Molecular dynamics integration and molecular vibrational theory. I. New symplectic integrators

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

New symplectic integrators have been developed by combining molecular dynamics integration with the standard theory of molecular vibrations to solve the Hamiltonian equations of motion. The presented integrators analytically resolve the internal high-frequency molecular vibrations by introducing a translating and rotating internal coordinate system of a molecule and calculating normal modes of an isolated molecule only. The translation and rotation of a molecule are treated as vibrational motions with the vibrational frequency zero. All types of motion are thus described in terms of the normal coordinates. The method's time reversibility requirement was used to determine the equations of motion for internal coordinate system of a molecule. The calculation of long-range forces is performed numerically within the generalized second-order leap-frog scheme, in the same way as in standard second-order symplectic methods. The new methods for integrating classical equations of motion using normal mode analysis allow us to use a long integration step and are applicable to any system of molecules with one equilibrium configuration. © *2005 American Institute of Physics*. [DOI: 10.1063/1.1884607]

I. INTRODUCTION

The standard integrators for solving the classical equations of motion are the second-order symplectic leap-frog Verlet (LFV) algorithm[1] and its variants. Their power lies in their simplicity since the only required information about the studied physical system are its interacting potential and the time scale of the fastest motion in the system, which determines the integration time step size. Therefore they are employed for solving dynamics problems in a variety of scientific fields, e.g., molecular dynamics (MD) simulation,[2,3] celestial mechanics,[4–6] and accelator physics.[7] However, in the case of MD integration, the integration time step size is severely limited due to the numerical treatment of the high-frequency molecular vibrations, which represent the fastest motion in the system.[8] Therefore a huge number of integration steps is usually required to accurately sample the phase space composed of all the coordinates and momenta of all the particles. This is a time consuming task and is often too demanding for the capabilities of contemporary computers.

One way of overcoming the limitation of the standard methods' integration time step size is to freeze the high-frequency motions using constraints as in SHAKE and RATTLE.[9,10] The problem of such an approach is that the treatment of bonds as constraints prevents the generation of low-frequency modes from coupled vibrations and it also prevents the relaxation of vibrations under the influence of an external field.[11] Also, the high-frequency peaks cannot be reproduced in computed vibrational spectra. An alternative way is to use multiple integration time stepping schemes, e.g., Verlet-I/r-RESPA,[12,13] where the high-frequency molecular vibrations are numerically integrated with corre-spondingly shorter integration time steps as used for remaining low-frequency motion. The increase of the integration time step size is obtained by adding extra information about the studied physical system with respect to the LFV algorithm, i.e., all time scales and different types of interactions in the system must be considered.[14–21] Thus there is a trade-off between a method's simplicity and general applicability to different kinds of physical systems (LFV algorithm) and a method's increased efficiency for a specific physical systems, i.e., molecular systems (Verlet-I/r-RESPA).

Another approach is the analytical treatment of high-frequency molecular vibrations, which requires the standard theory of molecular vibrations[22] to be built into the integration method. In this way the fast degrees of freedom are rigorously treated and not removed, as in the case of rigid-body dynamics,[23–25] where small molecules are treated as rigid bodies. The first attempt in this direction was introduced in Ref. 26 where an integration algorithm for MD simulations of an isolated linear molecule using the splitting of the total Hamiltonian into the high-frequency harmonic and low-frequency remaining part was presented. The high-frequency molecular vibrations, which were described only by the harmonic bond stretching potential, were resolved analytically using the normal coordinates. The method introduced in Ref. 26 was extended in Refs. 27 and 28 to also treat systems of linear molecules. An important methodological step in the development of this approach is presented in Ref. 29 where normal modes with frequency zero were first used for describing the rotation and translation of molecules. In addition the angle bending term was also included for analytical description of the vibrational motion. A study of the system density dependence of the size of the maximal allowed integration time step is presented in Ref. 30. However, the equations defining the equilibrium configuration of

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

each molecule and its motion, which are the fundamental components of the described method, have not been presented yet.

A major goal of the present paper is to present a new semianalytical symplectic MD integration method split integration symplectic method (SISM), which combines the standard theory of molecular vibrations and MD integration. The key feature in which the presented approach differs from all other similar approaches for MD integration is the analytical description of coupled molecular vibrations, which is possible only by using the normal coordinates[22] and a translating and rotating internal coordinate system of each molecule.[31,32] The description of molecular motion in terms of normal coordinates provided by the standard theory of molecular vibrations is also extended to translation and rotation, which are treated as vibration with zero normal mode frequency. The presented MD integration methods are thus the direct implementation of the standard theory of molecular vibrations to calculate the MD trajectories of a molecular system, where all degrees of freedom are treated classically. Similar approaches either apply the standard theory of molecular vibrations to analytically treat the quantum-mechanical intramolecular vibrational degrees of freedom,[33,34] or do not use the standard theory of molecular vibrations, in particular, the concept of Eckart frame,[31] and can therefore treat the molecular vibrations only as one-dimensional harmonic oscillators with a single high-frequency of oscillation and a fixed direction-in space, e.g., NAPA.[12,14]

Another goal is to describe two variants of the SISM. First, the multiple time stepping SISM (SISM-MTS) which besides an analytical treatment of high-frequency molecular vibrational motion uses also a shorter integration time step for the numerical integration of high-frequency anharmonic vibration terms and a correspondingly longer time step for the dynamics generated by the electrostatic and van der Waals interactions. Second, the equilibrium SISM (SISM-EQ), which also treats high-frequency vibrations analytically but the van der Waals and electrostatic potential energies are calculated from the equilibrium positions of atoms instead of from the actual positions of the atoms in a molecule. In this way a new averaging function, introduced here, is derived from the standard theory of molecular vibrations, to mollify the impulse of the electrostatic and van der Waals forces as proposed in Refs. 35–37, which improves the stability of the presented integration methods. Combining the SISM-MTS and SISM-EQ approaches leads to the equilibrium multiple time stepping SISM (SISM-MTS-EQ) method by which the optimum enlargement of the integration time step is achived.

Shematic presentation of the new symplectic integrators is given in Fig. 1. These methods are applied in the accompanying papers[38,39] to various simulations of different systems of molecules with one equilibrium configuration and no internal rotation. The results indicate that the presented integrators, due to the analytical treatment of high-frequency motions, allow considerably longer integration time steps than the standard LFV algorithm for the same computational accuracy and computational cost per integration step.
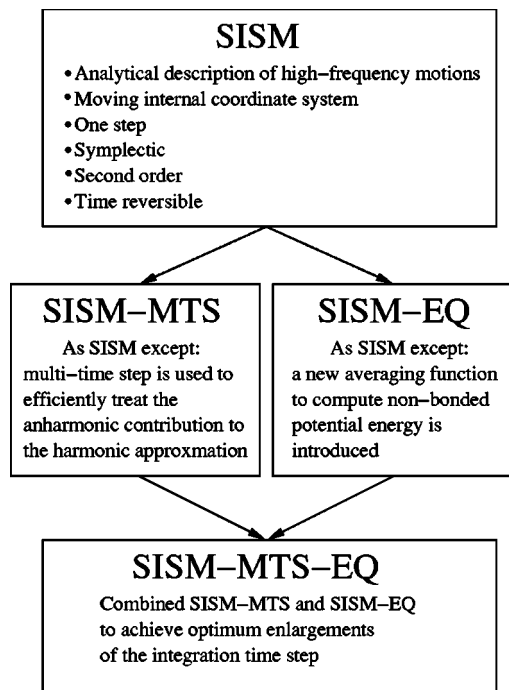

FIG. 1. New symplectic integrators.

II. MOLECULAR MOTION CHARACTERIZED IN NORMAL COORDINATES

To describe the atoms' motion in molecules by normal coordinates we have modified the dynamical molecular model introduced in Ref. 32 for molecules with only one equilibrium configuration and no internal rotation. Using this approach we are able to define the dynamics of the internal coordinate system of a molecule and describe vibrational, translational, and rotational motion in normal coordinates.

First, we introduce a fixed Cartesian coordinate system defined by the orthogonal right-handed triad of unit vectors $\mathbf{e}_i$, $i = 1, 2, 3$ and denote the instantaneous position vectors of N atoms of a molecule of the system relative to its origin by $\mathbf{r}_\alpha$, $\alpha = 1, 2, \ldots, N$ as presented in Fig. 2. The molecule center-of-mass vector $\mathbf{R}$ is defined as

$$\mathbf{R} = \sum_\alpha m_\alpha \mathbf{r}_\alpha \bigg/ \sum_\alpha m_\alpha. \tag{1}$$

Next, we introduce the translating and rotating internal coordinate system of the molecule, which is attached to the molecule and moves with it. In the standard theory of molecular vibrations,[22] the Eckart frame,[31,32] defined in Appendix A, is chosen to be the internal coordinate system. The Eckart frame is the internal coordinate system where the coupling between the vibrational and rotational degrees of freedom of a molecule is zero at equilibrium. In principle, one can also choose an alternative reference frame for the internal coordinate system, as we propose in the present work.

The equilibrium positions of atoms relative to the molecule's center of mass are then given by vectors $\mathbf{c}_\alpha$,

$$\mathbf{c}_\alpha = \sum_i c_i^\alpha \mathbf{f}_i, \tag{2}$$

where the orthogonal right-handed triad of unit vectors $\mathbf{f}_i$, $i = 1, 2, 3$ with the origin in the center of mass of a molecule
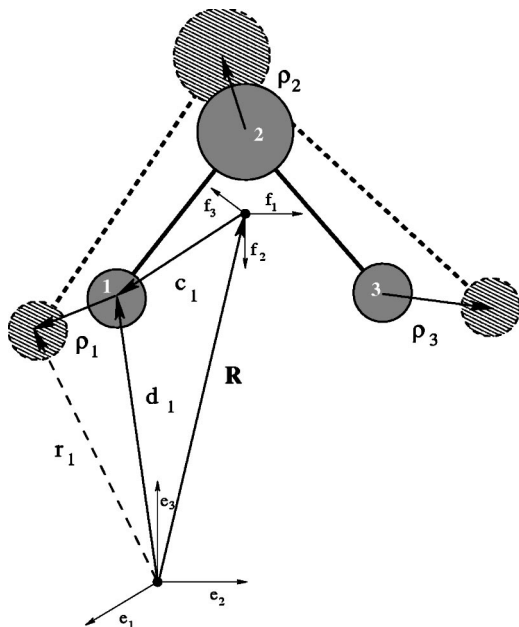
FIG. 2. Atom displacements in a triatomic molecule in the Cartesian and the internal coordinate system.

defines the internal coordinate system of the molecule (see Fig. 2).

The equilibrium positions of atoms in the Cartesian coordinate system are given by

$$\mathbf{d}_\alpha = \mathbf{R} + \mathbf{c}_\alpha. \tag{3}$$

The displacement vectors of atoms from their equilibrium positions are therefore defined as

$$\boldsymbol{\rho}_\alpha = \mathbf{r}_\alpha - \mathbf{d}_\alpha. \tag{4}$$

If the displacement vectors are written in terms of the internal coordinate system, the relative Cartesian displacement coordinates are

$$\begin{aligned}
\boldsymbol{\rho}_\alpha &= (\boldsymbol{\rho}_\alpha \cdot \mathbf{f}_1)\mathbf{f}_1 + (\boldsymbol{\rho}_\alpha \cdot \mathbf{f}_2)\mathbf{f}_2 + (\boldsymbol{\rho}_\alpha \cdot \mathbf{f}_3)\mathbf{f}_3 \\
&= \Delta x_\alpha \mathbf{f}_1 + \Delta y_\alpha \mathbf{f}_2 + \Delta z_\alpha \mathbf{f}_3,
\end{aligned} \tag{5}$$

where $\mathbf{f}_1, \mathbf{f}_2,$ and $\mathbf{f}_3$ point along $x, y,$ and z directions of the internal coordinate system, respectively.

Using the internal coordinate system, the velocity of each atom can be split into three parts:[22]

$$\mathbf{v}_\alpha = \dot{\mathbf{r}}_\alpha = \dot{\mathbf{R}} + \boldsymbol{\Omega} \times (\mathbf{r}_\alpha - \mathbf{R}) + \boldsymbol{\Delta}\mathbf{v}_\alpha = \mathbf{v}_\alpha^{\text{trans}} + \mathbf{v}_\alpha^{\text{rot}} + \mathbf{v}_\alpha^{\text{vib}}, \tag{6}$$

where $\dot{\mathbf{R}}$ is the center-of-mass velocity and represents the translational contribution, $\boldsymbol{\Omega} \times (\mathbf{r}_\alpha - \mathbf{R})$ is the contribution due to rotation of the internal coordinate system, and $\boldsymbol{\Delta}\mathbf{v}_\alpha$ is the velocity of the α atom in the internal coordinate system

$$\boldsymbol{\Delta}\mathbf{v}_\alpha = \frac{d\Delta x_\alpha}{dt}\mathbf{f}_1 + \frac{d\Delta y_\alpha}{dt}\mathbf{f}_2 + \frac{d\Delta z_\alpha}{dt}\mathbf{f}_3 \tag{7}$$

and represents the vibrational contribution.

The angular velocity $\boldsymbol{\Omega}$ of the Eckart frame is obtained by differentiating equation for Eckart conditions, Eq. (A4), with respect to time

$$\boldsymbol{\Omega} = \mathbf{J}'^{-1} \cdot \sum_\alpha m_\alpha \mathbf{c}_\alpha \times (\dot{\mathbf{r}}_\alpha - \dot{\mathbf{R}}), \tag{8}$$

where $\mathbf{J}' \in \mathbb{R}^{3\times3}$ is defined as[40,41]

$$\mathbf{J}' = \sum_\alpha m_\alpha \{[(\mathbf{r}_\alpha - \mathbf{R}) \cdot \mathbf{c}_\alpha]\mathbf{I} - (\mathbf{r}_\alpha - \mathbf{R}) \otimes \mathbf{c}_\alpha\}. \tag{9}$$

Here $\otimes$ is a tensor product of two vectors defined as $\mathbf{A} = \mathbf{a} \otimes \mathbf{b}(A_{ik} = a_i b_k)$ and $\mathbf{I} \in \mathbb{R}^{3\times3}$ represents the identity matrix. Note that $\mathbf{J}'$ in the case of nonrigid molecules differs from the tensor of a molecule's instantaneous moment of inertia.[42] The unit vectors $\mathbf{f}_i$, $i=1, 2, 3$, of the Eckart frame vary according to

$$\dot{\mathbf{f}}_i = \boldsymbol{\Omega} \times \mathbf{f}_i, \quad i = 1,2,3. \tag{10}$$

The kinetic energy can be, by virtue of Eq. (A4), expressed as[22]

$$\begin{aligned}
T &= \frac{1}{2}\sum_\alpha m_\alpha \mathbf{v}_\alpha^2 = \frac{1}{2}\dot{\mathbf{R}}^2\sum_\alpha m_\alpha \\
&\quad + \frac{1}{2}\sum_\alpha m_\alpha[\boldsymbol{\Omega} \times (\mathbf{r}_\alpha - \mathbf{R})] \cdot [\boldsymbol{\Omega} \times (\mathbf{r}_\alpha - \mathbf{R})] \\
&\quad + \frac{1}{2}\sum_\alpha m_\alpha \Delta v_\alpha^2 + \boldsymbol{\Omega} \cdot \sum_\alpha m_\alpha(\boldsymbol{\rho}_\alpha \times \boldsymbol{\Delta}\mathbf{v}_\alpha) \\
&= T_{\text{trans}} + T_{\text{rot}} + T_{\text{vib}} + T_{\text{Coriolis}}.
\end{aligned} \tag{11}$$

The first term represents the translational energy of a molecule, the second rotational energy, the third vibrational energy, and the fourth one the coupling between rotation and vibration—the Coriolis energy. Translation is completely separated from the vibrational and rotational motion while rotation and vibration are coupled.

Using Eq. (6) we can write Eq. (7) as

$$\begin{aligned}
\boldsymbol{\Delta}\mathbf{v}_\alpha &= \frac{d\Delta x_\alpha}{dt}\mathbf{f}_1 + \frac{d\Delta y_\alpha}{dt}\mathbf{f}_2 + \frac{d\Delta z_\alpha}{dt}\mathbf{f}_3 = \Delta v_{\alpha_x}\mathbf{f}_1 + \Delta v_{\alpha_y}\mathbf{f}_2 \\
&\quad + \Delta v_{\alpha_z}\mathbf{f}_3 = (\boldsymbol{\Delta}\mathbf{v}_\alpha \cdot \mathbf{f}_1)\mathbf{f}_1 + (\boldsymbol{\Delta}\mathbf{v}_\alpha \cdot \mathbf{f}_2)\mathbf{f}_2 + (\boldsymbol{\Delta}\mathbf{v}_\alpha \cdot \mathbf{f}_3)\mathbf{f}_3 \\
&= \{[\mathbf{v}_\alpha - \dot{\mathbf{R}} - \boldsymbol{\Omega} \times (\mathbf{r}_\alpha - \mathbf{R})] \cdot \mathbf{f}_1\}\mathbf{f}_1 + \{[\mathbf{v}_\alpha - \dot{\mathbf{R}} - \boldsymbol{\Omega} \\
&\quad \times (\mathbf{r}_\alpha - \mathbf{R})] \cdot \mathbf{f}_2\}\mathbf{f}_2 + \{[\mathbf{v}_\alpha - \dot{\mathbf{R}} - \boldsymbol{\Omega} \times (\mathbf{r}_\alpha \\
&\quad - \mathbf{R})] \cdot \mathbf{f}_3\}\mathbf{f}_3.
\end{aligned} \tag{12}$$

From Eq. (12) follows that only the vibrational part of the atom velocity $\boldsymbol{\Delta}\mathbf{v}_\alpha = \mathbf{v}_\alpha^{\text{vib}}$ is expressed in terms of the relative Cartesian displacement coordinates and consequently in terms of normal coordinates.[22]

Alternatively, we can describe also rotation and translation of a molecule in terms of the normal coordinates. To do so the whole atom velocity needs to be expressed in terms of the relative Cartesian displacement coordinates

$$\begin{aligned}
\boldsymbol{\Delta}\mathbf{v}_\alpha &= \Delta v_{\alpha_x}\mathbf{f}_1 + \Delta v_{\alpha_y}\mathbf{f}_2 + \Delta v_{\alpha_z}\mathbf{f}_3 \\
&= (\boldsymbol{\Delta}\mathbf{v}_\alpha \cdot \mathbf{f}_1)\mathbf{f}_1 + (\boldsymbol{\Delta}\mathbf{v}_\alpha \cdot \mathbf{f}_2)\mathbf{f}_2 + (\boldsymbol{\Delta}\mathbf{v}_\alpha \cdot \mathbf{f}_3)\mathbf{f}_3 \\
&= (\mathbf{v}_\alpha \cdot \mathbf{f}_1)\mathbf{f}_1 + (\mathbf{v}_\alpha \cdot \mathbf{f}_2)\mathbf{f}_2 + (\mathbf{v}_\alpha \cdot \mathbf{f}_3)\mathbf{f}_3.
\end{aligned} \tag{13}$$

From Eqs. (6) and (13) follows

$$\Delta \mathbf{v}_\alpha = \mathbf{v}_\alpha = \dot{\mathbf{r}}_\alpha = \dot{\mathbf{R}} + \mathbf{\Omega} \times (\mathbf{r}_\alpha - \mathbf{R}) + \Delta \mathbf{v}_\alpha, \tag{14}$$

which implies that

$$\dot{\mathbf{R}} + \mathbf{\Omega} \times (\mathbf{r}_\alpha - \mathbf{R}) = 0, \tag{15}$$

and therefore

$$\dot{\mathbf{R}} = 0, \tag{16}$$

$$\mathbf{\Omega} = 0. \tag{17}$$

Equation (16) requires the origin of the internal coordinate system to be at rest and from Eq. (17) follows that $\dot{\mathbf{f}}_i = \mathbf{\Omega} \times \mathbf{f}_i = 0$, $i = 1, 2, 3$. These conditions demand that the internal coordinate system is fixed (the velocity in the internal coordinate system must be equal to the velocity in the Cartesian coordinate system), which means that the equilibrium configuration of the molecule is also at rest,

$$\Delta \mathbf{v}_\alpha = \dot{\boldsymbol{\rho}}_\alpha - \mathbf{\Omega} \times \boldsymbol{\rho}_\alpha = \dot{\boldsymbol{\rho}}_\alpha = \dot{\mathbf{r}}_\alpha - \dot{\mathbf{d}}_\alpha = \dot{\mathbf{r}}_\alpha = \mathbf{v}_\alpha \Rightarrow \dot{\mathbf{d}}_\alpha = 0. \tag{18}$$

It is important to emphasize that the dynamics of the internal coordinate system obeying the conditions (16) and (17) is different from the dynamics of the Eckart frame [see Eq. (8)], which is employed in the standard theory of molecular vibrations. Due to the translational and rotational motion, the position of the center of mass as well as the orientation of a molecule change. The requirements (16) and (17) reflect the change in understanding of the translational and rotational motion of a molecule which are here treated as vibrations with zero frequency. The motion of the center of mass and the orientation of a molecule are thus described by the equations of motion for the normal coordinates with frequency zero.

The vibrational potential energy can be expanded in a power series of the relative displacements of atoms from their equilibrium positions as

$$V_{\text{vib}} = V_{\text{vib}_0} + \sum_{i=1}^{3N} \left(\frac{\partial V_{\text{vib}}}{\partial \Delta q_i} \right)_0 \Delta q_i$$

$$+ \frac{1}{2} \sum_{i,j=1}^{3N} \left(\frac{\partial^2 V_{\text{vib}}}{\partial \Delta q_i \, \partial \Delta q_j} \right)_0 \Delta q_i \Delta q_j + \cdots$$

$$= \frac{1}{2} \sum_{i,j=1}^{3N} \left(\frac{\partial^2 V_{\text{vib}}}{\partial \Delta q_i \, \partial \Delta q_j} \right)_0 \Delta q_i \Delta q_j + \cdots, \tag{19}$$

where

$$\Delta \mathbf{q} = (\Delta q_1, \Delta q_2, \dots, \Delta q_{3N})$$

$$= (\Delta x_1, \Delta y_1, \Delta z_1, \dots, \Delta x_N, \Delta y_N, \Delta z_N)$$

is a vector of the relative displacements of atoms from their equilibrium positions. We set $V_{\text{vib}_0} = 0$ and $(\partial V_{\text{vib}} / \partial \Delta q_i)_0 = 0$ because there are no forces acting on atoms in the equilibrium positions.[22]

The standard theory of molecular vibrations considers only motions in which all coordinates vary by infinitesimal amounts. In molecular systems, however, the atoms' coordi-

nates change by finite amounts due to the translational and rotational motion of molecules. By introducing the Eckart frame for the internal coordinate system of a molecule, the problem of vibrational motion of a molecule can be considered independently from its rotation and translation.[22] Using the Eckart frame we can define equilibrium positions of the atoms in a molecule in such a way that the displacements of the atoms are sufficiently small, so that the normal mode analysis can be applied.

In normal mode analysis, where it is assumed that Δq_i are sufficiently small, only quadratic terms are kept in the expansion (19) and all higher terms are neglected,[22]

$$V_{\text{vib}} \approx V_{\text{harm}} = \frac{1}{2} \sum_{i,j=1}^{3N} \left(\frac{\partial^2 V_{\text{vib}}}{\partial \Delta q_i \, \partial \Delta q_j} \right)_0 \Delta q_i \Delta q_j$$

$$= \frac{1}{2} \sum_{i,j=1}^{3N} \left(\frac{\partial^2 V_{\text{harm}}}{\partial \Delta q_i \, \partial \Delta q_j} \right)_0 \Delta q_i \Delta q_j = \frac{1}{2} \sum_{i,j=1}^{3N} H_{ij} \Delta q_i \Delta q_j$$

$$= \frac{1}{2} \Delta \mathbf{q} \cdot \mathbf{H} \cdot \Delta \mathbf{q}. \tag{20}$$

Here

$$H_{ij} = H_{ji} = \left(\frac{\partial^2 V_{\text{vib}}}{\partial \Delta q_i \, \partial \Delta q_j} \right)_0 = \left(\frac{\partial^2 V_{\text{harm}}}{\partial \Delta q_i \, \partial \Delta q_j} \right)_0 \tag{21}$$

are the elements of the Hessian $\mathbf{H} \in \mathbb{R}^{3N \times 3N}$, symmetric matrix of the second derivatives of the vibrational potential energy.

To determine the vibrational motions of the system, the eigenvalues and eigenvectors of the mass-weighted Hessian $\mathbf{M}^{-1/2} \cdot \mathbf{H} \cdot \mathbf{M}^{-1/2}$ have to be calculated.[22,43–45] This leads to solving a secular equation

$$\det(\mathbf{M}^{-1/2} \cdot \mathbf{H} \cdot \mathbf{M}^{-1/2} - \lambda \mathbf{I}) = 0, \tag{22}$$

where $\mathbf{M} \in \mathbb{R}^{3N \times 3N}$ is a diagonal mass matrix. The diagonal elements are $M_{11} = m_1$, $M_{22} = m_1$, $M_{33} = m_1, \dots, M_{3N-2,3N-2} = m_N$, $M_{3N-1,3N-1} = m_N$, $M_{3N,3N} = m_N$. For a nonlinear molecule composed of N atoms, $3N-6$ nonzero eigenvalues, $\omega_k = \sqrt{\lambda_k}$, provide the normal, or fundamental, frequencies of vibration and their associated eigenvectors, normal modes, give the directions and relative amplitudes of the atomic displacements in each mode. Six of $3N$ roots in the Eq. (22) are zero. They correspond to three translations and three rotations of a molecule as a whole.[22]

Once the normal modes and normal frequencies are determined the normal coordinates Q_k, $k = 1, 2, \dots, 3N$ can be introduced, which are in terms of the relative Cartesian displacement coordinates defined by a linear orthogonal transformation and represent independent degrees of freedom. The columns of the transformational matrix $\mathbf{A}$ between the normal and relative Cartesian displacement coordinates,[42]

$$Q_k = \sum_{i=1}^{3N} \sqrt{M_{ii}} A_{ik} \Delta q_i, \quad k = 1, 2, \dots, 3N, \tag{23}$$

are the eigenvectors of $\mathbf{M}^{-1/2} \cdot \mathbf{H} \cdot \mathbf{M}^{-1/2}$.

The equations of motion for the normal coordinates are

$$\ddot{Q}_k + \lambda_k Q_k = \ddot{Q}_k + \omega_k^2 Q_k = 0, \quad k = 1, 2, \ldots, 3N \quad (24)$$

and hold only for positive ω_k.[22]

$3N-6$ normal coordinates, which correspond to $\omega_k > 0$, describe molecular vibrations. The remaining six normal coordinates describing three translations and rotations of a molecule are introduced in such a way that the conditions required to define the Eckart frame can be written as[22]

$$Q_k = 0, \quad \dot{Q}_k = 0, \quad k = 3N - 5, \ldots, 3N. \quad (25)$$

The Eckart frame is therefore the internal coordinate system, which determines such equilibrium positions of atoms that their displacements are zero in the normal modes with $\omega_k = 0$ that describe rotational and translational motions of a molecule.

The momentum P_k conjugate to the normal coordinate Q_k is defined as

$$P_k = \frac{\partial T}{\partial \dot{Q}_k} = \dot{Q}_k + f(m_1, \ldots, m_N, Q_1, \ldots, Q_{3N-6}, \boldsymbol{\Omega}), \quad (26)$$

where $f(m_1, \ldots, m_N, Q_1, \ldots, Q_{3N-6}, \boldsymbol{\Omega}) \neq 0$ is a function of $\boldsymbol{\Omega}, m_\alpha, \alpha = 1, \ldots, N$, and $Q_k, k = 1, \ldots, 3N-6$ and stems from the coupling between rotational and vibrational degrees of freedom.[22]

In accordance with Eq. (12)

$$2T_{\text{vib}} = \sum_{\alpha=1}^{N} m_\alpha \Delta v_\alpha^2 = \sum_{k=1}^{3N} \dot{Q}_k^2 = \sum_{k=1}^{3N-6} \dot{Q}_k^2. \quad (27)$$

In the standard theory of molecular vibrations the translational and rotational degrees of freedom are described by the equations of motion for the center of mass $\dot{\mathbf{R}} = \Sigma_\alpha m_\alpha \mathbf{v}_\alpha / \Sigma_\alpha m_\alpha$ and orientation of a molecule, Eq. (10). As it turns out this kind of approach is very difficult to implement in an integration method for MD simulation[27,28] due to the Coriolis coupling between rotation and vibration.

Alternatively, we use Eq. (13) for the velocity transformation. The total kinetic energy of a molecule can then be in terms of the normal coordinates written as

$$T = \frac{1}{2} \sum_\alpha m_\alpha \mathbf{v}_\alpha^2 = \frac{1}{2} \sum_\alpha m_\alpha \Delta \mathbf{v}_\alpha^2 = \frac{1}{2} \sum_{k=1}^{3N} \dot{Q}_k^2. \quad (28)$$

In this sum, as opposed to Eqs. (11) and (27), also included are the normal coordinates with $\omega_k = 0$ for which $\dot{Q}_k$ is, in general, nonzero. In this case the momentum conjugate to the normal coordinate Q_k is

$$P_k = \frac{\partial T}{\partial \dot{Q}_k} = \dot{Q}_k. \quad (29)$$

The expression (29) differs from Eq. (26), where only the vibrational part of the velocity is expressed in terms of the normal coordinates, in that there is no coupling term between rotation and vibration.

The total kinetic energy can now be written in terms of P_k as

$$T = \frac{1}{2} \sum_{k=1}^{3N} P_k^2, \quad (30)$$

which already takes the Hamiltonian form, as opposed to Eq. (11), since it is expressed with momenta only. This is an important fact for the application of Hamilton mechanics[42] in MD simulation where the Hamiltonian of the system must be expressed as a function of generalized coordinates and their conjugate momenta, and not as a function of coordinates and velocities.[46]

Using Eq. (29), the equations of motion (24) take the Hamiltonian form as

$$\frac{d}{dt} P_k = -\omega_k^2 Q_k, \quad \frac{d}{dt} Q_k = P_k, \quad k = 1, 2, \ldots, 3N. \quad (31)$$

The general solution of this Hamilton system of equations is

$$Q_k(t) = A_k \sin(\omega_k t + \varepsilon_k), \quad (32)$$

$$P_k(t) = \omega_k A_k \cos(\omega_k t + \varepsilon_k), \quad (33)$$

where A_k and ε_k are the amplitude and the vibrational phase determined from the initial values $Q_k(0)$ and $P_k(0)$, respectively.

The particular solution of the system (31) can be written as[2]

$$\begin{bmatrix} P_k(t) \\ Q_k(t) \end{bmatrix} = \begin{bmatrix} \cos(\omega_k t) & -\omega_k \sin(\omega_k t) \\ \frac{1}{\omega_k} \sin(\omega_k t) & \cos(\omega_k t) \end{bmatrix} \begin{bmatrix} P_k(0) \\ Q_k(0) \end{bmatrix}. \quad (34)$$

Equation (34) describes vibrational motion corresponding to the normal mode k with $\omega_k > 0$.

Since here the internal coordinate system is not equal to the Eckart frame, which is not fixed at any time, the requirements $Q_k = 0$ and $\dot{Q}_k = 0$ for the normal coordinates corresponding to $\omega_k = 0$ given by Eq. (25) are no longer valid. The equations of motion for the translation and rotation of a molecule in terms of the normal coordinates, obtained from Eq. (34) for the normal coordinates with $\omega_k = 0$ and using $\lim_{x \to 0} (\sin x / x) = 1$, are

$$P_k(t) = P_k(0), \quad (35)$$

$$Q_k(t) = P_k(0)t + Q_k(0). \quad (36)$$

Since rotation is described by Eq. (36) in the same way as translation, it follows that Eq. (36) cannot be used for the description of rotation of a molecule for infinitely long times t. Equations (34) and (36) hold only for finite time intervals. In the integration method for MD simulation, in which Eqs. (34) and (36) are used for the propagation of the normal coordinates, the internal coordinate system cannot be fixed, but must be moved at least once per integration step so that Eq. (36) can be used to describe both the rotation and translation of a molecule. At the same time the internal coordinate system of a molecule must be fixed at that point in the algorithm where Eqs. (34) and (36) are applied, so that Eq. (13) holds for the transformation of the velocity.

Expressing the total velocity in terms of the normal coordinates simplifies the treatment of vibrational and rotational degrees of freedom. In Eq. (28), the problem of cou-

pling between vibration and rotation is avoided since they are decoupled, which is not the case in Eq. (11). Also, there are no system forces present in the internal coordinate system because it is fixed during the propagation of the normal coordinates by Eqs. (34) and (36). This greatly simplifies the transformations between Cartesian and relative Cartesian displacement coordinates. Thus the description of all degrees of freedom of a molecule in terms of the normal coordinates, which we have used in our MD integration approach, is more convenient for the development of a new integration method for MD simulation than the standard approach in which only vibrational degrees of freedom are treated with the normal coordinates.

III. NEW SYMPLECTIC INTEGRATORS

The new symplectic methods for MD integration are developed by implementing the derived dynamics of the internal coordinate system of a molecule into MD integration scheme.

In MD simulations for each atom of the system the Hamilton equations are solved

$$\frac{d\boldsymbol{\eta}}{dt} = \{\boldsymbol{\eta}, H\} = \hat{L}_H \boldsymbol{\eta}, \tag{37}$$

where $\hat{L}_H$ is the Lie operator, $\{,\}$ is the Poisson bracket,[42] and $\boldsymbol{\eta} = (\mathbf{q}, \mathbf{p})$ is a vector of the coordinates of all the particles and their conjugate momenta.

The formal solution of the Hamiltonian system (37) can be written in terms of Lie operators as

$$\boldsymbol{\eta}|_{t_k + \Delta t} = \exp(\Delta t \hat{L}_H) \, \boldsymbol{\eta}|_{t_k}, \tag{38}$$

and 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.[42]

A. Split integration symplectic method

The first step in the development of a new symplectic integration method is to split the Hamiltonian H of a system into two parts,[26,27]

$$H = H_0 + H_r, \tag{39}$$

where H_0 is the part of the Hamiltonian that can be solved analytically and H_r is the remaining part.

Next, a second-order approximation for Eq. (38), known as the generalized leap-frog scheme,[47,48] is used

$$\boldsymbol{\eta}|_{t_{k+1}} = \exp\left(\frac{\Delta t}{2} \hat{L}_{H_0}\right) \exp(\Delta t \hat{L}_{H_r}) \exp\left(\frac{\Delta t}{2} \hat{L}_{H_0}\right) \boldsymbol{\eta}|_{t_k}$$
$$+ O(\Delta t^3), \tag{40}$$

which defines the SISM. The whole integration time step combines the analytical evolution of H_0 with a correction from H_r resolved by numerical integration.

The model Hamiltonian has the following form:

$$H = \sum_i \frac{\mathbf{p}_i^2}{2m_i} + \frac{1}{2} \sum_{\text{bonds}} k_b(b - b_0)^2 + \frac{1}{2} \sum_{\text{angles}} k_\theta(\theta - \theta_0)^2$$
$$+ \frac{1}{2} \sum_{\text{torsions}} V_0(\cos\phi - \cos\phi_0)^2 + \sum_{i>j} \frac{e_i e_j}{4\pi\epsilon_0 r_{ij}}$$
$$+ \sum_{i>j} 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right], \tag{41}$$

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 bond lengths and angles, respectively, k_b and k_θ are corresponding force constants, ϕ_0 are the reference values for the torsion angles, and V_0 are the corresponding barrier heights, 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.

The Hamiltonian (41) is a typical MD Hamiltonian that describes a system of molecules with only one equilibrium configuration and no internal rotation. We assume that the height of the barrier of the torsional potential is large enough that the motion of atoms in the vicinity of the minimum of the torsional potential can be treated as a harmonic vibration around the equilibrium configuration. The vibrational potential energy defined by Eq. (19) for an individual molecule, is therefore the sum of vibrational potential energies of all the molecules in the system,

$$V_{\text{vib}} = \sum_{j'=1}^{m} V_{\text{vib}_{j'}} = \frac{1}{2} \sum_{\text{bonds}} k_b(b - b_0)^2 + \frac{1}{2} \sum_{\text{angles}} k_\theta(\theta - \theta_0)^2$$
$$+ \frac{1}{2} \sum_{\text{torsions}} V_0(\cos\phi - \cos\phi_0)^2, \tag{42}$$

where $V_{\text{vib}_{j'}}$ is the vibrational potential energy of the j'th molecule.

The pure harmonic Hamiltonian H_0 in the splitting (39) is defined as the sum of vibrational energies of all the molecules in the system,

$$H_0 = T + V_{\text{harm}} = \sum_{j'=1}^{m} (T_{j'} + V_{\text{harm}_{j'}}), \tag{43}$$

where $T = \sum_i \mathbf{p}_i^2 / 2m_i$ is the kinetic energy of all the atoms in the system, $T_{j'}$ is the kinetic energy of the j'th molecule, V_{harm} is the harmonic vibrational potential energy, which is for an individual molecule defined by Eq. (20), $V_{\text{harm}_{j'}}$ is the corresponding harmonic vibrational potential energy of the j'th molecule, and m is the number of all the molecules in the system.

The remaining part of the Hamiltonian,

$$H_r = H - H_0 = V_{\text{nb}} + V_{\text{ah}}, \tag{44}$$

is then equal to the sum of the nonbonded potential energy

$$V_{\mathrm{nb}} = \sum_{i>j} \frac{e_i e_j}{4\pi\epsilon_0 r_{ij}} + \sum_{i>j} 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^6\right] \quad (45)$$

and the anharmonic vibrational potential energy of higher terms (cubic, quartic, etc.) in terms of displacements of atoms from their equilibrium positions

$$V_{\mathrm{ah}} = V_{\mathrm{vib}} - V_{\mathrm{harm}}. \quad (46)$$

The underlying principle to enable the SISM to permit longer integration time steps lies in the analytical treatment of high-frequency vibrations described by H_0. The propagation scheme (40) enables us to treat the time evolution of the vibrational, rotational, and translational degrees of freedom of each molecule [described by $\exp((\Delta t/2)\hat{L}_{H_0})$] independently of all other molecules in the system because the total intermolecular interactions are described by a seperate term $\exp(\Delta t \hat{L}_{H_r})$. Each molecule is treated as an isolated molecule when propagating by $\exp((\Delta t/2)\hat{L}_{H_0})$. Propagation by $\exp((\Delta t/2)\hat{L}_{H_0})$ can therefore be solved analytically using the theory described in previous section. Since the normal mode analysis has to be carried out only once at the outset of the simulation for an isolated molecule only, a translating and rotating internal coordinate system of a molecule must be defined. For the introduced internal system of a molecule the model described in previous section was incorporated into the SISM.

The description of time evolution of the rotational and translational degrees of freedom of a molecule in terms of the normal coordinates with vibrational frequency zero ($\omega = 0$) is closely related to the introduction of the moving and rotating internal coordinate system of a molecule. Therefore the total linear momentum of every atom in a molecule is expressed in terms of the normal coordinates, so that Eqs. (34)–(36) hold for every molecule in the system.

The transformations between Cartesian, relative Cartesian, and normal coordinates of a molecule in the system used here are as follows. Let $\mathbf{e}_i$, $i=1, 2, 3$, $\mathbf{e}_j \cdot \mathbf{e}_k = \delta_{jk}$ be the unit vectors of the Cartesian coordinate system. The transformation from the Cartesian coordinates $X_\alpha, Y_\alpha, Z_\alpha$, where $\mathbf{r}_\alpha = X_\alpha \mathbf{e}_1 + Y_\alpha \mathbf{e}_2 + Z_\alpha \mathbf{e}_3$ and $\mathbf{R} = X\mathbf{e}_1 + Y\mathbf{e}_2 + Z\mathbf{e}_3$, to relative Cartesian displacement coordinates $\Delta x_\alpha, \Delta y_\alpha, \Delta z_\alpha$ is given by Eq. (5), where the displacements of the atoms from their equilibrium positions $\boldsymbol{\rho}_\alpha$ and the unit vectors of the internal coordinate system $\mathbf{f}_i$, $i=1, 2, 3$ are expressed in terms of the Cartesian coordinates

$$\mathbf{f}_i = f_{i_X}\mathbf{e}_1 + f_{i_Y}\mathbf{e}_2 + f_{i_Z}\mathbf{e}_3, \quad i=1,2,3, \quad (47)$$

$$\boldsymbol{\rho}_\alpha = \left(X_\alpha - X - \sum_{i=1}^3 c_i^\alpha f_{i_X}\right)\mathbf{e}_1 + \left(Y_\alpha - Y - \sum_{i=1}^3 c_i^\alpha f_{i_Y}\right)\mathbf{e}_2$$
$$+ \left(Z_\alpha - Z - \sum_{i=1}^3 c_i^\alpha f_{i_Z}\right)\mathbf{e}_3. \quad (48)$$

The subscripts X, Y, Z denote x, y, z components of Cartesian coordinate system, respectively. By virtue of Eqs. (47) and (48) the relative Cartesian displacement coordinates are expressed as

$$\Delta x_\alpha = \boldsymbol{\rho}_\alpha \cdot \mathbf{f}_1 = \left(X_\alpha - X - \sum_{i=1}^3 c_i^\alpha f_{i_X}\right)f_{1_X}$$
$$+ \left(Y_\alpha - Y - \sum_{i=1}^3 c_i^\alpha f_{i_Y}\right)f_{1_Y}$$
$$+ \left(Z_\alpha - Z - \sum_{i=1}^3 c_i^\alpha f_{i_Z}\right)f_{1_Z},$$

$$\Delta y_\alpha = \boldsymbol{\rho}_\alpha \cdot \mathbf{f}_2 = \left(X_\alpha - X - \sum_{i=1}^3 c_i^\alpha f_{i_X}\right)f_{2_X}$$
$$+ \left(Y_\alpha - Y - \sum_{i=1}^3 c_i^\alpha f_{i_Y}\right)f_{2_Y}$$
$$+ \left(Z_\alpha - Z - \sum_{i=1}^3 c_i^\alpha f_{i_Z}\right)f_{2_Z}, \quad (49)$$

$$\Delta z_\alpha = \boldsymbol{\rho}_\alpha \cdot \mathbf{f}_3 = \left(X_\alpha - X - \sum_{i=1}^3 c_i^\alpha f_{i_X}\right)f_{3_X}$$
$$+ \left(Y_\alpha - Y - \sum_{i=1}^3 c_i^\alpha f_{i_Y}\right)f_{3_Y}$$
$$+ \left(Z_\alpha - Z - \sum_{i=1}^3 c_i^\alpha f_{i_Z}\right)f_{3_Z}.$$

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

$$\mathbf{r}_\alpha = \mathbf{d}_\alpha + \boldsymbol{\rho}_\alpha, \quad (50)$$

in which

$$\boldsymbol{\rho}_\alpha = \Delta x_\alpha \mathbf{f}_1 + \Delta y_\alpha \mathbf{f}_2 + \Delta z_\alpha \mathbf{f}_3 = (\Delta x_\alpha f_{1_X} + \Delta y_\alpha f_{2_X}$$
$$+ \Delta z_\alpha f_{3_X})\mathbf{e}_1 + (\Delta x_\alpha f_{1_Y} + \Delta y_\alpha f_{2_Y} + \Delta z_\alpha f_{3_Y})\mathbf{e}_2$$
$$+ (\Delta x_\alpha f_{1_Z} + \Delta y_\alpha f_{2_Z} + \Delta z_\alpha f_{3_Z})\mathbf{e}_3. \quad (51)$$

For each separate component, Eq. (50) can be written as

$$X_\alpha = X + \sum_{i=1}^3 c_i^\alpha f_{i_X} + \Delta x_\alpha f_{1_X} + \Delta y_\alpha f_{2_X} + \Delta z_\alpha f_{3_X},$$

$$Y_\alpha = Y + \sum_{i=1}^3 c_i^\alpha f_{i_Y} + \Delta x_\alpha f_{1_Y} + \Delta y_\alpha f_{2_Y} + \Delta z_\alpha f_{3_Y}, \quad (52)$$

$$Z_\alpha = Z + \sum_{i=1}^3 c_i^\alpha f_{i_Z} + \Delta x_\alpha f_{1_Z} + \Delta y_\alpha f_{2_Z} + \Delta z_\alpha f_{3_Z}.$$

Using Eq. (13) where $\mathbf{f}_i$, $i=1, 2, 3$, and $\mathbf{v}_\alpha$ are expressed in terms of their components in the Cartesian coordinate system as in the case of transformation of coordinates, the transformation of momenta takes the following form:

$$\Delta p_{\alpha_x} = \mathbf{p}_\alpha \cdot \mathbf{f}_1 = p_{\alpha_X} f_{1_X} + p_{\alpha_Y} f_{1_Y} + p_{\alpha_Z} f_{1_Z},$$

$$\Delta p_{\alpha_y} = \mathbf{p}_\alpha \cdot \mathbf{f}_2 = p_{\alpha_X} f_{2_X} + p_{\alpha_Y} f_{2_Y} + p_{\alpha_Z} f_{2_Z},$$

$$\Delta p_{\alpha_z} = \mathbf{p}_\alpha \cdot \mathbf{f}_3 = p_{\alpha_X} f_{3_X} + p_{\alpha_Y} f_{3_Y} + p_{\alpha_Z} f_{3_Z}, \tag{53}$$

where $\mathbf{p}_\alpha = m_\alpha \mathbf{v}_\alpha$ is the linear momentum, $\mathbf{p}_\alpha = p_{\alpha_X} \mathbf{e}_1 + p_{\alpha_Y} \mathbf{e}_2 + p_{\alpha_Z} \mathbf{e}_3$.

Back transformation of momenta is the same as in the case of the transformation of the coordinates given by Eq. (51) as

$$\begin{aligned}
\mathbf{p}_\alpha &= \Delta p_{\alpha_x} \mathbf{f}_1 + \Delta p_{\alpha_y} \mathbf{f}_2 + \Delta p_{\alpha_z} \mathbf{f}_3 = (\Delta p_{\alpha_x} f_{1_X} + \Delta p_{\alpha_y} f_{2_X} \\
&\quad + \Delta p_{\alpha_z} f_{3_X}) \mathbf{e}_1 + (\Delta p_{\alpha_x} f_{1_Y} + \Delta p_{\alpha_y} f_{2_Y} + \Delta p_{\alpha_z} f_{3_Y}) \mathbf{e}_2 \\
&\quad + (\Delta p_{\alpha_x} f_{1_Z} + \Delta p_{\alpha_y} f_{2_Z} + \Delta p_{\alpha_z} f_{3_Z}) \mathbf{e}_3.
\end{aligned} \tag{54}$$

Transformation from the relative Cartesian displacement coordinates to the normal coordinates is given by Eq. (23) with the help of Eq. (29) as

$$Q_l = \sum_{i=1}^{3N} \sqrt{M_{ii}} A_{il} \Delta q_i, \tag{55}$$

$$P_l = \sum_{i=1}^{3N} \frac{1}{\sqrt{M_{ii}}} A_{il} \Delta p_i, \tag{56}$$

where

$$\begin{aligned}
\mathbf{\Delta q} &= (\Delta q_1, \Delta q_2, \ldots, \Delta q_N) \\
&= (\Delta x_1, \Delta y_1, \Delta z_1, \ldots, \Delta x_N, \Delta y_N, \Delta z_N)
\end{aligned}$$

is a vector of relative Cartesian displacement coordinates and their corresponding momenta are

$$\begin{aligned}
\mathbf{\Delta p} &= (\Delta p_1, \Delta p_2, \ldots, \Delta p_{3N}) \\
&= (\Delta p_{1_x}, \Delta p_{1_y}, \Delta p_{1_z}, \ldots, \Delta p_{N_x}, \Delta p_{N_y}, \Delta p_{N_z}) \\
&= (m_1 \Delta v_{1_x}, m_1 \Delta v_{1_y}, m_1 \Delta v_{1_z}, \ldots, m_N \Delta v_{N_x}, m_N \Delta v_{N_y}, m_N \Delta v_{N_z})
\end{aligned}$$

The subscripts x, y, x denote x, y, z components of the internal coordinate system, respectively.

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

$$\Delta q_i = \frac{1}{\sqrt{M_{ii}}} \sum_{l=1}^{3N} A_{il} Q_l, \tag{57}$$

$$\Delta p_i = \sqrt{M_{ii}} \sum_{l=1}^{3N} A_{il} P_l, \tag{58}$$

since the transformation matrix $\mathbf{A}$ is orthogonal.

1. Dynamics of internal coordinate system

The internal coordinate system of a molecule changes with time because the rotation of a molecule changes the orientation of the internal coordinate system while the translational motion of a molecule shifts its origin. As follows from derivation of Eqs. (13)–(18), the conditions (16) and (17) require the internal coordinate system of a molecule to remain fixed during the propagation by $\exp((\Delta t/2) \hat{L}_{H_0})$. As a consequence the explicit form of the internal coordinate system is required only while transforming between normal and Cartesian coordinates before and after the propagation by

$\exp((\Delta t/2) \hat{L}_{H_0})$. We should note that the internal coordinate system is not equal to the Eckart frame, which also changes during the propagation by $\exp((\Delta t/2) \hat{L}_{H_0})$ according to Eq. (10). Therefore, equations of motion for the internal coordinate system cannot be derived simply by discretization of Eq. (10). It is crucial in the development of the new method to derive the equations of motion that describe the dynamics of the internal coordinate system of a molecule. Note that for symmetry reasons, different molecular configurations can be related to the same equilibrium configuration and consequently to the same Eckart frame. Therefore, our internal coordinate system can be derived by means of Eqs. (A1)–(A3) from some arbitrary configuration $\mathbf{w}_k$. To remove the arbitrariness of $\mathbf{w}_k$ we need an extra condition.

Because the SISM is symplectic and symmetric it is time reversible.[2] According to the identity espressing the time reversibility condition

$$\exp(-\Delta t \hat{L}_H) \exp(\Delta t \hat{L}_H) = \mathbf{I}, \tag{59}$$

where $\mathbf{I}$ is the identity matrix, all equations defining the new integration method must be time reversible. Since the action of the propagator $\exp((\Delta t/2) \hat{L}_{H_0})$ depends on the internal coordinate system the dynamics of the internal coordinate system must be also time reversible. So, if at the end of a given integration step, the direction of time is reversed, then the second propagator $\exp((\Delta t/2) \hat{L}_{H_0})$ in the current integration step and the first propagator $\exp((-\Delta t/2) \hat{L}_{H_0})$ in the subsequent integration step must satisfy the condition

$$\exp\left(-\frac{\Delta t}{2} \hat{L}_{H_0}\right) \exp\left(\frac{\Delta t}{2} \hat{L}_{H_0}\right) = \mathbf{I}. \tag{60}$$

Therefore, Eq. (60) represents a strong restriction to the choice of the dynamics of the internal coordinate system, which enables us to combine symplectic MD integration and molecular vibrational theory.

The linear momentum $\mathbf{p}_k = (p_{1_k}, p_{2_k}, \ldots, p_{3N_k}) = (\mathbf{p}_{1_k}, \mathbf{p}_{2_k}, \ldots, \mathbf{p}_{N_k})$ of a given molecule with N atoms can be split into a vibrational, rotational, and translational contribution as in Eq. (6)

$$\mathbf{p}_k = \mathbf{p}_k^{\text{vib}} + \mathbf{p}_k^{\text{rot}} + \mathbf{p}_k^{\text{trans}}, \tag{61}$$

where $\mathbf{p}_k^{\text{vib}}$ is the vibrational part, $\mathbf{p}_k^{\text{rot}}$ is the rotational part, and $\mathbf{p}_k^{\text{trans}}$ is the translational part of the linear momentum, respectively. From the dynamics governed by the normal modes corresponding to vibrational frequency zero ($\omega=0$) as described by Eqs. (35) and (36) it follows that $\mathbf{p}_k^{\text{rot}} + \mathbf{p}_k^{\text{trans}}$ is a constant of motion during the propagation by $\exp((\Delta t/2) \hat{L}_{H_0})$. Only this part of the linear momentum changes the equilibrium configuration of a molecule (the rotational part changes the orientation, the translational part moves the origin of the internal coordinate system) whereas the vibrational part of the linear momentum determines only the displacements of atoms from their equilibrium positions.

If the initial coordinates $\mathbf{q}_k = (q_{1_k}, q_{2_k}, \ldots, q_{3N_k}) = (\mathbf{r}_{1_k}, \mathbf{r}_{2_k}, \ldots, \mathbf{r}_{N_k})$ and the initial momenta $\mathbf{p}_k$ of atoms from the beginning of the kth integration step are used to determine the coordinates $\mathbf{w}_k$ as

$$\mathbf{w}_k = \mathbf{q}_k + \mathbf{M}^{-1} \cdot \mathbf{p}_k \frac{\Delta t}{2} = \mathbf{q}_k + \mathbf{M}^{-1} \cdot (\mathbf{p}_k^{\text{vib}} + \mathbf{p}_k^{\text{rot}} + \mathbf{p}_k^{\text{trans}}) \frac{\Delta t}{2},$$
(62)

then coordinates $\mathbf{w}_k$ define the same internal coordinate system of a molecule as the actual atoms' coordinates in the middle of the integration step after the propagation by $\exp((\Delta t/2)\hat{L}_{H_0})$. The formal proof is given in Appendix B.

If the internal coordinate system is changed at the beginning of the kth integration step using the prediction in Eq. (62) and the direction of time is reversed at the end of the kth integration step then the same internal coordinate system of a molecule is predicted in the beginning of the $(k+1)$th integration step, as it was used in the kth integration step because the coordinates of atoms do not change during the propagation by $\exp(\Delta t \hat{L}_{H_r})$. This means that the new internal coordinate system of a molecule is equal to the old one if the direction of time is reversed. The time reversibility condition of Eq. (60) is thus satisfied by our method. The internal coordinate system of a molecule remains fixed over the whole integration step and only changes at the beginning of each integration step.

The prediction, Eq. (62), yields the exact equilibrium configuration of a molecule halfway through the integration time step. This fact leads to two important consequences. First, the error of the method remains of the order $O(\Delta t^3)$ because there is no additional approximation involved. Second, the internal coordinate system of the molecule moves with the molecule (the equilibrium positions are related via Eckart conditions to the actual atomistic positions) providing that the displacements of the atoms are always kept small so that the anharmonic part of the vibrational potential energy is always much smaller than the harmonic one and our normal mode approach is justified.

2. Algorithm

The algorithm employed for the SISM defined by Eq. (40), for each molecule in the system, consists of nine steps given below.

(0) Initialization: At the outset of a simulation the vibrational frequencies and normal mode vectors of the mass-weighted Hessian $\mathbf{M}^{-1/2} \cdot \mathbf{H} \cdot \mathbf{M}^{-1/2}$, where $\mathbf{H}$ is defined by Eq. (21) for an isolated molecule only, are computed. Also the transformational matrix $\mathbf{A}$ between the relative Cartesian displacement and normal coordinates is determined. The columns of $\mathbf{A}$ are the eigenvectors of $\mathbf{M}^{-1/2} \cdot \mathbf{H} \cdot \mathbf{M}^{-1/2}$.

(1) Definition of moving and rotating internal coordinate system: The internal coordinate system of a molecule is defined by Eqs. (1) and (A1)–(A3) in which the coordinates $\mathbf{w}_k$ from the prediction (62) are inserted as the coordinates of the atoms in the molecule. The coordinates $\mathbf{w}_k$ determine such configuration of the molecule that has the same Eckart frame as the true configuration of the same molecule in the middle of the current integration step during the evolution by $\exp(\Delta t \hat{L}_{H_r})$ [in accordance with Eq. (16) the same center-of-mass vector

$\mathbf{R}$ is used for transformations between different coordinate sets until the next move of the internal coordinate system]. Then the relative Cartesian displacement coordinates $\Delta\mathbf{q}_k, \Delta\mathbf{p}_k$ are obtained from $\mathbf{q}_k, \mathbf{p}_k$ using Eqs. (49) and (53). The normal coordinates Q_{i_k}, P_{i_k}, $i = 1, \dots, 3N$, at the beginning of the current integration step are determined using Eqs. (55) and (56).

(2) Propagation by $\exp((\Delta t/2)\hat{L}_{H_0})$. Rotation of Q_{i_k}, P_{i_k}, $i = 1, \dots, 3N$, in phase space by corresponding vibrational frequency ω_i for $\Delta t/2$ to obtain Q'_{i_k}, P'_{i_k}, $i = 1, \dots, 3N$, using Eq. (34) for vibration ($\omega_i \neq 0$) and Eqs. (35) and (36) for translation and rotation ($\omega_i = 0$).

(3) Transformation of the normal coordinates Q'_{i_k}, P'_{i_k}, $i = 1, \dots, 3N$, to the relative Cartesian displacement coordinates $\Delta\mathbf{q}'_k, \Delta\mathbf{p}'_k$ using Eqs. (57) and (58).

(4) Transformation of the relative Cartesian displacement coordinates $\Delta\mathbf{q}'_k, \Delta\mathbf{p}'_k$ to the Cartesian coordinates $\mathbf{q}'_k, \mathbf{p}'_k$ using Eqs. (52) and (54).

(5) Evolution by $\exp(\Delta t \hat{L}_{H_r})$. The numerical integration of momenta (force calculation):

$$\mathbf{p}''_k = \mathbf{p}'_k - \Delta t \frac{\partial H_r}{\partial \mathbf{q}},$$

where $\partial/\partial\mathbf{q} = (\partial/\partial q_1, \partial/\partial q_2, \dots, \partial/\partial q_{3n})$ and n is the number of all atoms in the system. Since H_r is the function of coordinates, $\partial H_r/\partial\mathbf{p} = 0$, only momenta are changed in this step ($\mathbf{q}''_k, \mathbf{q}'_k$).

(6) Back transformation from the Cartesian coordinates $\mathbf{q}''_k, \mathbf{p}''_k$ to the relative Cartesian displacement coordinates $\Delta\mathbf{q}''_k, \Delta\mathbf{p}''_k$ using Eqs. (49) and (53).

(7) Back transformation from the relative Cartesian displacement coordinates $\Delta\mathbf{q}''_k, \Delta\mathbf{p}''_k$ to the normal coordinates Q''_{i_k}, P''_{i_k}, $i = 1, \dots, 3N$, using Eqs. (55) and (56).

(8) Propagation by $\exp((\Delta t/2)\hat{L}_{H_0})$. Again, rotation of the normal coordinates in phase space for $\Delta t/2$ to obtain Q'''_{i_k}, P'''_{i_k}, $i = 1, \dots, 3N$, using Eq. (34) for vibration ($\omega_i \neq 0$) and Eqs. (35) and (36) for translation and rotation ($\omega_i = 0$).

(9) Use Eqs. (57), (58), (52), and (54) to obtain $\mathbf{q}_{k+1}, \mathbf{p}_{k+1}$. Go to (1) until the desired number of integration steps is reached.

The above equations hold for every atom in a molecule and for every molecule in the system where the potential V_{nb} in Eq. (44) is the only function that depends on all the coordinates of all the atoms in the system not only on the coordinates of atoms in a single molecule. The SISM is schematically presented in Fig. 3.

Due to the introduction of the translating and rotating internal coordinate system of a molecule, the Hessian $\mathbf{H}$ defined by Eq. (21) is diagonalized only once at the beginning of a simulation, in step (0). It was assumed that molecules have only one equilibrium configuration and no internal rotation. Since the displacements of atoms from their moving equilibrium positions are always sufficiently small, Eq. (20) holds at any moment. Therefore the Hessian $\mathbf{H}$, defined by Eq. (21), is the same constant matrix for the entire simulation.

Definition of the
Internal Coordinate System
Cartesian Coordinates

Rel. Displ. Coordinates

$\Delta t/2$

Propagation with H_0
Normal Coordinates

Rel. Displ. Coordinates

Δt

Propagation with $H_r = H_r (q)$
Cartesian Coordinates

Rel. Displ. Coordinates

$\Delta t/2$

Propagation with H_0
Normal Coordinates

Rel. Displ. Coordinates

Cartesian Coordinates
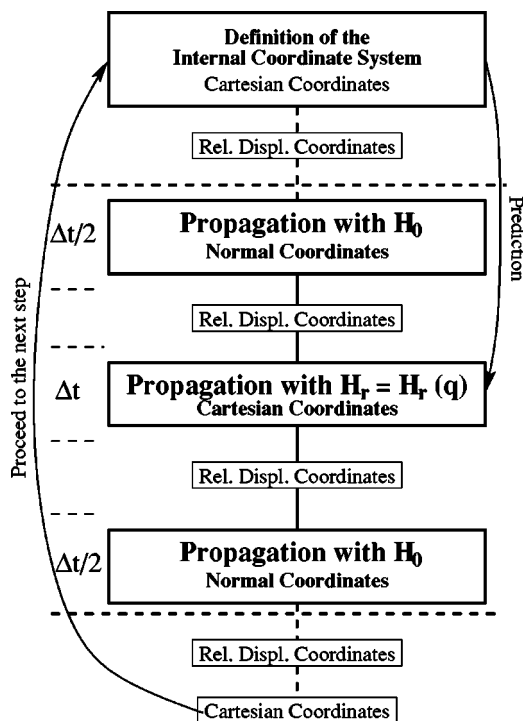
Proceed to the next step

Predicton

FIG. 3. Solution scheme for SISM.

The translational and rotational motions of a molecule are described in terms of the normal coordinates. Equations (35) and (36) are equal for both motions which means that the rotation is numerically integrated. Equations (35) and (36) are therefore valid only for some Δt that is determined by the amplitude of displacements of atoms from the equilibrium so that Eq. (20) holds.

The proposed integration method differs from all other integration schemes for MD simulation[49,50] in that the rotational and translational degrees of freedom are described in terms of the normal coordinates and the translating and rotating internal coordinate system of a molecule is introduced. In the presented method the high-frequency harmonic motions that stem from molecular vibrations are treated analytically, which enables the SISM to use longer integration time steps. The time consuming numerical calculation of non-bonded force, performed by the propagator $\exp(\Delta t \hat{L}_{H_r})$, is performed in the same way as in the standard methods, e.g., LFV.[1] The SISM's computation time per integration time step is therefore approximately the same as in the standard methods and the SISM's speed-up over the standard methods is due to longer allowed integration time steps.

The molecular vibrations can be considered as the forced vibration with the electrostatic and van der Waals interactions playing the role of the external driving force. The SISM adiabatically couples the high-frequency vibrations of atoms in a molecule to the slow degrees of freedom in the system. The propagator $\exp(\Delta t \hat{L}_{H_r})$ changes the momenta of atoms on account of electrostatic, van der Waals, and anharmonic interactions determined by V_{ah} defined by Eq. (46). The vibrational phase ε_k in every normal mode of the molecular vibration is changed accordingly. The changed momenta and displacements of atoms are transformed to the normal coor-

dinates, and because Eq. (34) is used in the propagation by $\exp((\Delta t/2)\hat{L}_{H_0})$, the information about the amplitude and the vibrational phase of atoms in each normal mode is completely conserved.[42] Since the total vibrational potential energy V_{vib} is not the quadratic function of relative Cartesian displacement coordinates the high-frequency anharmonic terms collected in V_{ah} remain in H_r. Despite that $V_{ah} \ll V_{harm}$ the term V_{ah} may not be neglected. The potential V_{ah} couples the normal modes of a molecule and thus enables the flow of the energy between them. A molecular system could not reach thermostatic equilibrium at all in the absence of the electrostatic and van der Waals interactions if the term V_{ah} is neglected. Therefore not only the harmonic but all terms in the expansion by Eq. (19) of V_{vib} are taken into account in the SISM. For this reason the SISM could be also applied, already at this stage of development, to MD simulation of molecules with more than one equilibrium configuration and/or internal rotations.[51–56] However, in this case due to the substantial anharmonic term V_{ah} the integration time step size could not be longer than the corresponding integration time step size of the LFV algorithm and no speed-up over the LFV is gained in this case.

B. Multiple time stepping SISM

The anharmonic potential V_{ah} defined by Eq. (46) is the only high-frequency term in the Hamiltonian (41) that is integrated numerically in the SISM. The largest high-frequency contribution in V_{ah} comes from the anharmonic terms of the bond-stretching potential. To overcome the difficulties caused by the troublesome V_{ah} term, we use a similar approach as the Verlet-I/r-RESPA method[12,13] except that we use different splitting of the Hamiltonian and analytically resolve the harmonic part of high-frequency motions.

First we split the Hamiltonian H defined by Eq. (41) as

$$H = H_1 + H_2, \tag{63}$$

$$H_1 = V_{nb}, \tag{64}$$

$$H_2 = H_0 + V_{ah}, \tag{65}$$

in which H_0 is defined by Eq. (43), the potential V_{nb} is defined by Eq. (45), and V_{ah} by Eq. (46).

Using Strang splitting (40), the following approximation for the propagator $\exp(\Delta t \hat{L}_H)$

$$\exp(\Delta t \hat{L}_H) = \exp\left(\frac{\Delta t}{2} L_{V_{nb}}\right)$$
$$\times \left[\exp\left(\frac{\delta t}{2}\hat{L}_{H_0}\right)\exp(\delta t \hat{L}_{V_{ah}})\exp\left(\frac{\delta t}{2}\hat{L}_{H_0}\right)\right]^n$$
$$\times \exp\left(\frac{\Delta t}{2}\hat{L}_{V_{nb}}\right) + O(\Delta t^3), \tag{66}$$

is used to derive the 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 using Eqs. (34)–(36) in the same way
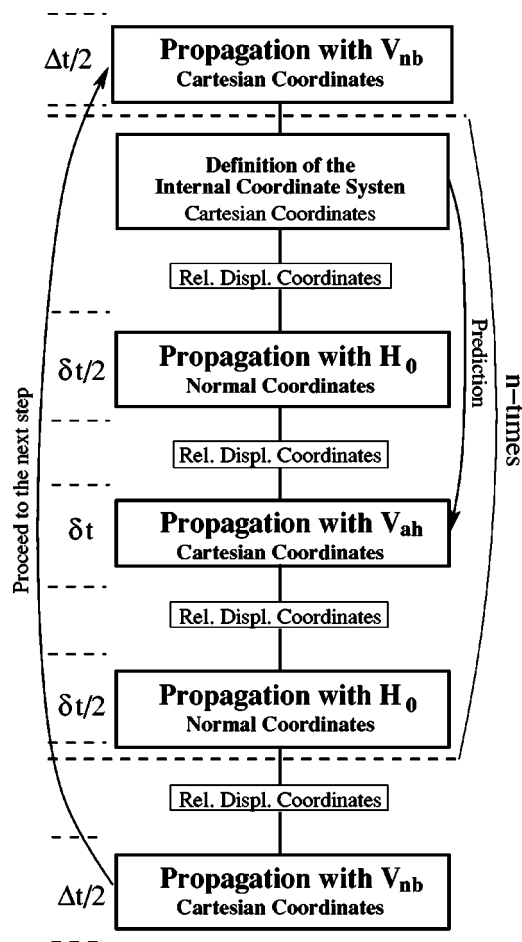
FIG. 4. Solution scheme for SISM-MTS.

as in the SISM described in the preceding section. The SISM-MTS is schematically shown in Fig. 4.

The positions $\mathbf{w_k}$, which are used to determine the new translating and rotating internal coordinate system of a molecule, are predicted using Eq. (62), in which positions and momenta after the first propagation by $\exp((\Delta t/2)\hat{L}_{V_{nb}})$ are inserted instead of the initial $\mathbf{q}_k$ and $\mathbf{p}_k$ at the beginning of the integration step. In Eq. (62) the integration time step Δt is substituted by δt. In the SISM-MTS the internal coordinate system is thus moved n times per integration step.

The difference between the SISM-MTS and Verlet-I/r-RESPA is in the different splitting of the Hamiltonian and in the analytical propagation by $\exp((\delta t/2)\hat{L}_{H_0})$ using Eqs. (34)–(36) in the SISM-MTS. In the scheme defined by Eq. (66) the nonbonded forces are calculated only once per integration step and the intramolecular linear forces are calculated n times per integration step. Since this calculation is much less time demanding than the calculation of long-range forces, $\delta t = \Delta t/n$ can be chosen as small as is required for an accurate sampling of the motion generated by the high-frequency interactions V_{ah}.

C. Equilibrium SISM

In the SISM defined by Eq. (40), the problem of the high-frequency anharmonic interactions defined by V_{ah} in Eq. (46) cannot be addressed by introducing shorter time

steps as in the SISM-MTS because the coordinates of atoms remain unchanged during the propagation by $\exp(\Delta t \hat{L}_{H_r})$. Since the total kinetic energy T is included in H_0, the propagator $\exp((\Delta t/2)\hat{L}_{H_0})$ is the only propagator that moves the coordinates in the SISM. Further splitting of $\exp(\Delta t \hat{L}_{H_r})$ into smaller time steps is senseless because the forces remain the same. At the same time the problem of resonances that are induced by the fastest normal modes of molecular vibrations[37] cannot be avoided even in the case of the SISM-MTS. However, we can apply the idea of the mollified impulse method[35–37] (MOLLY) in which the potential of the slow forces is computed at time averaged values of atom positions. In this way the components of slow forces, i.e., the electrostatic and van der Waals forces, are filtered out in the directions that excite the molecular vibrations generated by the fast forces, i.e., the forces determined by the bond-stretching and angle-bending potential that are susceptible to resonances.

The procedure in the SISM-EQ is that the potential of nonbonded forces is computed from the equilibrium positions and that the corresponding force is multiplied by the corresponding Jacobian as follows:

$$V_{nb}(\mathbf{q}) \rightarrow V_{nb}[\mathbf{d}(\mathbf{q})], \tag{67}$$

$$\mathbf{F}_{nb}(\mathbf{q}) \rightarrow \mathcal{J}^T \cdot \mathbf{F}_{nb}[\mathbf{d}(\mathbf{q})], \tag{68}$$

where V_{nb} is the sum of Coulomb and van der Waals potentials given by Eq. (45), $\mathbf{F}_{nb} = -\partial V_{nb}/\partial\mathbf{q}$ is the corresponding force, $\partial/\partial\mathbf{q} = (\partial/\partial X_1, \partial/\partial Y_1, \partial/\partial Z_1, \ldots, \partial/\partial X_n, \partial/\partial Y_n, \partial/\partial Z_n)$, $\mathbf{q} = (q_1, \cdots, q_{3n}) = (X_1, Y_1, Z_1, \ldots, X_n, Y_n, Z_n)$ are the Cartesian coordinates of all atoms in the system with n atoms, $\mathbf{d}(\mathbf{q}) \in \mathbb{R}^{3n}$ are the equilibrium positions defined for every molecule in the system by Eqs. (1)–(3) and (A1)–(A3), and $\rightarrow$ denotes substitution. The matrix $\mathcal{J} \in \mathbb{R}^{3n \times 3n}$ is the Jacobian due to the change of coordinates. It is calculated numerically by the finite difference method

$$\mathcal{J}_{ij} = \frac{d_i(q_1, q_2, \ldots, q_j + h, \ldots, q_{3n}) - d_i(q_1, q_2, \ldots, q_j - h, \ldots, q_{3n})}{2h},$$

$$\tag{69}$$

where h is 0.001. Because the equilibrium positions of atoms in a molecule depend only on the coordinates of atoms of the same molecule, the Jacobian $\mathcal{J}$ is a sparse block diagonal matrix.

For the same reasons as presented in Ref. 36 it is also valid for the SISM-EQ that the components of the external force $\mathbf{F}_{nb}$ in Eq. (68) are annihilated in the directions along which the high-frequency molecular vibrations are excited. This leads to weaker anharmonic interactions defined by V_{ah} in Eq. (46) and consequently to enhanced stability of the integrator. Another advantage of computing the potential of the nonbonded forces with the equilibrium configuration of the molecule is that the partial charges on the atoms in the molecule, which are constant during MD simulation, are determined only for the equilibrium configuration. The quantum potential is different for all the other deformed configurations of the molecule and therefore different partial charges corresponding to these configurations should be used in MD

simulation. By computing the potential of the electrostatic forces with the equilibrium positions instead of the actual ones, this problem is circumvented.

The SISM-EQ is equivalent to the SISM defined by Eq. (40) except that the modified electrostatic and Lennard-Jones potential are used. The SISM-EQ therefore conserves

$$H = T(\mathbf{p}) + V_{\text{vib}}(\mathbf{q}) + V_{\text{nb}}[\mathbf{d}(\mathbf{q})]. \tag{70}$$

The SISM-EQ is similar to equilibrium MOLLY (Ref. 36) except that there are some important differences. The first difference is in the applied integration scheme for the propagation of the coordinates and momenta of the atoms. The SISM-EQ uses the integration scheme of the SISM defined by Eq. (40) for the propagation of the coordinates and momenta of the atoms whereas the equilibrium MOLLY uses the scheme of the Verlet-I/r-RESPA method.[12,13] Another essential difference is that in the SISM-EQ different positions $\mathbf{d}(\mathbf{q})$, which define the equilibrium positions of the atoms, are used. In the equilibrium MOLLY the equilibrium positions $\mathbf{d}(\mathbf{q})$ are obtained by projecting the coordinates of the atoms $\mathbf{q}$ onto the manifold in the configuration space defined by the reference values of the bond lengths and bond angles.[36] This projection is performed by an iterative numerical procedure similar to SHAKE.[9] On the contrary, in the SISM-EQ Eqs. (1)–(3) and (A1)–(A3) are used to define the equilibrium positions of atoms. This is more convenient because the equilibrium positions of atoms determined in this way are in accordance with the standard theory of molecular vibrations and no numerical iteration is required.

D. Equilibrium multiple time stepping SISM

Combining integration scheme the SISM-MTS given by Eq. (66) and the SISM-EQ introduced in the preceding section for the propagation of the coordinates and momenta of atoms, leads to the SISM-MTS-EQ. We use the equilibrium positions $\mathbf{d}(\mathbf{q})$ defined by Eqs. (1)–(3) and (A1)–(A3) as in the SISM-EQ method. This function $\mathbf{d}(\mathbf{q})$ represents a new averaging function to compute nonbonded potential energy. It mollifies the impulse of the slow forces and thus together with the SISM-MTS leads to the SISM-MTS-EQ integration method.

IV. CONCLUSIONS

In the present work we combined the symplectic molecular dynamics integration and molecular vibrational theory to derive a new efficient SISM for MD integration and its variants.

An increase of efficiency of the newly developed integration methods in comparison with standard symplectic integration LFV method was achieved by the analytical treatment of high-frequency motions using the normal mode analysis, which is carried out once at the outset of a simulation for an isolated molecule only. The method increases the longest size of the integration time step from the size that is determined by the high-frequency bond-stretching and angle-bending interactions to the size that is determined by the van der Waals and electrostatic interactions acting among particles in the system. These nonbonded interactions are in the

SISM treated numerically within the generalized leap-frog scheme in the same way as in the LFV method.

The SISM differs from other similar integration methods in that it uses the translating and rotating internal coordinate system of each molecule in the system, which enables us to perform the normal mode analysis only once, at the beginning of the simulation. Since the SISM is symplectic and symmetric it is time reversible. From the time reversibility condition we derived the equations of motion for the dynamics of the internal coordinate system of each molecule. This dynamics represents the core of the method. With the introduction of the translating and rotating internal coordinate system we achieve that the displacements of the atoms from the instantaneous equilibrium positions are always sufficiently small and the theory of small vibrations can be applied. The coupled vibrations of atoms in molecules are decoupled using the normal coordinates, which are also used to describe the translation and rotation of molecules.

We have also derived two improvements of the SISM, which efficiently handle those high-frequency anharmonic vibrational terms in the Hamiltonian that are integrated numerically in the SISM. The SISM-MTS uses, for the numerical integration of high-frequency anharmonic vibration terms, a shorter integration time step and for the dynamics, which is generated by the electrostatic and van der Waals interactions, a correspondingly longer time step. The equilibrium SISM (SISM-EQ) uses a new averaging function, first introduced here, to mollify the impulse of the slow forces and the integration scheme of the introduced SISM to propagate in phase space. The van der Waals and electrostatic potential energies are calculated using the equilibrium positions of atoms, which are defined by the Eckart frame, instead of in the actual positions of the atoms in a molecule. In this way the components of intermolecular forces are eliminated in the directions along which the high-frequency molecular vibrations are excited, which improves the stability of the presented integration methods. Combining methods SISM-MTS and SISM-EQ leads to the method SISM-MTS-EQ with which the optimum enlargement of the integration time step was achieved.[38,39]

APPENDIX A

The Eckart internal coordinate system[31,32] is introduced as follows: first we define a fixed Cartesian coordinate system and denote the instantaneous position vectors of N atoms of a molecule of the system relative to its origin by $\mathbf{r}_\alpha$, $\alpha = 1, 2, \ldots, N$.

The Eckart frame is then defined by the orthogonal right-handed triad of unit vectors $\mathbf{f}_i$, $i = 1, 2, 3$, with the origin in the center of mass of a molecule. The unit vectors are

uniquely defined by the instantaneous positions $\mathbf{r}_\alpha$ of atoms in the molecule, by the masses of atoms m_α, and by the constant equilibrium distances of the atoms from the molecule's center of mass c_i^α, $i=1$, 2, 3, defined by the static molecular model[32] and satisfying $\sum_\alpha m_\alpha c_i^\alpha = 0$. The internal coordinate system can be attached to the equilibrium configuration of a molecule in different ways, each of which yields a different Eckart frame. Nevertheless, once we have made a definite choice, the Eckart frame is defined in a unique way.

The unit vectors $\mathbf{f}_1, \mathbf{f}_2$, and $\mathbf{f}_3$ are determined from three Eckart vectors $\boldsymbol{\mathcal{F}}_1, \boldsymbol{\mathcal{F}}_2$, and $\boldsymbol{\mathcal{F}}_3$, which are introduced as[32]

$$\boldsymbol{\mathcal{F}}_i = \sum_\alpha m_\alpha c_i^\alpha \mathbf{r}_\alpha, \tag{A1}$$

where m_α is the mass of the α atom in the molecule. In the example of nonlinear and nonplanar molecules $\mathbf{f}_i$, $i=1$, 2, 3, are computed as

$$(\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3) = (\boldsymbol{\mathcal{F}}_1, \boldsymbol{\mathcal{F}}_2, \boldsymbol{\mathcal{F}}_3) \cdot \boldsymbol{\mathcal{F}}^{-1/2}, \tag{A2}$$

where $\boldsymbol{\mathcal{F}}$ is a symmetric positive definite Gram matrix defined as

$$\boldsymbol{\mathcal{F}} = \begin{bmatrix} \boldsymbol{\mathcal{F}}_1 \cdot \boldsymbol{\mathcal{F}}_1 & \boldsymbol{\mathcal{F}}_1 \cdot \boldsymbol{\mathcal{F}}_2 & \boldsymbol{\mathcal{F}}_1 \cdot \boldsymbol{\mathcal{F}}_3 \\ \boldsymbol{\mathcal{F}}_2 \cdot \boldsymbol{\mathcal{F}}_1 & \boldsymbol{\mathcal{F}}_2 \cdot \boldsymbol{\mathcal{F}}_2 & \boldsymbol{\mathcal{F}}_2 \cdot \boldsymbol{\mathcal{F}}_3 \\ \boldsymbol{\mathcal{F}}_3 \cdot \boldsymbol{\mathcal{F}}_1 & \boldsymbol{\mathcal{F}}_3 \cdot \boldsymbol{\mathcal{F}}_2 & \boldsymbol{\mathcal{F}}_3 \cdot \boldsymbol{\mathcal{F}}_3 \end{bmatrix}. \tag{A3}$$

$\boldsymbol{\mathcal{F}}^{-1/2}$ is a positive definite matrix for which $\boldsymbol{\mathcal{F}}^{-1/2} \cdot \boldsymbol{\mathcal{F}}^{-1/2} = \boldsymbol{\mathcal{F}}^{-1}$ holds, where $\boldsymbol{\mathcal{F}}^{-1}$ is the positive definite inverse of $\boldsymbol{\mathcal{F}}$.[32]

The vectors $\mathbf{f}_i$, $i=1$, 2, 3, satisfy the Eckart conditions for the orientation of the Eckart frame[31,32]

$$\sum_\alpha m_\alpha \mathbf{c}_\alpha \times \boldsymbol{\rho}_\alpha = 0, \tag{A4}$$

which state that there is no angular momentum with respect to the internal coordinate system in the zeroth order of displacements of the atoms from their equilibrium positions.[22]

In the example of planar molecules, $c_3^\alpha = 0$ holds. Therefore,[32]

$$(\mathbf{f}_1, \mathbf{f}_2) = (\boldsymbol{\mathcal{F}}_1, \boldsymbol{\mathcal{F}}_2) \cdot \boldsymbol{\mathcal{F}}^{-1/2}, \tag{A5}$$

$$\mathbf{f}_3 = \mathbf{f}_1 \times \mathbf{f}_2, \tag{A6}$$

where $\times$ denotes the vector product of two vectors and $\boldsymbol{\mathcal{F}}$ is a symmetric 2×2 matrix defined by

$$\boldsymbol{\mathcal{F}} = \begin{bmatrix} \boldsymbol{\mathcal{F}}_1 \cdot \boldsymbol{\mathcal{F}}_1 & \boldsymbol{\mathcal{F}}_1 \cdot \boldsymbol{\mathcal{F}}_2 \\ \boldsymbol{\mathcal{F}}_2 \cdot \boldsymbol{\mathcal{F}}_1 & \boldsymbol{\mathcal{F}}_2 \cdot \boldsymbol{\mathcal{F}}_2 \end{bmatrix}. \tag{A7}$$

The example for linear molecules is described in Ref. 29.

APPENDIX B

Let the Eckart frame $\mathbf{f}_i$, $i=1$, 2, 3, be located at the center of mass of a molecule $\mathbf{R}$ defined by the position vector

$$\mathbf{w}_\alpha = \mathbf{R} + \mathbf{c}_\alpha + \boldsymbol{\rho}_\alpha, \quad \alpha = 1, \dots, N, \tag{B1}$$

where the equilibrium positions of atoms relative to the center of mass of the molecule are given by the vectors $\mathbf{c}_\alpha$ defined by Eq. (2) and the displacements of atoms from their equilibrium positions $\boldsymbol{\rho}_\alpha$ fulfill the Eckart conditions in Eq. (A4),

$$\sum_\alpha m_\alpha \mathbf{c}_\alpha \times \boldsymbol{\rho}_\alpha = 0. \tag{B2}$$

Arbitrary position vector $\mathbf{r}_\alpha$, $\alpha = 1, \dots, N$, with the same center of mass $\mathbf{R}$ defines the same Eckart frame as $\mathbf{w}_\alpha$, $\alpha = 1, \dots, N$, only if it can be written as[32]

$$\mathbf{r}_\alpha = \mathbf{R} + \mathbf{c}_\alpha + \boldsymbol{\eta}_\alpha, \tag{B3}$$

where $\boldsymbol{\eta}_\alpha$ are the displacements of atoms from the equilibrium positions in this changed configuration of the molecule $\mathbf{r}_\alpha$, $\alpha = 1, \dots, N$, which also fulfill the Eckart conditions

$$\sum_\alpha m_\alpha \mathbf{c}_\alpha \times \boldsymbol{\eta}_\alpha = 0. \tag{B4}$$

From Eqs. (B1) and (B3) it follows that the position vectors which define the same Eckart frame have the same equilibrium configuration, only the displacements are different.

Let $\mathbf{w}_\alpha$, $\alpha = 1, \dots, N$, define the predicted Eckart frame in the middle of the integration step after the propagation by $\exp((\Delta t/2)\hat{L}_{H_0})$,

$$\mathbf{w}_\alpha = \mathbf{r}_{\alpha_0} + \frac{\Delta t}{2} \mathbf{v}_{\alpha_0}, \tag{B5}$$

where $\mathbf{r}_{\alpha_0}$ and $\mathbf{v}_{\alpha_0}$ are the atom coordinates and velocities at the beginning of the integration step, respectively. The positions $\mathbf{w}_\alpha$, $\alpha = 1, \dots, N$, can then be written in the form (B1) with the displacements $\boldsymbol{\rho}_\alpha$ satisfying the Eckart conditions (B2).

The initial velocity can be split into the vibrational, rotational, and translational part, respectively,

$$\mathbf{v}_{\alpha_0} = \mathbf{v}_{\alpha_0}^{\text{vib}} + \mathbf{v}_{\alpha_0}^{\text{rot}} + \mathbf{v}_{\alpha_0}^{\text{trans}}. \tag{B6}$$

The term $\mathbf{v}_{\alpha_0}^{\text{rot}} + \mathbf{v}_{\alpha_0}^{\text{trans}}$ is the only part of the velocity that moves the position of the equilibrium configuration of the molecule (the center of mass of the molecule is moving according to $\mathbf{v}_{\alpha_0}^{\text{trans}}$, the orientation of the molecule changes according to $\mathbf{v}_{\alpha_0}^{\text{rot}}$). Since the velocity $\mathbf{v}_{\alpha_0}^{\text{rot}} + \mathbf{v}_{\alpha_0}^{\text{trans}}$ is constant during propagation by $\exp((\Delta t/2)\hat{L}_{H_0})$ and the initial coordinates and momenta of atoms at the beginning of the integration step are also used in the prediction (B5), the actual positions of atoms $\mathbf{r}_\alpha$, $\alpha = 1, \dots, N$, in the middle of the integration step after the propagation by $\exp((\Delta t/2)\hat{L}_{H_0})$ determine the same equilibrium configuration of the molecule $\mathbf{R} + \mathbf{c}_\alpha$ as $\mathbf{w}_\alpha$, $\alpha = 1, \dots, N$. The positions $\mathbf{r}_\alpha$, $\alpha = 1, \dots, N$, can therefore be expressed in the form (B3).

Next, $\boldsymbol{\eta}_\alpha$ is split as

$$\boldsymbol{\eta}_\alpha = \boldsymbol{\eta}_{\alpha_{\omega=0}} + \boldsymbol{\eta}_{\alpha_{\omega \neq 0}}, \tag{B7}$$

where $\boldsymbol{\eta}_{\alpha_{\omega=0}}$ are the displacements of atoms from the equilibrium positions $\mathbf{R} + \mathbf{c}_\alpha$ due to rotational and translational motion, and $\boldsymbol{\eta}_{\alpha_{\omega \neq 0}}$ are the displacements due to vibrational

motion. They are determined either using the normal coordinates (rotation and translation correspond to zero normal frequencies and vibrations to nonzero normal frequencies) or by the condition

$$\sum_{\alpha} m_{\alpha} \mathbf{c}_{\alpha} \times \boldsymbol{\eta}_{\alpha_{\omega \neq 0}} = 0, \tag{B8}$$

since vibration has no angular momentum because rotation and vibration are decoupled by Eq. (28).

Since $\mathbf{r}_{\alpha}$, $\alpha = 1, \ldots, N$, and $\mathbf{w}_{\alpha}$, $\alpha = 1, \ldots, N$, describe the positions of atoms with the same equilibrium positions $\mathbf{R} + \mathbf{c}_{\alpha}$,

$$\boldsymbol{\eta}_{\alpha_{\omega = 0}} = 0 \tag{B9}$$

holds in the middle of the integration step after the propagation by $\exp((\Delta t/2)\hat{L}_{H_0})$. Then it follows that

$$\sum_{\alpha} m_{\alpha} \mathbf{c}_{\alpha} \times \boldsymbol{\eta}_{\alpha} = \sum_{\alpha} m_{\alpha} \mathbf{c}_{\alpha} \times (\boldsymbol{\eta}_{\alpha_{\omega = 0}} + \boldsymbol{\eta}_{\alpha_{\omega \neq 0}})$$

$$= \sum_{\alpha} m_{\alpha} \mathbf{c}_{\alpha} \times \boldsymbol{\eta}_{\alpha_{\omega \neq 0}} = 0. \tag{B10}$$

The predicted positions $\mathbf{w}_{\alpha}$, $\alpha = 1, \ldots, N$, and the actual positions $\mathbf{r}_{\alpha}$, $\alpha = 1, \ldots, N$, in the middle of the integration step after the propagation by $\exp((\Delta t/2)\hat{L}_{H_0})$ therefore define one and the same Eckart frame.

[1] L. Verlet, Phys. Rev. **159**, 98 (1967).
[2] J. M. Sanz-Serna and M. P. Calvo, *Numerical Hamiltonian Problems* (Chapman & Hall, London, 1994).
[3] B. J. Leimkuhler, S. Reich, and R. D. Skeel, Volumes in Math. Appls. **82**, 1 (1997).
[4] J. Wisdom and M. Holman, Astron. J. **102**, 1528 (1991).
[5] J. Wisdom, M. Holman, and J. Touma, Fields Inst. Commun. **10**, 217 (1996).
[6] J. Laskar and P. Robutel, Celest. Mech. **80**, 39 (2001).
[7] L. Nadolski and J. Laskar, *Proceedings of EPAC*, 2002.
[8] T. Schlick, E. Barth, and M. Mandziuk, Annu. Rev. Biophys. Biomol. Struct. **26**, 181 (1997).
[9] J. P. Ryckaert, G. Ciccotti, and H. J. C. Berendsen, J. Comput. Phys. **23**, 327 (1977).
[10] H. C. Anderson, J. Comput. Phys. **52**, 24 (1983).
[11] H. J. C. Berendsen, 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, and R. D. Skeel (Springer, Berlin, 1999), pp. 3–36.
[12] M. E. Tuckerman, B. J. Berne, and G. J. Martyna, J. Chem. Phys. **97**, 1990 (1992).
[13] H. Grubmüller, H. Heller, A. Windemuth, and K. Schulten, Mol. Simul. **6**, 121 (1991).
[14] M. E. Tuckerman, G. J. Martyna, and B. J. Berne, J. Chem. Phys. **93**, 1287 (1990).
[15] M. E. Tuckerman and G. J. Martyna, J. Chem. Phys. **94**, 1465 (1991).
[16] M. E. Tuckerman and G. J. Martyna, J. Chem. Phys. **94**, 6811 (1991).
[17] M. E. Tuckerman and B. J. Berne, J. Chem. Phys. **95**, 8362 (1991).
[18] D. D. Humphreys, R. A. Friesner, and B. J. Berne, J. Phys. Chem. **98**, 6885 (1994).
[19] P. Procacci and B. J. Berne, J. Chem. Phys. **101**, 2421 (1994).
[20] T. C. Bishop, R. D. Skeel, and K. Schulten, J. Comput. Chem. **18**, 1785 (1997).
[21] Q. Ma, J. A. Izaguirre, and R. D. Skeel, SIAM J. Sci. Comput. (USA) **24**, 1951 (2003).
[22] E. B. Wilson, J. C. Decius, and P. C. Cross, *Molecular Vibrations* (McGraw-Hill, New York, 1955).
[23] N. Matubayasi and M. Nakahara, J. Chem. Phys. **110**, 3291 (1999).
[24] T. F. Miller III, M. Eleftherion, P. Pattnaik, A. Ndirango, D. Newns, and G. J. Martyna, J. Chem. Phys. **116**, 8649 (2002).
[25] M. Ikegutchi, J. Comput. Chem. **25**, 529 (2004).
[26] D. Janežič and F. Merzel, J. Chem. Inf. Comput. Sci. **35**, 321 (1995).
[27] D. Janežič and F. Merzel, J. Chem. Inf. Comput. Sci. **37**, 1048 (1997).
[28] D. Janežič and F. Merzel, 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, and R. D. Skeel (Springer, Berlin, 1999), pp. 332–348.
[29] D. Janežič and M. Praprotnik, Int. J. Quantum Chem. **84**, 2 (2001).
[30] D. Janežič and M. Praprotnik, J. Chem. Inf. Comput. Sci. **43**, 1922 (2003).
[31] C. Eckart, Phys. Rev. **47**, 552 (1935).
[32] J. D. Louck and H. W. Galbraith, Rev. Mod. Phys. **48**, 69 (1976).
[33] M. Diraison, J.-C. Leicknam, G. Tarjus, and S. Bratos, Phys. Rev. E **50**, 2689 (1994).
[34] M. Diraison, J.-C. Leicknam, G. Tarjus, and S. Bratos, Chem. Phys. Lett. **258**, 348 (1996).
[35] B. Garcia-Archilla, J. M. Sanz-Serna, and R. D. Skeel, SIAM J. Sci. Comput. (USA) **20**, 930 (1998).
[36] J. A. Izaguirre, S. Reich, and R. D. Skeel, J. Chem. Phys. **110**, 9853 (1999).
[37] 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, and R. D. Skeel (Springer, Berlin, 1999), pp. 318–331.
[38] M. Praprotnik and D. Janežič, J. Chem. Phys. **122**, 174102 (2005), following paper.
[39] M. Praprotnik and D. Janežič, J. Chem. Phys. **122**, 174103 (2005), this issue.
[40] Y. M. Rhee and M. S. Kim, J. Chem. Phys. **107**, 1394 (1997).
[41] S. T. Park, J. H. Moon, and M. S. Kim, J. Chem. Phys. **107**, 9899 (1997).
[42] H. Goldstein, *Classical Mechanics*, 2nd ed. (Addison-Wesley, New York, 1980).
[43] B. R. Brooks, D. Janežič, and M. Karplus, J. Comput. Chem. **16**, 1522 (1995).
[44] D. Janežič and B. R. Brooks, J. Comput. Chem. **16**, 1543 (1995).
[45] D. Janežič, R. M. Venable, and B. R. Brooks, J. Comput. Chem. **16**, 1554 (1995).
[46] W. F. van Gunsteren and A. E. Mark, J. Chem. Phys. **108**, 6109 (1998).
[47] H. F. Trotter, Proc. Am. Math. Soc. **10**, 545 (1959).
[48] G. Strang, SIAM (Soc. Ind. Appl. Math.) J. Numer. Anal. **5**, 506 (1968).
[49] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Clarendon, Oxford, 1987).
[50] A. R. Leach, *Molecular Modelling*, 2nd ed. (Pearson, Harlow, 2001).
[51] A. Sayvetz, J. Chem. Phys. **7**, 383 (1939).
[52] J. B. Howard, J. Chem. Phys. **5**, 442 (1937).
[53] J. B. Howard, J. Chem. Phys. **5**, 451 (1937).
[54] B. Kirtman, J. Chem. Phys. **37**, 2516 (1962).
[55] B. Kirtman, J. Chem. Phys. **41**, 775 (1964).
[56] B. Kirtman, J. Chem. Phys. **49**, 2257 (1968).