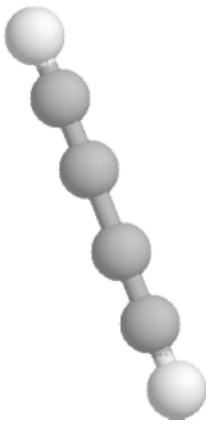

FIGURE 3. Butadiyne molecule after 100 steps of MD simulation using SISM with time step of 1 fs with bending term included in the potential function.

for the system of 20 butadiyne molecules for various time steps using SISM for a trajectory length of 9 ps. The density of the system was chosen to be $\varrho = 0.1$ g/cm^3. The results presented in Figure 5 show the LFV method for the same model system for time steps of 1, 2, and 3 fs. The total energy oscillates around its initial value without exhibiting growth, while for a time step of 3 fs the excise growth is seen. In comparison with SISM, which conserves the total energy for time steps of 1, 2, 3, 4, and also 4.5 fs.

To test the accuracy and the efficiency of SISM with the LFV algorithm, we compared computational performances for the same level of accuracy. The relative error in energy, δE, defined as

$$\delta E = \frac{|E_0 - E_i|}{E_0}, \tag{51}$$

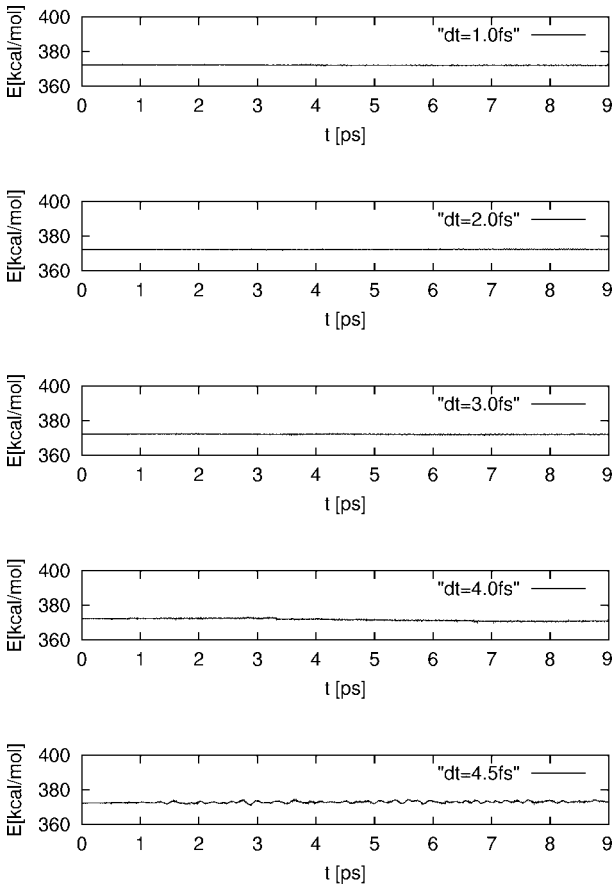

FIGURE 4. Time history of the total energy for the system of 20 butadiyne molecules for time steps of 1, 2, 3, 4, and 4.5 fs using SISM for a trajectory of length of 9 ps.

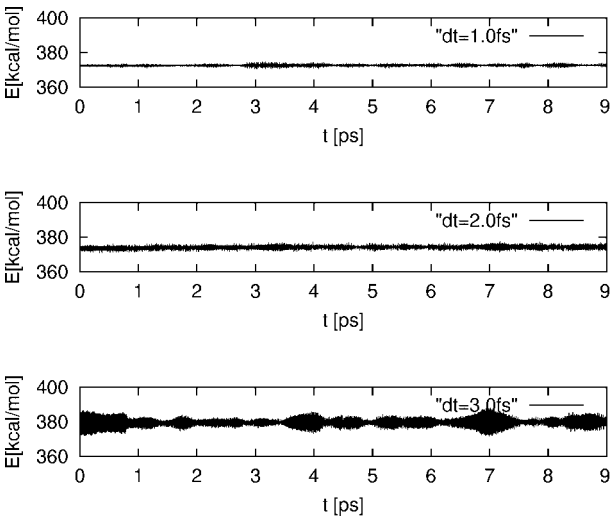

FIGURE 5. Time history of the total energy for the system of 20 butadiyne molecules for time steps of 1, 2, and 3 fs using LFV for a trajectory of length of 9 ps.

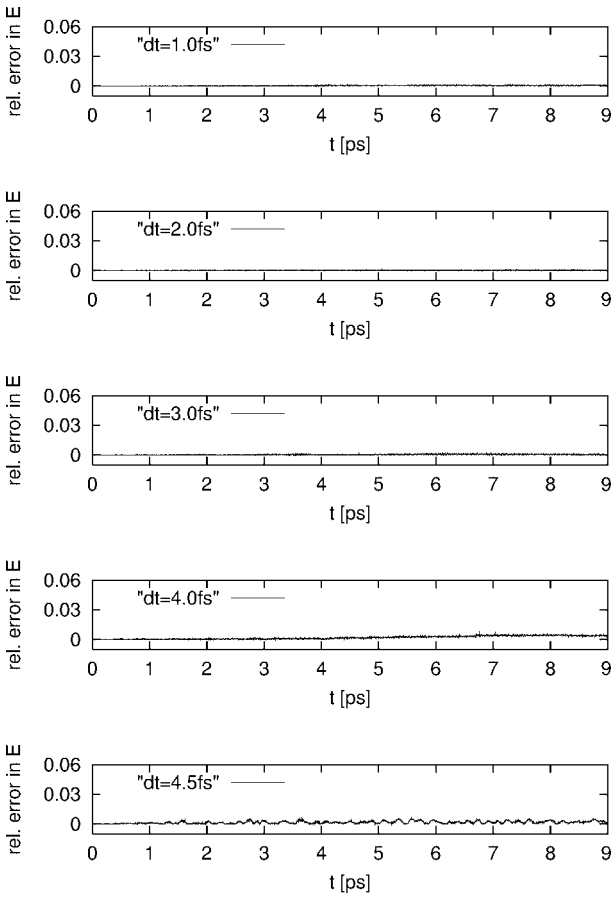

FIGURE 6. Relative error in total energy for the SISM. Results are plotted for time steps of 1, 2, 3, 4, and 4.5 fs for a system of 20 butadiyne molecules for a trajectory of length of 9 ps.

where E_0 is the initial energy, and E_i is the energy at time step i that was monitored for both methods.

The results of the relative error in the total energy for test molecules, a system of 20 butadiyne molecules, for different time steps using two different methods (SISM and LFV) are presented in Figures 6 and 7. It can be observed that the value δE using the LFV algorithm is in all examples presented larger than the value obtained by SISM. The results show that LFV is giving the acceptable relative error in total energy only for a time step up to 1 fs in comparison with SISM, which gives a similar error for a time step of 4.5 fs. It is obvious that for the same level of accuracy, the SISM allows to use up to five times larger then time step the LFV method. Furthermore, the LFV method is numerically stable for only short time steps.

The CPU time required by the two methods (SISM and LFV) for 100 MD steps computed on the PII/200 MHz processor for different system sizes n and equal time step 1 fs are compared in Table I. The computation cost per integration step is approximately the same for both methods so that the speed up of SISM over LFV is determined mainly by the difference in time step, which is significant.

Symplectic integration methods, which are also symmetric, are time reversible. Since SISM is symplectic and symmetric, it must be time reversible. To test this fact we calculated the trajectories for the model system of 20 butadiyne molecules for negative time also. In Table II the total energy for our model system using SISM, for two different time steps (1 and 2 fs) and different integration times. Results demonstrate the time reversibility of the SISM.

Conclusions

The present work presents a design and analysis of a split integration symplectic method (SISM) for MD simulation, which uses normal mode analysis for an analytical treatment of high-frequency motions within a second-order generalized leapfrog Verlet (LFV) scheme.

Even if the complexity of the SISM seems to be high compared to that of the LFV method, it is not so since the most demanding part of the SISM, i.e., the calculation of nonbonding forces and energy, is the

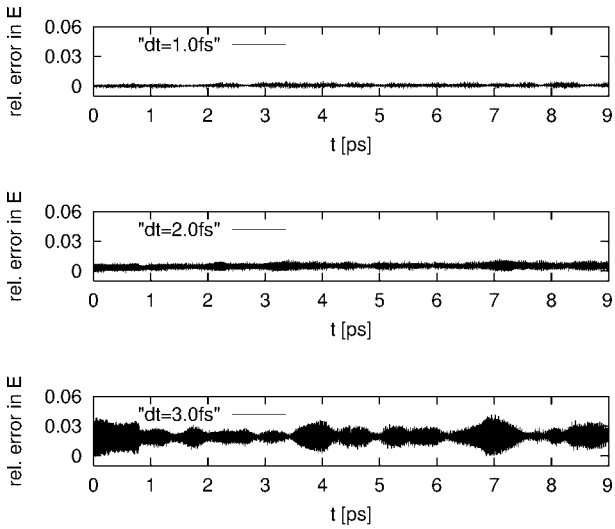

FIGURE 7. Relative error in total energy for the LFV algorithm. Results are plotted for time steps of 1, 2, and 3 fs for a system of 20 butadiyne molecules for a trajectory of length of 9 ps.

same as that of LFV. All the extra work (coordinate transformations) is thus prevailed by nonbonding force and energy calculation. It was demonstrated that for the model system of 20 linear butadiyne molecules SISM allows use of a step size up to five times larger then standard LFV method, and the computational cost per step is approximately the same for the two methods. This is in agreement with our previous work [12–14] where no bending term was included in the potential energy function.

Further improvements in efficiency were achieved by implementing the methods on computers with highly parallel architecture. It was shown that SISM offers better performances in comparison with LFV for both sequential and parallel implementation. SISM performs in parallel as LFV which means the speed up is gained due to a longer time step, which can be used by SISM [20–22].

Much work remains to be done in the development of this approach to explore the advantages and limitations of the method. The method will be extended to force fields that include torsional terms in the MD potential.

TABLE I

CPU time (s) for SISM and LFV for 100 MD steps of the system of butadiyne molecules for different system sizes *n* for equal time steps of 1 fs on the PII/200 MHz.

n	*t*(SISM) (s)	*t*(LFV) (s)
10	13.46	12.91
20	52.11	51.51
30	115.73	115.82
40	204.58	205.96
50	318.60	321.51

TABLE II

Total energy of the system in kcal/mol at different times of integration and time steps Δt.

	Δt (ps)	
t (ps)	0.001	0.002
0.000	372.4295391	372.4295391
0.050	372.4342472	372.4319225
0.100	372.4251498	372.4202091
0.150	372.4243393	372.4352490
0.200	372.4196697	372.4295630
0.250	372.4299782	372.4267537
0.300	372.4463068	372.4519875
0.350	372.4327027	372.4373704
0.400	372.4477789	372.4657149
0.450	372.4510523	372.4557296
0.450	372.4510523	372.4557296
0.400	372.4477789	372.4657149
0.350	372.4327027	372.4373704
0.300	372.4463068	372.4519875
0.250	372.4299782	372.4267537
0.200	372.4196697	372.4295630
0.150	372.4243393	372.4352490
0.100	372.4251498	372.4202091
0.050	372.4342472	372.4319225
0.000	372.4295391	372.4295391

ACKNOWLEDGMENTS

The authors express their thanks to Dr. Franci Merzel and Dr. Milan Hodošček for helpful discussions that arose during the course of this work as well as to Urban Borštnik for assistance with Figure 1.

References

1. Allen, M. P.; Tildesley, D. J. Computer Simulation of Liquids; Clarendon: Oxford, 1987.
2. McCammon, J. A.; Harvey, S. C. Dynamics of Proteins and Nucleic Acids; Cambridge University Press: Cambridge, 1987.

3. Smith, J. C. In van Gunsteren, W. F.; Weiner, P. K.; Wilkinson, A. J., Eds. Computer Simulations of Biomolecules, Vol. 3; Kluwer/ESCOM: Leiden, 1997; pp. 305–360.
4. Loncharich, R. J.; Brooks, B. R. Proteins 1989, 6, 32–45.
5. Schlick, T.; Barth, E.; Mandziuk, M. Ann Rev Biophys Biomol Struct 1997, 26, 181–222.
6. Pastor, R. W.; Brooks, B. R.; Szabo, A. Mol Phys 1988, 65, 1409–1419.
7. Grubmüller, H.; Heller, H.; Windemuth, A.; Schulten, K. Molec Simul 1991, 6, 121–142.
8. Sanz-Serna, J. M.; Calvo, M. P. Numerical Hamiltonian Problems; Chapman and Hall: London, 1994.
9. Brooks, B. R.; Janežič, D.; Karplus, M. J Comput Chem 1995, 16, 1522–1542.
10. Janežič, D.; Brooks, B. R. J Comput Chem 1995, 16, 1543–1553.
11. Janežič, D.; Venable, R. M.; Brooks, B. R. J Comput Chem 1995, 16, 1554–1566.
12. Janežič, D.; Merzel, F. J Chem Inf Comput Sci 1995, 35, 321–326.
13. Janežič, D.; Merzel, F. J Chem Inf Comput Sci 1997, 37, 1048–1054.
14. Janežič, D.; Merzel, F. In Deuflhard, P.; Hermans, J.; Leimkuhler, B.; Mark, A.; Reich, S.; Skeel, R., Eds. Computational Molecular Dynamics: Challenges, Methods, Ideas; Lecture Notes in Computational Science and Engineering, Vol. 4; Springer: Berlin, 1999; pp. 332–348.
15. Goldstein, H. Classical Mechanics; Addison-Wesley: Reading, MA, 1965.
16. Forest, E.; Ruth, R. D. Phys D 1990, 43, 105–117.
17. Wisdom, J.; Holman, M. Astron J 1991, 102, 1528–1538.
18. Humphreys, D. D.; Friesner, R. A.; Berne, B. J. J Chem Phys 1994, 98, 6885–6892.
19. Leimkuhler, B. J.; Reich, S.; Skeel, R. D. In Mesirov, J.; Schulten, K., Eds. IMA Volumes in Mathematics and Its Applications, Vol. 82; Springer: Berlin, 1995; pp. 161–186.
20. Trobec, R.; Jerebic, I.; Janežič, D. Parallel Computing 1993, 19, 1029–1039.
21. Trobec, R.; Merzel, F.; Janežič, D. J Chem Inf Comput Sci 1997, 37, 1055–1062.
22. Trobec, R.; Šterk, M.; Praprotnik, M.; Janežič, D. Int J Quantum Chem, to appear.