JCTC Journal of Chemical Theory and Computation

ARTICLE

pubs.acs.org/JCTC

Overcoming the Barrier on Time Step Size in Multiscale Molecular Dynamics Simulation of Molecular Liquids

Igor P. Omelyan*,†,‡,§ and Andriy Kovalenko*,‡,†

†Department of Mechanical Engineering, University of Alberta, Mechanical Engineering Bldg. 4-9, Edmonton, AB, T6G 2G8, Canada

‡National Institute for Nanotechnology, 11421 Saskatchewan Dr., Edmonton, Alberta, T6G 2M9, Canada

§Institute for Condensed Matter Physics, National Academy of Sciences of Ukraine, 1 Svientsitskii Street, UA-79011, Lviv, Ukraine

ABSTRACT: We propose and validate a new multiscale technique, the extrapolative isokinetic Nóse−Hoover chain orientational (EINO) motion multiple time step algorithm for rigid interaction site models of molecular liquids. It nontrivially combines the multiple time step decomposition operator method with a specific extrapolation of intermolecular interactions, complemented by an extended isokinetic Nosé−Hoover chain approach in the presence of translational and orientational degrees of freedom. The EINO algorithm obviates the limitations on time step size in molecular dynamics simulations. While the best existing multistep algorithms can advance from a 5 fs single step to a maximum 100 fs outer step, we show on the basis of molecular dynamics simulations of the TIP4P water that our EINO technique overcomes this barrier. Specifically, we have achieved giant time steps on the order of 500 fs up to 5 ps, which now become available in the study of equilibrium and conformational properties of molecular liquids without a loss of stability and accuracy.

1. INTRODUCTION

The method of molecular dynamics (MD) is one of the most powerful tools for the investigation of various properties in liquids.[1−4] This especially concerns such systems as water and its solutions as well as more complex biophysical fluids, including solvated proteins, which are of interest for modern chemistry and medicine. A characteristic feature of these systems is the coexistence of dynamical processes with vastly different time scales, extending from femtoseconds up to the millisecond region.[5−8] For instance, in water, the fastest motion relates to the intramolecular vibrations of atoms, an intermediate scale arises from the strong short-range intermolecular potentials, while slow dynamics appears due to the weak long-range van der Waals and Coulombic interactions.

In MD simulations, however, the size of time steps is restricted to rather small values in order to avoid numerical instabilities and achieve a desired accuracy when integrating the equations of motion. This imposes severe limitations on the efficiency of MD calculations. It is obvious that longer time steps are more preferable because then the number of discretization points decreases, reducing the computational costs. Many multiple time stepping (MTS) techniques have been devised over the years to enlarge the time step size during the MD integration. They include the generalized Verlet integrator,[9] reversible reference system propagator algorithm (RESPA),[10−12] normal mode theories,[13−16] mollified impulse schemes,[17−19] Langevin dynamics approaches,[20−22] and canonical Nosé−Hoover-like[23−26] and isokinetic[27] thermostatting versions of RESPA, as well as more recent developments.[28,29]

The MTS concept implies that faster components of motion are integrated with inner (smaller) time steps with respect to an outer (larger) one which is employed to handle slow dynamics. This leads to a speedup of the calculations since the costly long-range interactions can then be sampled less frequently than the cheap short-range forces. However, the size of the outer time step in the microcanonical MTS integrators[9−12] (e.g., RESPA) cannot be taken as being very large by simply increasing the number of inner loops, even though the distant interactions are sufficiently small. A rapid energy growth occurs when the time interval between the weak force updates exceeds half of the period related to the fastest motion.[30−32] Such growth is well-known as resonance instabilities.[33−35]

Within the normal mode[13−16] and Langevin-type algorithms,[19−22] the resonance effects are suppressed by adding friction and random forces. The mollified impulse schemes[17−20] diminish the multistep artifacts by modifying the potential energy. In the canonical ensemble, the appearance of the resonance phenomena can be postponed to larger steps by exploiting extra phase-space variables related to a Nosé−Hoover chain thermostat.[23−26] Alternatively, the non-Hamiltonian equations of motion which are free from the resonance instabilities can be obtained with the help of the isokinetic ensemble.[27] Relatively recently, it has been shown[28,29] that a very efficient elimination of the resonant modes is achieved by conjugating the canonical chain method[23] with the isokinetic dynamics.[27] The resulting impulsive isokinetic Nosé−Hoover chain RESPA (INR) algorithm was applied to MD simulations of water to prove that time steps on the order of 100 fs are possible.

Despite this, the INR integrator was designed for fluids with only translational degrees of freedom. Usually, the orientational degrees are implicitly parametrized by atomic Cartesian coordinates subject to intramolecular constraints.[36,37] However, to satisfy them within the isokinetic integrators, the cumbersome Shake- or Rattle-like iterative procedures must be involved.[27] The necessity to extend the MTS consideration to rotational motion is motivated by the fact that the standard force fields treat

Received: March 4, 2011
Published: November 16, 2011

water and other solvent molecules as rigid bodies.[38] Moreover, in a protein, it is useful to model hydrogen-containing groups as rigid moieties since the corresponding links have the highest frequencies in the molecule. The lower frequency bonds can be interpreted as flexible. Clearly, longer time steps can then be employed. Note that the rigid-body approximation yields enough accurate results without affecting the physically important distribution functions.[39]

The existing rotational motion MD algorithms in the microcanonical,[40−47] canonical,[48−50] isokinetic,[51,52] and Langevin[53] ensembles are applicable solely for simple single time step (STS) dynamics, and they cannot handle much more complicated MTS integration. Surprisingly, up to now, only one paper[26] dealt, in fact, with the construction of MTS schemes for the propagation of orientational variables. However, it has been devoted to a canonical scheme with no emphasis on overcoming the resonance instabilities. No rigid-body MTS algorithm was derived within the isokinetic approach, which appears to be more efficient[27−29] than the canonical method. The MTS dynamics in the presence of orientational degrees is more complex and requires a special study.

In this paper, we develop an idea of combining the isokinetic dynamics with the Nosé−Hoover-like thermostatting by writing down the non-Hamiltonian equations for both translational and rotational motions. They are then explicitly integrated using the multiple time step decomposition operator method complemented by special force and torque extrapolations (section 2). This results in a completely new MTS algorithm which cannot be reduced to the impulsive INR integrator even in the absence of orientational degrees of freedom. As is demonstrated for a rigid model of water, the new approach significantly increases the maximum acceptable size of the outer step and pushes it up to several picoseconds (section 3). Concluding remarks are also provided (section 4).

2. THEORY

2.1. Interaction Site Models and Basic Equations. Let us consider a collection with N rigid molecules, each composed of M interacting sites. The usual (microcanonical) equations of translational and rotational motion for such as a system can be cast in the form $d\mathbf{\Gamma}/dt = L\mathbf{\Gamma}(t)$, where[46]

$$L = \sum_{i=1}^{N}\left[\mathbf{V}_i\frac{\partial}{\partial\mathbf{R}_i} + \mathbf{W}(\mathbf{\Omega}_i)\mathbf{S}_i : \frac{\partial}{\partial\mathbf{S}_i} + \frac{\mathbf{F}_i}{\mu}\times\frac{\partial}{\partial\mathbf{V}_i} + \mathbf{J}^{-1}(\mathbf{G}_i + (\mathbf{J}\mathbf{\Omega}_i)\times\mathbf{\Omega}_i)\frac{\partial}{\partial\mathbf{\Omega}_i}\right] \quad (1)$$

is the Liouville operator and $\mathbf{\Gamma} = \{\mathbf{R},\mathbf{V},\mathbf{S},\mathbf{\Omega}\}$ denotes the set of all phase variables. Here, $\mathbf{R}_i$ and $\mathbf{V}_i$ are the translational position and velocity, respectively, of the center of mass $\mu = \sum_{a=1}^{M} m_a$ of the ith molecule, while $\mathbf{S}_i$ and $\mathbf{\Omega}_i$ are correspondingly its attitude matrix and principal angular velocity. The matrix $\mathbf{W}(\mathbf{\Omega})$ is skewsymmetric and linear in $\mathbf{\Omega}$ so that, for instance, $\mathbf{W}(\mathbf{\Omega})\mathbf{J}\mathbf{\Omega} = (\mathbf{J}\mathbf{\Omega})\times\mathbf{\Omega}$ with $\mathbf{J}$ being the molecular matrix of moments of inertia.

The total force exerted on molecule i due to the site−site interactions φ_{ab} can be expressed in terms of the atomic counterparts $\mathbf{F}_{ia} = -\sum_{j\neq i}^{N}\sum_{b=1}^{M}\hat{\mathbf{r}}_{ij}^{ab}\varphi'_{ab}(r_{ij}^{ab})$ as $\mathbf{F}_i = \sum_{a=1}^{M}\mathbf{F}_{ia}$, where $\hat{\mathbf{r}}_{ij}^{ab} = (\mathbf{r}_{ia} - \mathbf{r}_{jb})/r_{ij}^{ab}$ with $r_{ij}^{ab} = |\mathbf{r}_{ia} - \mathbf{r}_{jb}|$ and $\varphi'(r) = d\varphi(r)/dr$. Then, the principal torque is $\mathbf{G}_i = \mathbf{S}_i\sum_{a=1}^{M}(\mathbf{r}_{ia} - \mathbf{R}_i)\times\mathbf{F}_{ia} = \sum_{a=1}^{M}\boldsymbol{\delta}_a\times(\mathbf{S}_i\mathbf{F}_{ia})$, where $\mathbf{r}_{ia} = \mathbf{R}_i + \mathbf{S}_i^{+}\boldsymbol{\delta}_a$ denotes the position of atom a within molecule i, while $\boldsymbol{\delta}_a$ is the time-independent location of site a in the body-fixed frame, so that $\mathbf{J} = \sum_{a=1}^{M}[(\boldsymbol{\delta}_a\times\boldsymbol{\delta}_a)\mathbf{I} - \boldsymbol{\delta}_a\boldsymbol{\delta}_a]m_a$. The functions φ_{ab} present the sum of the Lennard-Jones $4\varepsilon_{ab}[(\sigma_{ab}/r_{ij}^{ab})^{12} - (\sigma_{ab}/r_{ij}^{ab})^{6}]$ and Coulombic q_aq_b/r_{ij}^{ab} potentials, where q_a is the atomic charge. The values for parameters ε_{ab}, σ_{ab}, q_a, and $\boldsymbol{\delta}_a$ depend on the concrete interaction site model chosen to describe a fluid. In the case of water, the most popular is the rigid TIP4P[57] one with $M = 4$ sites.

Because in the rigid-body approximation the intramolecular degrees of freedom are frozen, we will have only two scales of time. This follows from the fact that the total intermolecular forces $\mathbf{F} = \mathbf{F}_s + \mathbf{F}_w$ and torques $\mathbf{G} = \mathbf{G}_s + \mathbf{G}_w$ consist of the strong (s) and weak (w) parts related to the short- and long-range interactions which cause the fast and slow processes, respectively.

2.2. Standard MTS Decomposition Method. In the standard MTS decomposition approach,[10,11] the Liouville operator is split as $L = A + B_s + B_w$ into one kinetic A and two potential $B_{s,w}$ terms. Taking into account eq 1, the explicit expressions for them are $A = \mathbf{V}\times\partial/\partial\mathbf{R} + \mathbf{W}(\mathbf{\Omega})\mathbf{S}\partial/\partial\mathbf{S}$ as well as $B_s = \mu^{-1}\mathbf{F}_s\times\partial/\partial\mathbf{V} + \mathbf{J}^{-1}(\mathbf{G}_s + (\mathbf{J}\mathbf{\Omega})\times\mathbf{\Omega})\times\partial/\partial\mathbf{\Omega}$ and $B_w = \mu^{-1}\mathbf{F}_w\times\partial/\partial\mathbf{V} + \mathbf{J}^{-1}\mathbf{G}_w\times\partial/\partial\mathbf{\Omega}$, where subindex i has been omitted to simplify notation and the strong-weak components of $\mathbf{F}$ and $\mathbf{G}$ have been used. Then, the time evolution propagator e^{Lh} is factorized into analytically integrable single-exponential operators as $e^{[(A+B_s+B_w)+\mathcal{O}(h^2)]h} = e^{B_w(h/2)}[e^{B_s(h/2n)}e^{A(h/n)}e^{B_s(h/2n)}]^n e^{B_w(h/2)}$, where h is the size of the outer time step, n is the number of inner loops, and $\mathcal{O}(h^2)$ is the second-order error function. For any time t, the solution can be presented in the form

$$\Gamma(t) = e^{Lt}\Gamma(0) = [e^{B_w\frac{h}{2}}[e^{B_s\frac{h}{2n}}e^{A\frac{h}{n}}e^{B_s\frac{h}{2n}}]^n e^{B_w\frac{h}{2}}]^l\Gamma(0) + \mathcal{O}(h^2) \quad (2)$$

where $\mathcal{O}(h^2) \sim l\mathcal{O}(h^2)$ denotes the global error and $l = t/h$ is the total number of steps.

From eq 2 at $n = 1$, one reproduces the well-known Verlet integrator,[54,55] while for $n \geq 2$, we come to the RESPA scheme.[10,11] In the latter case, the reference system ($L_{ref} = A + B_s$) is integrated with a time step that is h/n smaller than h related to the weak contribution B_w. This speeds up the calculations because at $n > 1$ the expensive long-range forces are recalculated not so frequently. The action of the exponentials $e^{A(h/n)}$, $e^{B_s(h/(2n))}$, and $e^{B_w(h/2)}$ on a phase space point $\mathbf{\Gamma}$ can be given analytically.[46] In particular,

$$e^{A\frac{h}{n}}\{\mathbf{R},\mathbf{S}\} = \left\{\mathbf{R} + \mathbf{V}\frac{h}{n}, \mathbf{\Theta}\left(\mathbf{\Omega},\frac{h}{n}\right)\mathbf{S}\right\} \quad (3)$$

where the changes of $\mathbf{R}$ and $\mathbf{S}$ correspond to free translational and orientational motions at fixed $\mathbf{V}$ and $\mathbf{\Omega}$. The matrix $\mathbf{\Theta}(\mathbf{\Omega},h/n) = \exp(\mathbf{W}(\mathbf{\Omega})h/n)$ denotes the three-dimensional rotation around vector $\mathbf{\Omega}$ on angle $\Omega h/n$.

As was mentioned in the Introduction, the RESPA scheme is hampered by the resonance instabilities already at relatively small values of h, even through the reference system is integrated exactly (i.e., when $n \gg 1$). We will now study the problem of how to eliminate these instabilities in the most efficient way.

2.3. Extrapolative Isokinetic Nosé−Hoover Chain Approach. *2.3.1. Non-Hamiltonian Equations of Motion.* Within the isokinetic ensemble,[27] the MTS instabilities are prevented by fixing the total kinetic energy T of the system (in our case $T = \sum_{i=1}^{N}[\sum_{\alpha}^{x,y,z}\mu V_{i,\alpha}^2/2 + \sum_{\alpha}^{X,Y,Z}J_\alpha\Omega_{i,\alpha}^2/2]$, where J_α are the diagonal elements of $\mathbf{J}$). This energy is a collective quantity which depends on velocities of all particles. Thus, further improvements in stability

can be achieved by introducing a complete set of kinetic constraints, each one concerning only a particular degree of freedom. Obviously, such an approach should provide a better suppression of the resonant modes because it allows one to control individual kinetic energies.

The main idea consists of coupling each physical degree with its own one-dimensional, one-particle imaginary subsystem. Like real bodies, such subsystems can be characterized by some masses m and moments of inertia j_α as well as by the translational $\nu_{i,\alpha}$ and angular $w_{i,\alpha}$ velocities. Then, the desired individual constraints can be built in the extended phase space by writing

$$\begin{aligned} T^{\mathrm{t}}_{i,\alpha} &= \mu V^2_{i,\alpha}/2 + m\nu^2_{i,\alpha} = k_{\mathrm{B}}\mathscr{J}/2 \\ T^{\mathrm{r}}_{i,\alpha} &= J_\alpha \Omega^2_{i,\alpha}/2 + j_\alpha w^2_{i,\alpha}/2 = k_{\mathrm{B}}\mathscr{J}/2 \end{aligned} \tag{4}$$

where k_{B} denotes the Boltzmann constant, $\mathscr{J}$ is the required temperature, and α relates either to three Cartesian (x,y,z) or principal (X,Y,Z) components for the cases of translational or angular velocities. In view of eq 4, the virtual bodies can be treated as external baths or thermostats which do not allow one to exceed the fixed level $k_{\mathrm{B}}\mathscr{J}/2$ of the kinetic energy for any real degree of freedom. The quantities $\mu V^2_{i,\alpha}/2$ and $J_\alpha \Omega^2_{i,\alpha}/2$ will fluctuate within the interval $[0,k_{\mathrm{B}}\mathscr{J}/2]$ due to the physical interactions and energy exchange between the real system and thermostats.

The nonholonomic relations (eq 4) to be satisfied require the introduction of constraint forces $-\lambda^{\mathrm{t}}_{i,\alpha}\partial T^{\mathrm{t}}_{i,\alpha}/\partial V_{i,\alpha} = -\lambda^{\mathrm{t}}_{i,\alpha}\mu V_{i,\alpha}$ and torques $-\lambda^{\mathrm{r}}_{i,\alpha}\partial T^{\mathrm{r}}_{i,\alpha}/\partial \Omega_{i,\alpha} = -\lambda^{\mathrm{r}}_{i,\alpha} J_\alpha \Omega_{i,\alpha}$ for the physical system as well as their counterparts $-\lambda^{\mathrm{t}}_{i,\alpha} m\nu_{i,\alpha}$ and $-\lambda^{\mathrm{r}}_{i,\alpha} j_\alpha w_{i,\alpha}$ for the virtual degrees. Further, each such subsystem can be in turn coupled with its own chain of $\mathscr{M}$ thermostats, described by the velocity variables $\nu_{j,i,\alpha}$ and $w_{j,i,\alpha}$, where $j = 1, 2, ..., \mathscr{M}$ with $\nu_{i,\alpha} \equiv \nu_{1,i,\alpha}$ and $w_{i,\alpha} \equiv w_{1,i,\alpha}$. Acting now in the spirit of the Nosé−Hoover (NH) chain approach[23] and taking into account the constraint forces and torques, the equations of motion for the thermostat variables can be cast in the form $d\nu_{1,\alpha}/dt = -\lambda^{\mathrm{t}}_{\alpha}\nu_{1,\alpha} - \nu_{1,\alpha}\nu_{2,\alpha}$ and $d\nu_{j,\alpha}/dt = (\nu^2_{j-1,\alpha} - 1/\tau^2_{\mathrm{t}}) - \nu_{j,\alpha}\nu_{j+1,\alpha}$ as well as $dw_{1,\alpha}/dt = -\lambda^{\mathrm{r}}_{\alpha}w_{1,\alpha} - w_{1,\alpha}w_{2,\alpha}$ and $dw_{j,\alpha}/dt = (w^2_{j-1,\alpha} - 1/\tau^2_{\mathrm{r},\alpha}) - w_{j,\alpha}w_{j+1,\alpha}$ for $j = 2, ..., \mathscr{M}$ with $\nu_{\mathscr{M}+1} = w_{\mathscr{M}+1} = 0$. Here, subscript i was again hidden for simplicity. The four relaxation times τ_{t} and $\tau_{\mathrm{r},\alpha}$ will determine the strength of coupling of the system with the translational and rotational thermostats, so that $m = \tau^2_{\mathrm{t}}k_{\mathrm{B}}\mathscr{J}/2$ and $j_\alpha = \tau^2_{\mathrm{r},\alpha}k_{\mathrm{B}}\mathscr{J}/2$. This is justified by the fact that we have only one mass μ while three $(\alpha = X,Y,Z)$ moments J_α of inertia of the molecule.

The Lagrangian multipliers $\lambda^{\mathrm{t}}_{i,\alpha}$ and $\lambda^{\mathrm{r}}_{i,\alpha}$ can be found by differentiating eq 4 with respect to time, i.e., $dT^{\mathrm{t,r}}_{i,\alpha}/dt = 0$, and using the above equations of motion for thermostat variables complemented by the equations $dV_{i,\alpha}/dt = \mu^{-1}F_{i,\alpha} - \lambda^{\mathrm{t}}_{i,\alpha}V_{i,\alpha}$ and $d\Omega_{i,\alpha}/dt = J^{-1}_{\alpha}(G_{i,\alpha} + (J_\beta - J_\gamma)\Omega_{i,\beta}\Omega_{i,\gamma}) - \lambda^{\mathrm{r}}_{i,\alpha}\Omega_{i,\alpha}$ for translational and angular velocities. This yields

$$\begin{aligned} \lambda^{\mathrm{t}}_{i,\alpha} &= \left(V_{i,\alpha}F_{i,\alpha} - \tfrac{1}{2}k_{\mathrm{B}}\mathscr{T}\tau^2_t\nu^2_{1,i,\alpha}\nu_{2,i,\alpha}\right)/(2T^{\mathrm{t}}_{i,\alpha}) \\ \lambda^{\mathrm{r}}_{i,\alpha} &= \Big(\Omega_{i,\alpha}G_{i,\alpha} + (J_\beta - J_\gamma)\Omega_{i,\alpha}\Omega_{i,\beta}\Omega_{i,\gamma} \\ &\quad - \frac{1}{2}k_{\mathrm{B}}\mathscr{T}\tau^2_{r,\alpha}w^2_{1,i,\alpha}w_{2,i,\alpha}\Big)/(2T^{\mathrm{r}}_{i,\alpha}) \end{aligned} \tag{5}$$

Note that the kinetic constraints (eq 4) provide true canonical distributions in position space (Appendix A).

The isokinetic Nosé−Hoover chain (INC) equations of translational and rotational motion we just derived can be rewritten in the compact form

$$\frac{d\Gamma_{\mathrm{inc}}}{dt} = L_{\mathrm{inc}}\Gamma_{\mathrm{inc}}(t) \tag{6}$$

with L_{inc} being the INC Liouville operator and $\mathbf{\Gamma}_{\mathrm{inc}} = \{\mathbf{R},\mathbf{V},\mathbf{S},\boldsymbol{\Omega};\boldsymbol{\nu},\mathbf{w}\}$ denoting the extended phase space, where vectors $\boldsymbol{\nu}$ and $\mathbf{w}$ represent the whole set of scalar quantities $\nu_{j,i,\alpha}$ and $w_{j,i,\alpha}$ (subscript i will further be omitted at all). Applying the strong−weak force $\mathbf{F} = \mathbf{F}_{\mathrm{s}} + \mathbf{F}_{\mathrm{w}}$ and torque $\mathbf{G} = \mathbf{G}_{\mathrm{s}} + \mathbf{G}_{\mathrm{w}}$ decompositions, one finds in view of eq 5 that $L_{\mathrm{inc}} = A + B_{\mathrm{s}} + B_{\mathrm{w}} + B_{\mathrm{inc}}$, where now

$$\begin{aligned} B_{\mathrm{s,w}} &= \sum_{\alpha}^{x,y,z} B^{\mathrm{t}}_{\mathrm{s,w},\alpha} + \sum_{\alpha}^{X,Y,Z} B^{\mathrm{r}}_{\mathrm{s,w},\alpha} \\ &= \sum_{\alpha}^{x,y,z} F_{\mathrm{s,w},\alpha}\left[\left(\frac{1}{\mu} - \frac{V^2_\alpha}{2\mathrm{T}^{\mathrm{t}}_\alpha}\right)\frac{\partial}{\partial V_\alpha} - \frac{V_\alpha \nu_{1,\alpha}}{2\mathrm{T}^{\mathrm{t}}_\alpha}\frac{\partial}{\partial \nu_{1,\alpha}}\right] \\ &\quad + \sum_{\alpha}^{X,Y,Z}(G_{\mathrm{s,w},\alpha} + \zeta_{\mathrm{s,w}}(J_\beta - J_\gamma)\Omega_\beta\Omega_\gamma)\left[\left(\frac{1}{J_\alpha} - \frac{\Omega^2_\alpha}{2\mathrm{T}^{\mathrm{r}}_\alpha}\right)\frac{\partial}{\partial \Omega_\alpha}\right. \\ &\quad \left. - \frac{\Omega_\alpha w_{1,\alpha}}{2\mathrm{T}^{\mathrm{r}}_\alpha}\frac{\partial}{\partial w_{1,\alpha}}\right] \end{aligned} \tag{7}$$

at $\zeta_{\mathrm{s}} = 1$ and $\zeta_{\mathrm{w}} = 0$, while (α,β,γ) designate the three cyclic permutations (X,Y,Z), (Y,Z,X), and (Z,X,Y). The chain thermostat contribution is

$$\begin{aligned} B_{\mathrm{inc}} &= \sum_{\alpha}^{x,y,z}\Big[B_{V,\nu,\alpha} + \sum_{j=2}^{\mathscr{M}} B_{\nu_j,\alpha}\Big] \\ &\quad + \sum_{\alpha}^{X,Y,Z}\Big[B_{\Omega,w,\alpha} + \sum_{j=2}^{\mathscr{M}} B_{w_j,\alpha}\Big] \\ &= \sum_{\alpha}^{x,y,z} B^{\mathrm{t}}_{\mathrm{inc},\alpha} + \sum_{\alpha}^{X,Y,Z} B^{\mathrm{r}}_{\mathrm{inc},\alpha} \end{aligned} \tag{8}$$

with

$$B_{V,\nu,\alpha} = \frac{1}{2}V_\alpha\nu^2_{1,\alpha}\nu_{2,\alpha}\tau^2_t\frac{\partial}{\partial V_\alpha} + \left(\frac{1}{2}\nu^3_{1,\alpha}\nu_{2,\alpha}\tau^2_{\mathrm{t}} - \nu_{1,\alpha}\nu_{2,\alpha}\right)\frac{\partial}{\partial \nu_{1,\alpha}} \tag{9}$$

and

$$B_{\nu_j,\alpha} = \left(\nu^2_{j-1,\alpha} - \frac{1}{\tau^2_{\mathrm{t}}} - \nu_{j,\alpha}\nu_{j+1,\alpha}\right)\frac{\partial}{\partial \nu_{j,\alpha}} \tag{10}$$

The expressions for $B_{\Omega,w,\alpha}$ and $B_{wj,\alpha}$ are similar to those of eqs 9 and 10 with formal replacement of V by Ω, ν by w, and τ_{t} by $\tau_{\mathrm{r},\alpha}$.

2.3.2. Extrapolative Decomposition of the Evolution Operator. Taking into account eqs 7−10, the solution to the non-Hamiltonian equations of motion (eq 6) can be obtained with the help of the decomposition method (section 2.2). As a result, for the MTS integration ($n \geq 2$), one finds

$$\Gamma_{\mathrm{inc}}(t) = \left[e^{L_{\mathrm{ws}}\frac{h}{n}}\left[e^{B_{\mathrm{inc}}\frac{h}{2n}}e^{B_{\mathrm{s}}\frac{h}{2n}}e^{A\frac{h}{n}}e^{B_{\mathrm{s}}\frac{h}{2n}}e^{B_{\mathrm{inc}}\frac{h}{2n}}\right]^{n-2}e^{L_{\mathrm{sw}}\frac{h}{n}}\right]^{l}\Gamma_{\mathrm{inc}}(0) + \mathscr{O}(h^2) \tag{11}$$

with

$$e^{L_{\mathrm{ws}}\frac{h}{n}} = e^{B_{\mathrm{inc}}\frac{h}{2n}}e^{B^{\mathrm{I}}_{\mathrm{sw}}\frac{h}{2n}}e^{A\frac{h}{n}}e^{B_{\mathrm{s}}\frac{h}{2n}}e^{B_{\mathrm{inc}}\frac{h}{2n}}, e^{L_{\mathrm{sw}}\frac{h}{n}} = e^{B_{\mathrm{inc}}\frac{h}{2n}}e^{B_{\mathrm{s}}\frac{h}{2n}}e^{A\frac{h}{n}}e^{B^{\mathrm{II}}_{\mathrm{sw}}\frac{h}{2n}}e^{B_{\mathrm{inc}}\frac{h}{2n}} \tag{12}$$

where $B^{I}_{sw} = B_s + nB^{I}_{w}$ and $B^{II}_{sw} = B_s + nB^{II}_{w}$. The time evolution propagation given by eqs 11 and 12 presents an analog of the RESPA scheme (eq 2) in the case of the INC dynamics. Like RESPA, it corresponds to an impulsive approach, where the translational and angular velocities are updated instantaneously two times per outer step h (at its beginning and its end) by $e^{B_w^{I}(h/2)}$ and $e^{B_w^{II}(h/2)}$ in the weak force $F^{I,II}_{w,\alpha}$ and torque $G^{I,II}_{w,\alpha}$ fields (B_w depends on them according to eq 7). The indexes I and II mean that the forces and torques are calculated at two spatial configurations $\{\mathbf{R}(t'),\mathbf{S}(t')\}$ and $\{\mathbf{R}(t'+h),\mathbf{S}(t'+h)\}$ corresponding to two consecutive moments of time $t' = l'h$ and $t' + h$, where $l' = 1, 2, ..., l$. The arising resonance instabilities are damped out here by the thermostat stabilizing terms $e^{B_{inc}(h/(2n))}$. It is obvious, however, that with increasing the size of the time step h, the resonance effects will grow appreciably. This can prevent their proper suppressing even at a sufficiently large strength of coupling of the system with the thermostatting baths.

A more efficient stabilization can be achieved within an extrapolative approach. Indeed, the smoothly varying long-range weak forces $F_{w,\alpha}$ and torques $G_{w,\alpha}$ we can hold constants $F^{I}_{w,\alpha}$ and $G^{I}_{w,\alpha}$ during the outer time interval $[t',t'+h]$. Further, such constants can be added to the quickly changing strong components $F_{s,\alpha}$ and $G_{s,\alpha}$ when performing the inner step propagation. Then, the time evolution (eqs 11 and 12) transforms to

$$\Gamma_{inc}(t) = [e^{B_{inc}\frac{h}{2n}}e^{B_{ws}\frac{h}{2n}}e^{A\frac{h}{n}}e^{B_{ws}\frac{h}{2n}}e^{B_{inc}\frac{h}{2n}}]^{nl}\Gamma_{inc}(0) + \mathcal{O}(h^2) \quad (13)$$

where $B_{ws} = B_s + B^{I}_{w}$. This avoids the resonance effects since the periodic impulsive propagations of velocities by $e^{B_w^{I,II}(h/2)}$ are now absent. In eq 13, the velocities are updated continuously ($2n$ times per step h) by $e^{B_w^{I}(h/(2n))}$ in the constant weak force-torque fields $F^{I}_{w,\alpha}$ and $G^{I}_{w,\alpha}$. The next constant values $F^{II}_{w,\alpha}$ and $G^{II}_{w,\alpha}$ will be used only after passing the current outer step.

Of course, the extrapolative method produces some uncertainties to the full energy since the resulting approximate forces $F_{ws,\alpha} = F_{s,\alpha} + F^{I}_{w,\alpha}$ and torques $G_{ws,\alpha} = G_{s,\alpha} + G^{I}_{w,\alpha}$ cannot be related to any conservative potentials. However, such uncertainties have a nonresonance nature, and they are suppressed by $e^{B_{inc}(h/(2n))}$ using the same INC thermostatting baths (eqs 8−10), as in the case of the impulsive method. As a result, larger values of h can be employed, despite the fact that the extrapolative algorithm (eq 13) is not reversible in time. However, this is not so important in our case because we deal with the non-Hamiltonian dynamics, where (unlike the microcanonical ensemble) the total energy should not be conserved exactly.

Note that the idea of using the force extrapolation is not new. It was exploited earlier in the context of Langevin normal mode integrators.[13−16] However, its combination with the isokinetic Nosé−Hoover chain approach complemented with the decomposition operator method is proposed here for the first time. Moreover, the introduced constant extrapolation of the long-range torque cannot be obtained by simply fixing the atomic forces, because it depends on the orientation of the molecule as well. As will be shown in section 3, the usual atomic force extrapolation does not lead to stable phase trajectories.

Therefore, the propagation $\mathbf{\Gamma}_{inc}(t)$ of phase variables from their initial values $\mathbf{\Gamma}_{inc}(0)$ to arbitrary time t in the future can be readily carried out by consecutively applying the exponential operators $e^{A(h/n)}$ (see eq 3), $e^{B_{inc}(h/(2n))}$, and $e^{B_{ws}(h/(2n))}$ in the order defined by eq 13. The analytical expressions for the action of $e^{B_{inc}(h/(2n))}$ and $e^{B_{ws}(h/(2n))}$ on $\mathbf{\Gamma}_{inc}$ can be found in Appendix B.

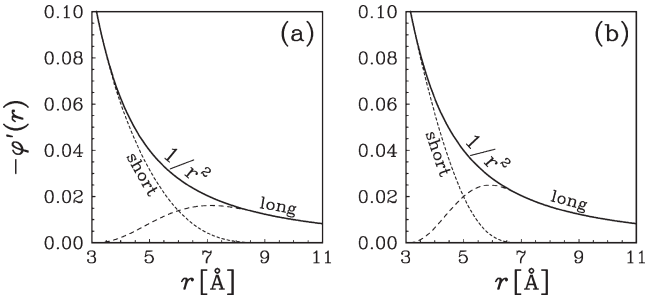

Figure 1. Schematic representation of the short- and long-range parts of the Coulombic interaction for the cutoff radii $r_c = 9$ Å [subset a] and $r_c = 7$ Å [subset b].

This completes the extrapolative isokinetic NH chain orientational (EINO) motion MTS algorithm. The impulsive version (eqs 11 and 12) will be referred to as simply INO. The latter was originally introduced in refs 63 and 64. In the absence of orientational degrees of freedom, the INO integrator reduces to INR.[28,29] This is contrary to the proposed EINO approach, which cannot be reproduced from the impulsive INR scheme even in the case of pure translational motion.

3. APPLICATION TO WATER AND DISCUSSION

3.1. Details of MD Simulations. The EINO algorithm derived in the preceding section will now be tested in MD simulations. The system considered is the rigid TIP4P model[57] of water ($M = 4$). We have involved $N = 512$ molecules placed in a cubic box of volume $V = L^3$ with periodic boundaries. The simulations were performed at a density of $N/V = 1$ g/cm^3 and a temperature of $\mathcal{T} = 293$ K. The total intermolecular forces $\mathbf{F}$ were evaluated with the help of the Ewald summation technique[58] at the cutoff radii $R_c = L/2 = 12.417$ Å and $\kappa_{max} = 8$ in the real and reciprocal spaces, respectively.

The strong component $\mathbf{F}_s$ has been determined in the spirit of the near/far distance-based approach[11,59−62] using the replacement of $\varphi'(r)$ by $\varphi'_s(r) = \phi(r)\,\varphi'(r)$ in the standard expression for $\mathbf{F}$ (section 2.1). The switching function was chosen in the form of the cubic spline[62] $\phi(r) = 1 - (10 - 15\eta + 6\eta^2)\eta^3$ with $\eta = 1 + (r - r_c)/(r_0 - r_c)$ to smoothly change its value from 1 to 0 when increasing the interatomic distance r from r_0 to $r_c < L/2$. The weak force part was then found by extracting $\mathbf{F}_s$ from $\mathbf{F}$, i.e., as $\mathbf{F}_w = \mathbf{F} - \mathbf{F}_s$. This appears to be more efficient than the straightforward real/reciprocal splitting of Coulombic interactions.[60,62] Having $\mathbf{F}_s$ and $\mathbf{F}_w$, the strong $\mathbf{G}_s$ and weak $\mathbf{G}_w$ components of the total intermolecular torques $\mathbf{G}$ were obtained applying usual relations (section 2.1) with formal replacement of $\mathbf{F}$ by $\mathbf{F}_s$ or $\mathbf{F}_w$. Two cases related to cutting-off of the short-range interaction at $r_c = 9$ Å and $r_c = 7$ Å have been considered. For illustration, the related short- and long-range parts $\varphi'_s(r)$ and $\varphi'(r) - \varphi_s'(r)$ are plotted in Figure 1 together with the total function $\varphi'(r)$, where $\varphi(r) = 1/r$ is the generic Coulomb potential. The switching-on parameter was set equal to $r_0 = 3$ Å.

The equations of motion were solved using the proposed EINO algorithm (eqs 13 and B1−B8) as well as its impulsive version INO (eqs 11 and 12). For the purpose of comparison, the best previously known MTS approaches in different statistical ensembles were applied, too. They include the energy-targeted microcanonical RESPA (ERESPA) scheme as well as the extended canonical NH (ENH) and isokinetic (ISO) integrators.

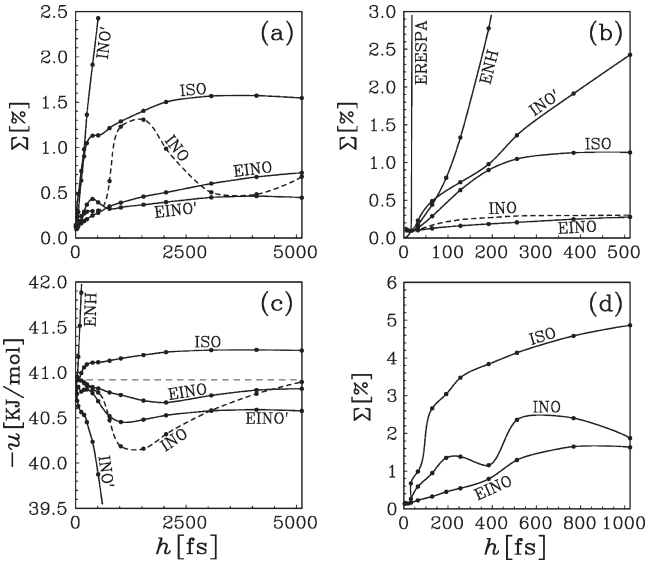

Figure 2. Uncertainty $\Sigma(h)$ in the calculation of the distribution functions by the MD simulations of the TIP4P water with different algorithms at various time steps h for r_c = 9 Å [subsets a and b] and r_c = 7 Å [subset d]. The mean potential energy $u(h)$ is plotted versus h in c for r_c = 9 Å.

The ERESPA, ENH, and ISO algorithms are described in detail in refs 63 and 64. Each MD run corresponded to its own size h of the outer time step. This size varied from run to run in a wide region from 4 fs up to 5120 fs. The inner step was fixed to h/n = 4 fs in all of the cases, meaning that the MTS parameter n changed from 1 at h = 4 fs to 1280 for h = 5120 fs. The choice h/n = 4 fs was dictated by the strength of the short-range intermolecular interactions. The total number l of outer steps was chosen in such a way as to cover nearly the same full propagation time $t = lh \sim 15$ ns at each given h.

During the EINO and INO propagations, we employed the three chains ($\mathscr{M}$ = 3) in the extended phase space at the relaxation time $\tau_t = \tau_{r,\alpha} = \tau$ = 10 fs. The triple concatenation with $n_t = n_\alpha^r$ = 8 was applied when integrating the thermostatting variables (Appendix B). The runs for τ = 40 and 400 fs without concatenation at $n_t = n_\alpha^r$ = 1 were examined too.

3.2. Numerical Results. The accuracy of the MD simulations was estimated by measuring the deviations of the oxygen−oxygen (OO), hydrogen−hydrogen (HH), and oxygen−hydrogen (OH) radial distribution functions $g(r)$ as well as the mean potential energy u of the system per molecule from their "exact" counterparts. The latter were precalculated using the Verlet integrator (i.e., RESPA at n = 1) with a tiny time step of h = 1 fs and a long simulation length of $l = t/h = 10^6$ to make the uncertainties negligibly small.[63,64] The normalized sum (multiplied on 1/3) of the three relative root-mean-square deviations $\Sigma = (\int_0^{L/2}[g(r) - g_0(r)]^2\, dr/\int_0^{L/2} g_0^2(r)\, dr)^{1/2}$ of $g(r)$ from the "exact" counterparts $g_0(r)$, related to the OO, HH, and OH distribution functions, is presented in Figure 2a for the ENH, ISO, INO, and EINO algorithms, depending on the size h of the outer time step. A more detailed behavior of $\Sigma(h)$ at not very large h is shown in subset b of Figure 2. The function of u on h is plotted in Figure 2c for each of the integrators (the "exact" level there is marked by the horizontal dashed line). The results in subsets a, b, and c correspond to the case r_c = 9 Å, while the dependencies of Σ on h at r_c = 7 Å are given in Figure 2d.

We see in Figure 2 that all of the curves for $\Sigma(h)$ or $u(h)$ start at small $h \sim 4$ fs with almost the same values, which are very close to their "exact" counterparts $\Sigma(0) = 0$ or $u(0) = -40.9$ kJ/mol (a slight discrepancy is due to statistical noise). However, the further behavior of $\Sigma(h)$ and $u(h)$ strongly depends on the type of the algorithm. For example, with increasing h, the microcanonical ERESPA integrator quickly loses stability, so that already at $h \gtrsim 20$ fs it is absolutely inadequate (see the almost vertical curve in Figure 2b). Note that the situation with the usual Verlet and RESPA algorithms is worse yet, where the maximum workable size of the time step cannot exceed 5 and 8 fs, respectively.[63,64] A better pattern can be observed for the canonical ENH and isokinetic ISO integrators. However, the best results are achieved by the proposed extrapolative EINO algorithm, which exhibits exceptional accuracy and stability. Indeed, the EINO deviations are minimal in all of the quantities investigated. For example, for r_c = 9 Å even at h = 512 fs, the EINO uncertainties Σ do not exceed a level of 0.3%, which is comparable with statistical noise. The impulsive INO integrator leads to an accuracy which is similar to that of the extrapolative EINO algorithm but only at not very large steps $h \lesssim 512$ fs. At longer $h \gtrsim 512$ fs, the advantage of the EINO method over the INO scheme becomes evident (cf. Figure 2a). For r_c = 7 Å, the latter is clearly inferior to the former in the whole region of h including small step sizes (cf. Figure 2d). Here, a decrease in the maximum allowable values for h is expected because the long-range interactions are stronger than in the case r_c = 9 Å. Nevertheless, even under these conditions, the extrapolative EINO algorithm can provide good precision ($\Sigma \sim 1.5\%$) in the picosecond region $h \sim 1024$ fs.

Note that the curves marked by INO and EINO in Figure 2 correspond to the value $\tau_t = \tau_{r,\alpha} = \tau$ = 10 fs. It provides a sufficiently strong coupling of the system with thermostats, because then the correlation time τ is only by a factor of 2.5 larger than the inner time step h/n = 4 fs. Upon increasing τ to 40 fs, the performance of the impulsive method drops drastically (see the dashed curve labeled by INO′). In particular, then the maximum acceptable outer steps reduce from $h \sim 512$ to 40 fs. At the same time, the extrapolative EINO algorithm is free of this negative feature. Even at τ = 400 fs, it continues to generate stable solutions in the whole region of time steps considered (see the solid curve labeled by EINO′). In addition, the EINO uncertainties only slightly increase with increasing h, giving a possibility of using the extrapolative approach up to giant values of the outer time step on the order of $h \sim 5120$ fs ≡ 5.12 ps, i.e., up to the picosecond region! We see in Figure 2a that the same level $\Sigma \approx 0.45\%$ of accuracy can be reached by the EINO, INO, ISO, ENH, and EOMTS integrators at the outer time steps h = 5120, 725, 90, 60, and 15 fs, respectively.

The superiority of the proposed extrapolative EINO approach over its impulsive INO counterpart can be explained by the fact that the former assumes only a spatial smoothness of the long-range interactions, which should not be necessarily small. On the other hand, the impulsive method requires both smoothness and weakness of these interactions. As a consequence, the nonresonance instabilities in the extrapolative method appear to be less sensitive to the increase of the outer time step size than the resonance effects of the impulsive scheme. This is confirmed in Figure 2, where the INO uncertainties $\Sigma(h)$ and $u(h)$, unlike the EINO ones, exhibit a resonance-like behavior with the existence of maxima and minima near certain values of h. Thus, the destabilizing resonant modes still exist in the impulsive INO method, despite the usage of the thermostats.

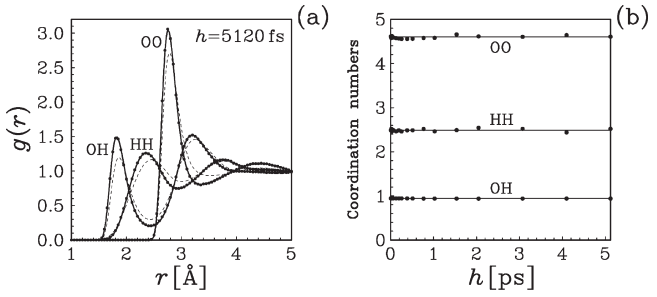

Figure 3. The OO, HH, and OH radial distribution functions of the TIP4P water obtained within the EINO algorithm for r_c = 9 Å at h = 5120 fs [circles in a] and the corresponding coordination numbers at various time steps $h \leq 5120$ fs [circles in b]. The "exact" results are plotted in subsets a and b by the solid curves and horizontal lines, respectively. The dashed curves in a correspond to the microcanonical EOMTS data at h = 19 fs.

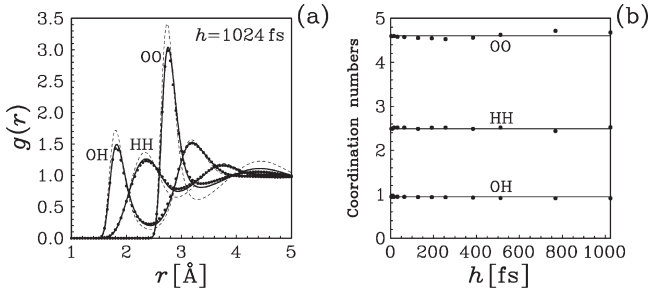

Figure 4. The same as for Figure 3 but at r_c = 7 Å and $h \leq 1024$ fs. The dashed curves in part a relate to a system with the long-range interactions excluded.

The OO, HH, and OH radial distribution functions $g(r)$ and their coordination numbers,[63,64] obtained for the TIP4P water by the EINO method for r_c = 9 Å (with τ = 400 fs) at h = 5120 fs and $h \leq 5120$ fs, are plotted in subsets a and b of Figure 3, respectively. The "exact" results (precalculated with the help of the Verlet integrator at h = 1 fs) are shown, too. The curves related to the microcanonical EOMTS scheme at h = 19 fs are included in Figure 3a as well. One can see that the deviations between the EOMTS data and the "exact" counterparts are too large already at h = 19 fs. On the other hand, the EINO algorithm still produces the radial distribution functions and coordination numbers with a high level of accuracy even at a huge time step of h = 5.12 ps. Indeed, the differences between the EINO and "exact" results are practically indistinguishable. Note that the largest acceptable outer time step reported earlier was h = 100 fs. It has been established within the translational motion INR algorithm[28,29] for a fully flexible water model. Thus, making the model rigid and using the proposed EINO approach have allowed us to significantly overcome this barrier. Now much larger (by a factor of 50) time step sizes on the order of 5120 fs are possible without affecting the structural properties and losing numerical stability. Such huge steps are up to 3 orders of magnitude longer than those feasible with the STS Verlet-like integrators.

The EINO distribution functions and coordination numbers corresponding to a more aggressive cutting-off with r_c = 7 Å (and τ = 10 fs) at $h \leq 1024$ fs are plotted in Figure 4. The deviations here between the solid curves and circles are more visible than for r_c = 9 Å (cf. Figure 3) because of the increased strength of $\mathbf{F}_w$ and $\mathbf{G}_w$. However, they are still sufficiently small to provide accurate results even at h = 1024 fs. Note that the long-range interactions influence significantly even structural properties. In order to demonstrate this, the curves related to a system with no long-range forces and torques ($\mathbf{F}_w = \mathbf{G}_w = 0$) are also included. We see that they differ appreciably from the "exact" counterparts in the whole range of varying the interatomic distance r.

The EINO simulations in the case of the standard extrapolation, i.e., when the long-range atomic forces $\mathbf{F}_{ia}^{(w)} = -\sum_{j\neq i}^{N}\sum_{b=1}^{M}(1 - \phi(r_{ij}^{ab}))\hat{\mathbf{r}}_{ij}^{ab}\varphi'_{ab}(r_{ij}^{ab})$ are held constant during outer time steps, have also been carried out. Unexpectedly, the pattern appeared to be significantly worse. It was established that already at relatively moderate $h \gtrsim 400$ fs, the atomic force extrapolation cannot be used because then the function Σ suddenly begins to increase after $l = t/h \lesssim 400$ steps, exceeding an unacceptable level of 10%. Note that in the atomic force extrapolation the molecular torque $\mathbf{G}_i^{(w)} = \sum_{a=1}^{M}\boldsymbol{\delta}_a \times (\mathbf{S}_i(t)\mathbf{F}_{ia}^{(w)})$ is not constant and varies in time due to the reorientation $\mathbf{S}_i(t)$ of the molecule. At first sight, such an extrapolation should look more precise since it takes into account the time dependence of $\mathbf{S}_i(t)$. However, this will be so in the Hamiltonian dynamics at tiny values of h when the changes of $\mathbf{F}_{ia}^{(w)}$ are small. In our case of the non-Hamiltonian propagation (which deals with huge h), it is much more important to provide a correct sampling of configurational points in phase space according to the canonical distribution (eq A1).

The molecular extrapolation just corresponds to the canonicity criteria. Indeed, the torque can be expressed in terms of the fixed (body-frame) dipole moment $\mathbf{m} = \sum_{a=1}^{M}\boldsymbol{\delta}_a q_a$ of the molecule and the electric field $\mathbf{E}_i(t) = \sum_{j\neq i}^{N}\sum_{b=1}^{M}(1 - \phi(|\mathbf{R}_i - \mathbf{r}_{jb}|))q_b(\mathbf{R}_i - \mathbf{r}_{jb})/|\mathbf{R}_i - \mathbf{r}_{jb}|^3$ (created by all the rest of the molecules due to the long-range contribution) as $\mathbf{G}_i^{(w)} \approx \mathbf{m} \times (\mathbf{S}_i(t)\mathbf{E}_i(t))$. Here, we have used that $-\varphi'_{ab}(r) = q_a q_b/r^2$ and $|\mathbf{r}_{ia} - \mathbf{r}_{jb}| \gg \sigma$ when calculating $\mathbf{F}_{ia}^{(w)}$, where σ is the diameter of the molecule. Thus, the proposed extrapolation implies that the combined quantity $\mathbf{S}_i\mathbf{E}_i$ remains fixed during the outer time step h when transiting the system from one phase space point to another. This preserves the true local canonical distribution $\exp[\mathbf{d}_i(t)\cdot\mathbf{E}_i(t)/(k_BT)] = \exp[\mathbf{m}\cdot(\mathbf{S}_i(t)\,\mathbf{E}_i(t))/(k_BT)]$ of the molecules, where $\mathbf{d}_i(t) = \sum_{a=1}^{M}(\mathbf{r}_{ia} - \mathbf{R}_i)q_a = \mathbf{S}_i^{+}(t)\mathbf{m}$ denotes the dipole moment in the laboratory frame and $-\mathbf{d}_i\cdot\mathbf{E}_i$ is the potential energy of the dipole in the electric field $\mathbf{E}_i$. On the other hand, the desired distribution can break significantly within the usual atomic extrapolation, where $\mathbf{S}_i\mathbf{E}_i$ varies in time, causing the strong instabilities at large h.

Rigorously speaking, the step by step propagation of phase variables in the EINO method can be interpreted as jumps from one relevant conformation to another without going through any physical path corresponding to the original (Hamiltonian) MD. That is why such huge outer time steps of order of $h \sim 5$ ps are possible here. The relaxation time of reorientation of a single water molecule just belongs to the picosecond region.[65] For instance, the normalized single dipole time correlation function $\langle\mathbf{d}_i(0)\cdot\mathbf{d}_i(t)\rangle/\langle\mathbf{d}_i(0)\cdot\mathbf{d}_i(0)\rangle$ decreases from 1 at t = 0 to 0.25 at t = 5 ps, lowering the precision of the molecular torque extrapolation when $h \gtrsim 5$ ps (here $\langle\ldots\rangle$ denotes statistical averaging). In view of this, the step size of $h \sim 5$ ps should be considered as the upper theoretical limit for accurate sampling of the canonical distribution within the EINO approach.

Consider now a question on the convergence of the results. The mean potential energy u and the error Σ in the radial distributions, obtained by the EINO propagation at τ = 400 fs

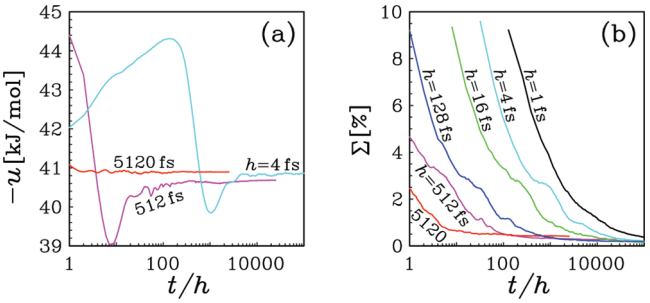

Figure 5. The EINO convergence of the potential energy [subset a] and the error in the radial distribution functions [subset b] with an increase in the length of the simulations at several fixed time steps, namely, h = 4, 16, 128, 512, and 5120 fs. The result in b for h = 1 fs corresponds to real dynamics obtained within the Verlet algorithm.

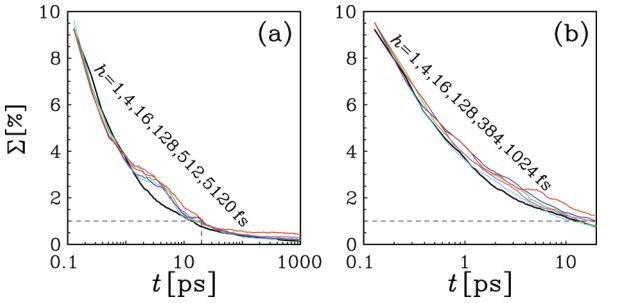

Figure 6. The EINO convergence of the error Σ in the radial distributions as a function of time t at several fixed step sizes h for r_c = 9 Å [τ = 400 fs, subset a] and r_c = 7 Å [τ = 10 fs, subset b]. Other notations are similar to those of Figure 5b.

and r_c = 9 Å, are shown in subsets a and b of Figure 5, respectively, as functions of the number $l = t/h$ of outer steps at several characteristic values of h. In all of the runs, the initial values of phase space variables $\mathbf{\Gamma}(0)$ were taken from the same thermodynamics state pre-equilibrated at $\mathscr{T}$ = 293 K but in the absence of weak long-range interactions (i.e., when $\mathbf{F}_w = \mathbf{G}_w = 0$). Then, the weak forces and torques were turned on at time t = 0, and the relaxation of the system from the perturbed state to the true configuration was observed. The potential energy was measured every outer time step when performing statistical averaging. Then, however, the measurements were carried out too frequently at small h, leading to a computational overload. In order to avoid this, the statistics for the radial distribution functions were accumulated after each fixed time interval of 128 fs for any h. Then, the costs due to the measurements will be negligibly small with respect to those spent on the integration of phase space variables.

We see in Figure 5a for the potential energy that the convergence at the largest stepsize h = 5120 fs is considerably faster than at moderate (h = 512 fs) and small (h = 4 fs) time steps. For instance, already after four outer time steps of size h = 5120 fs each, the potential energy almost achieves its limiting value and virtually does not change with a further increase in the simulation length. On the other hand, the asymptotic regime at h = 512 and 4 fs is reached only when $t/h \sim 40$ and 4×10^3, respectively. Thus, nearly the same time interval on the order of 20 ps is necessary to obtain reliable results in all three cases. Similar behavior is inherent in the error function Σ (see Figure 5b). Here, for instance, a level of Σ = 1% is reached after $t/h = 1.5 \times 10^4$, 4×10^3, 10^3, 150, 40, and 4 outer steps at h = 1, 4, 16, 128, 512, and 5120 fs, respectively. Such numbers are exactly inversely proportional to h, indicating that almost the same propagation time on the order of $t \sim 20$ ps is again enough to reproduce the radial distribution functions within 1% precision for all six sizes of h. This is confirmed in Figure 6, where the EINO curves are practically indistinguishable. Moreover, they are close to those corresponding to real dynamics at h = 1 fs (slight deviations in Figure 6a at intermediate t are explained by relaxation of the chain variables, which were set equal to zero, $w_j = 0$ and $v_j = 0$ with j = 2, ..., $\mathscr{M}$, at t = 0). Hence, the quasidynamics produced by the EINO method are free from the drag that would be caused by the introduction of the thermostats.

3.3. CPU Speedup. From the information already given, we can conclude that the EINO approach allows huge sizes of the outer time step which at $h \sim 5120$ fs are by a factor of 5120/8 = 640 larger than those of the standard RESPA scheme (for the latter, the maximal h is[63,64] on the order of 8 fs). Note, however, that the ideal 640-fold increase in computational efficiency cannot be realized because a nonzero portion $\theta > 0$ of CPU time is needed to evaluate the short-range forces and torques. This portion ($0 < \theta < 1$) can be measured in terms of the cost ratio of calculating such (cheap) forces and torques to (expensive) long-range interactions. Let h_{MTS} and h_{STS} be the maximal sizes of the time step which are possible to use by some MTS ($n > 1$) and STS ($n_{STS} = 1$) algorithms, respectively. Then, taking into account that nearly the same integration time t (see section 3.2) is required to reach an asymptotic behavior for all of the approaches considered, the actual relative speedup can be estimated as $\Lambda = n(1 - \xi)(1 + \theta)/(1 + n\theta)$. Here, $n = h_{MTS}/h_{STS}$ with h_{STS} = 4 fs, since the fixed inner step (h/n = 4 fs) is employed for any n to achieve the same precision corresponding to fast dynamics. The multiplier $0 < \xi < 1$ takes into account the overhead of the INO and EINO techniques on the propagation of extra (thermostatting) phase variables.

Thus, the ideal CPU speedup $\Lambda_{max} = n$ can be expected only in a hypothetical case when $\theta \rightarrow 0$ and $\xi \rightarrow 0$. In practical calculations $\Lambda < n$ because the quantities θ and ξ are always finite. These quantities depend on details of the simulations, implemented program code, and the compilers and platform used. The present calculations were performed on the SGI Altix 4700 supercomputer using the Linux Intel Fortran compiler. In the case of N = 512 with r_c = 9 Å and $R_c = L/2$ = 12.417 Å, we found that $\theta \approx 1/20$, while $\xi \approx 0.1$. With decreasing r_c to 7 Å, the ratio θ decreases to 1:50. The CPU speedups Λ for different MTS algorithms, obtained in the MD simulations of water with respect to the STS Verlet integrator (at h = 4 fs), are plotted in Figure 7 as functions of the size of the outer time step by the lower (r_c = 9 Å) and medium (r_c = 7 Å) lying curves. The symbols are related to the maximal values of h, which are still safely workable within a given algorithm.

As can be seen in Figure 7, the RESPA and ERESPA schemes only slightly (Λ = 2−3) increase the efficiency. A better pattern is observed for the ENH and ISO integrators which can reduce the overall CPU costs up to Λ = 10 times. The best results are achieved within the INO (triangles) and EINO (full circles) algorithms. They are able to speed up the MD calculations by factors of Λ = 20 and 40, respectively. Such an increase is explained by larger values of h, and thus n, which are inherent in these algorithms. Remember that to cover the same integration time t, the total number of steps t/h decreases with increasing h.

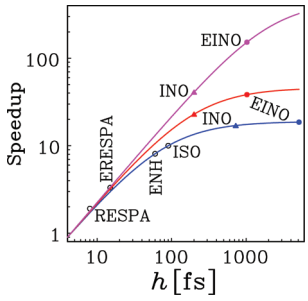

Figure 7. CPU speedup for different MTS algorithms in MD simulations of water relative to the STS integrator. The results obtained for $N = 512$ molecules with cutoff radii $r_c = 9$ Å and $r_c = 7$ Å appear as the lower and middle lying curves, respectively. The theoretical estimations for $N = 5120$ with $r_c = 7$ Å are presented by the upper curve. The symbols correspond to the maximum allowable values of h for each of the algorithms.

This leads to an increase in the overall computational efficiency since the extra INO and EINO thermostatting costs are minimal ($\xi \approx 0.1$). Note also that the values of Λ increase with decreasing the cutoff radius r_c from 9 Å to 7 Å because of lowering θ, despite the decrease of the maximum allowable size of the outer time step. This is so because then the computational costs spent on the calculation of the short-range interactions drop significantly. Moreover, taking into account that the ratio θ is approximately proportional to $(r_c/R_c)^3$, the speedup will further increase with an increase in the size of the system. For instance, for a collection of $N = 5120$ water molecules, it is expected that the relative speedup will be on the order of $\Lambda \approx 150$ (the upper lying curve in Figure 7). The choice of optimal values for r_c and R_c goes beyond the scope of the current paper and will be reported elsewhere.

4. CONCLUSION

We have proposed a new multiscale approach to overcome the restrictions on time step sizes in MD simulations of interaction site models of fluids with orientational degrees of freedom. It presents a nontrivial combination of the decomposition operator method with a special extrapolation of intermolecular interactions complemented by a modified isokinetic Nosé−Hoover chain thermostat. This has allowed us to substantially enlarge the size of the outer time step when propagating the phase space variables and, thus, significantly improve the efficiency of MD computations. As is shown on the basis of MD simulations for the rigid TIP4P model of ambient water, giant step sizes up to several picoseconds become possible without losing numerical stability and affecting equilibrium properties. Such steps are up to 3 orders of magnitude larger than those of single-scale Verlet-type integrators and by a factor of 50 longer than the maximal time steps feasible with the best previous multiple time step algorithms. This constitutes a considerable advantage for MD simulation of molecular liquids, with many applications in solution chemistry.

The new approach could be extended to more complex models of liquids and solutions in the presence of both rigid and flexible atomic groups, including solvated proteins and other biomolecules. The latter presents the biggest challenges, since it requires introducing three or more time scales. The fastest one is related to the internal bond and bending vibrations of atoms within the molecule. For large proteins, we should take into account very slow collective dynamics of molecular domains which can influence the atomic motion. Then, an interplay between the solvent hydrodynamics and solute movements will also take place, resulting in an extremely large separation between the time scales. Similar difficulties might arise when coupling the proposed MD methodology with the statistical−mechanical, 3D-RISM-KH molecular theory of solvation.[67−69] All of these problems will be a subject of future investigations.

APPENDIX A: CANONICITY OF THE INC DYNAMICS

It should be pointed out that the qusidynamics generated by the proposed INC equations of motion cannot provide the canonical distribution in velocity space because of the imposed individual kinetic constraints (eq 4). Nevertheless, the configurational part $\mathcal{Z}(\mathcal{J})$ of the extended partition function $\mathcal{G}(\mathcal{J})$ obtained within the INC approach does correspond to the true canonical distribution of the physical system (at temperature $\mathcal{J}$). This allows one to perform the genuine canonical averages of position- and orientation-dependent properties in equilibrium.

Indeed, taking into account eq 4, one sees that the partition function in the extended phase space (the basic system plus chain thermostats) is of the form

$$\begin{aligned}\mathcal{Q} = \int \prod_{i=1}^{N} \mathrm{d}\mathbf{V}_i\, \mathrm{d}\boldsymbol{\Omega}_i\, \mathrm{d}\boldsymbol{\nu}_i\, \mathrm{d}\mathbf{w}_i \\ \times \prod_{\alpha}^{x,y,z} \delta(\mu V_{i,\alpha}^2 + m_i \nu_{1,i,\alpha}^2 - k_B\mathcal{T}) \\ \times \prod_{\alpha}^{X,Y,Z} \delta(J_\alpha \Omega_{i,\alpha}^2 + j_{i,\alpha} w_{1,i,\alpha}^2 - k_B\mathcal{T}) \\ \times \prod_{a=1}^{M} \mathrm{d}\mathbf{r}_{ia}\, e^{\{-U(\mathbf{r}) + \sum_{i,\alpha} T_{i,\alpha}(V,\Omega,\nu,\mathrm{w})\}/(k_B\mathcal{T})} \\ \propto \prod_{i,a=1}^{N,M} \mathrm{d}\mathbf{r}_{ia}\, e^{-U(\mathbf{r})/(k_B\mathcal{T})} \propto \mathcal{Z}(\mathcal{T}) \end{aligned} \quad \text{(A1)}$$

where $\mathbf{r} \equiv \{\mathbf{r}_{ia}\}$ denotes the whole set of atomic positions, $U(\mathbf{r})$ is the full potential energy of the system, and $T_{i,\alpha}(V,\Omega,\nu,w) = \mu V_{i,\alpha}^2/2 + J_\alpha \Omega_{i,\alpha}^2/2 + \sum_{j=1}^{\mathcal{M}} (m\nu_{j,i,\alpha}^2 + j_\alpha w_{j,i,\alpha}^2)/2$ are the i,αth components of the total kinetic energy, which includes the real and imaginary velocity-type variables. Integrating in eq A1 over all of these variables and taking into account the presence of the δ functions gives the desired result $\mathcal{G} \propto \mathcal{Z}$.

APPENDIX B: ANALYTICAL EXPRESSIONS FOR SINGLE EXPONENTIAL OPERATORS

Apart from the possibility of using huge time steps, another great advantage of the proposed EINO method is that all of the single exponentials in eq 13 can be handled analytically. Really, in view of eq 7, the actions of operators $e^{B_{ws}h/(2n)}$ on translational phase variables $\mathbf{V}$ and $\boldsymbol{\nu}_1$ can first be factorized into the Cartesian components

$$e^{B_{ws}\frac{h}{2n}}\{\mathbf{V}, \boldsymbol{\nu}_1\} = \prod_{\alpha}^{x,y,z} e^{B_{ws,\alpha}^t \frac{h}{2n}}\{V_\alpha, \nu_{1,\alpha}\} \quad \text{(B1)}$$

and then expressed in terms of hyperbolic functions

$$\begin{aligned} & e^{B_{ws,\alpha}^t \frac{h}{2n}}\{V_\alpha, \nu_{1,\alpha}\} \\ & = \left\{ \frac{V_\alpha + \vartheta_\alpha^{-1} \tanh(h_{t,\alpha})}{1 + V_\alpha \vartheta_\alpha \tanh(h_{t,\alpha})}, \frac{\nu_{1,\alpha} \cosh^{-1}(h_{t,\alpha})}{1 + V_\alpha \vartheta_\alpha \tanh(h_{t,\alpha})} \right\} \end{aligned} \quad \text{(B2)}$$

dx.doi.org/10.1021/ct200157x | J. Chem. Theory Comput. 2012, 8, 6–16

with $\vartheta_\alpha = (T^{\mathrm{t}}_\alpha/\mu)^{-1/2}$ and $h_{\mathrm{t},\alpha} = \mu^{-1}F_{\mathrm{ws},\alpha}\vartheta_\alpha h/(2n)$. The expressions in the case of rotation motion for orientational phase variables $\mathbf{\Omega}$ and $\mathbf{w}_1$ are somewhat more complicated, namely,

$$e^{B_{\mathrm{ws}}\frac{h}{2n}}\{\mathbf{\Omega},\mathbf{w}_1\} = e^{B^{\mathrm{r}}_{\mathrm{ws},X}\frac{h}{4n}}\{\Omega_X, w_{1,X}\}\, e^{B^{\mathrm{r}}_{\mathrm{ws},Y}\frac{h}{4n}}\{\mathbf{\Omega}_Y, w_{1,Y}\} \times e^{B^{\mathrm{r}}_{\mathrm{ws},Z}\frac{h}{2n}}\{\mathbf{\Omega}_Z, w_{1,Z}\} e^{B^{\mathrm{r}}_{\mathrm{ws},Y}\frac{h}{4n}}\{\Omega_Y, w_{1,Y}\} \times e^{B^{\mathrm{r}}_{\mathrm{ws},X}\frac{h}{4n}}\{\Omega_X, w_{1,X}\} + \mathcal{O}(h^3) \tag{B3}$$

where

$$e^{B^{\mathrm{r}}_{\mathrm{ws},\alpha}\frac{h}{2n}}\{\Omega_\alpha, w_{1,\alpha}\} = \left\{\frac{\Omega_\alpha + \chi_\alpha^{-1}\tanh(h_{\mathrm{r},\alpha})}{1 + \Omega_\alpha\chi_\alpha\tanh(h_{\mathrm{r},\alpha})}, \frac{w_{1,\alpha}\cosh^{-1}(h_{\mathrm{r},\alpha})}{1 + \Omega_\alpha\chi_\alpha\tanh(h_{\mathrm{r},\alpha})}\right\} \tag{B4}$$

with $\chi_\alpha = (T^{\mathrm{r}}_\alpha/J_\alpha)^{-1/2}$, $h_{\mathrm{r},\alpha} = J_\alpha{}^{-1}G_\alpha\chi_\alpha h/(2n)$, and $G_\alpha = G_{\mathrm{ws},\alpha} + (J_\beta - J_\gamma)\Omega_\beta\Omega_\gamma$. It should be mentioned that the principal ($\alpha = X,Y,Z$) components of the orientational part of $B_{\mathrm{ws},\alpha}$ do not commute because of the existence of the inertial torque terms $(J_\beta - J_\gamma)\Omega_\beta\Omega_\gamma$ (see eq 7). Thus, contrary to the simple factorization (eq B1) used for the Cartesian ($\alpha = x,y,z$) components of the translational part of $B_{\mathrm{ws},\alpha}$, the extra decomposition (eq B3) has been exploited to achieve the desired $\mathcal{O}(h^3)$ one-step accuracy. Then, the partial operators $e^{B_{\mathrm{ws},\alpha}h/(2n)}$ acting on angular velocities $\mathbf{\Omega}$ and $\mathbf{w}_1$ will change only their αth component according to eq B4 at constant values of the remaining two parts β and γ.

Additional splitting is needed for analytical handling of the INC thermostat propagator $e^{B_{\mathrm{inc}}h/(2n)}$. In view of eqs 8−10, it can be decomposed as

$$e^{B_{\mathrm{inc}}\frac{h}{2n}} = \prod_{\alpha}^{x,y,z}\left[e^{B^{\mathrm{t}}_{\mathrm{inc},\alpha}\frac{h}{2nn_{\mathrm{t}}}}\right]^{n_{\mathrm{t}}} \prod_{\alpha}^{X,Y,Z}\left[e^{B^{\mathrm{r}}_{\mathrm{inc},\alpha}\frac{h}{2nn^{\mathrm{r}}_\alpha}}\right]^{n^{\mathrm{r}}_\alpha} \tag{B5}$$

where (for $\mathscr{M} \le 3$):

$$\begin{aligned} e^{B^{\mathrm{t}}_{\mathrm{inc},\alpha}\frac{h}{2nn_{\mathrm{t}}}} &= e^{B_{\nu_{3,\alpha}}\frac{h}{8nn_{\mathrm{t}}}}\, e^{B_{\nu_{2,\alpha}}\frac{h}{4nn_{\mathrm{t}}}}\, e^{B_{\nu_{3,\alpha}}\frac{h}{8nn_{\mathrm{t}}}}\, e^{B_{V,\nu,\alpha}\frac{h}{2nn_{\mathrm{t}}}}\, e^{B_{\nu_{3,\alpha}}\frac{h}{8nn_{\mathrm{t}}}} \times e^{B_{\nu_{2,\alpha}}\frac{h}{4nn_{\mathrm{t}}}}\, e^{B_{\nu_{3,\alpha}}\frac{h}{8nn_{\mathrm{t}}}}, \\ e^{B^{\mathrm{r}}_{\mathrm{inc},\alpha}\frac{h}{2nn^{\mathrm{r}}_\alpha}} &= e^{B_{w_{3,\alpha}}\frac{h}{8nn^{\mathrm{r}}_\alpha}}\, e^{B_{w_{2,\alpha}}\frac{h}{4nn^{\mathrm{r}}_\alpha}}\, e^{B_{w_{3,\alpha}}\frac{h}{8nn^{\mathrm{r}}_\alpha}}\, e^{B_{\Omega,w,\alpha}\frac{h}{2nn^{\mathrm{r}}_\alpha}}\, e^{B_{w_{3,\alpha}}\frac{h}{8nn^{\mathrm{r}}_\alpha}} \times e^{B_{w_{2,\alpha}}\frac{h}{4nn^{\mathrm{r}}_\alpha}}\, e^{B_{w_{3,\alpha}}\frac{h}{8nn^{\mathrm{r}}_\alpha}} \end{aligned} \tag{B6}$$

In eq B6, the internal loops with $n_{\mathrm{t}} \ge 1$ and $n^{\mathrm{r}}_\alpha \ge 1$ cycles have been introduced to improve the accuracy of the INC thermostat integration. This is necessary if the thermostatting correlation times are small, i.e., if $\tau_{\mathrm{t}} \sim h/n$ and $\tau_{\mathrm{r},\alpha} \sim h/n$. Then, n_{t} and n^{r}_α should be chosen in order to satisfy the inequalities $h/(2nn_{\mathrm{t}}) \ll \tau_{\mathrm{t}}$ and $h/(2nn^{\mathrm{r}}_\alpha) \ll \tau_{\mathrm{r},\alpha}$. The extra precision can be reached by applying the triple concatenation[56] of eq B6 at $h \equiv \varsigma h$, $(1 - 2\varsigma)h$, and again at $h \equiv \varsigma h$, where $\varsigma = 1/(2 - 2^{1/2})$. Such a concatenation reduces the uncertainty of the decompositions from $\mathcal{O}(h^3)$ to a negligibly small level of $\mathcal{O}(h^5)$. When $\tau_{\mathrm{t}} \gg h/n$ and $\tau_{\mathrm{r},\alpha} \gg h/n$, one can set $n_{\mathrm{t}} = 1$ and $n^{\mathrm{r}}_\alpha = 1$.

The action of the single exponential operators in eq B6 on the extended phase variables can be presented in terms of elementary functions as well. Using eqs 9 and 10, one finds for the translational components

$$\begin{aligned} e^{B_{V,\nu,\alpha}\frac{h}{2nn_{\mathrm{t}}}}\{V_\alpha, \nu_{1,\alpha}\} &= \left[1 + \frac{\nu^2_{1,\alpha}\tau^2_{\mathrm{t}}k_{\mathrm{B}}\mathscr{T}}{4T^{\mathrm{t}}_\alpha} \times (e^{-\nu_{2,\alpha}\frac{h}{nn_{\mathrm{t}}}} - 1)\right]^{-1/2}\{V_\alpha, e^{-\nu_{2,\alpha}\frac{h}{2nn_{\mathrm{t}}}}\nu_{1,\alpha}\}, \\ e^{B_{\nu_{2,\alpha}}\frac{h}{4nn_{\mathrm{t}}}}\nu_{2,\alpha} &= \nu_{2,\alpha}e^{-\nu_{3,\alpha}\frac{h}{4nn_{\mathrm{t}}}} + 2\left(\nu^2_{1,\alpha} - \frac{1}{\tau^2_{\mathrm{t}}}\right) \times \frac{e^{-\nu_{3,\alpha}\frac{h}{8nn_{\mathrm{t}}}}}{\nu_{3,\alpha}}\sinh\left(\nu_{3,\alpha}\frac{h}{8nn_{\mathrm{t}}}\right), \\ e^{B_{\nu_{3,\alpha}}\frac{h}{8nn_{\mathrm{t}}}}\nu_{3,\alpha} &= \nu_{3,\alpha} + \left(\nu^2_{2,\alpha} - \frac{1}{\tau^2_{\mathrm{t}}}\right)\frac{h}{8nn_{\mathrm{t}}} \end{aligned} \tag{B7}$$

The analogous expressions for the rotational components read

$$\begin{aligned} e^{B_{\Omega,w,\alpha}\frac{h}{2nn^{\mathrm{r}}_\alpha}}\{\Omega_\alpha, w_{1,\alpha}\} &= \left[1 + \frac{w^2_{1,\alpha}\tau^2_{\mathrm{r},\alpha}k_{\mathrm{B}}\mathscr{T}}{4T^{\mathrm{r}}_\alpha} \times (e^{-w_{2,\alpha}\frac{h}{nn^{\mathrm{r}}_\alpha}} - 1)\right]^{-1/2}\{\Omega_\alpha, e^{-w_{2,\alpha}\frac{h}{2nn^{\mathrm{r}}_\alpha}}w_{1,\alpha}\}, \\ e^{B_{w_{2,\alpha}}\frac{h}{4nn^{\mathrm{r}}_\alpha}}w_{2,\alpha} &= w_{2,\alpha}e^{-w_{3,\alpha}\frac{h}{4nn^{\mathrm{r}}_\alpha}} + 2\left(w^2_{1,\alpha} - \frac{1}{\tau^2_{\mathrm{r},\alpha}}\right) \times \frac{e^{-w_{3,\alpha}\frac{h}{8nn^{\mathrm{r}}_\alpha}}}{w_{3,\alpha}}\sinh\left(w_{3,\alpha}\frac{h}{8nn^{\mathrm{r}}_\alpha}\right), \\ e^{B_{w_{3,\alpha}}\frac{h}{8nn^{\mathrm{r}}_\alpha}}w_{3,\alpha} &= w_{3,\alpha} + \left(w^2_{2,\alpha} - \frac{1}{\tau^2_{\mathrm{r},\alpha}}\right)\frac{h}{8nn^{\mathrm{r}}_\alpha} \end{aligned} \tag{B8}$$

Note that the simultaneous transformations of $(V_\alpha,\nu_{1,\alpha})$ and $(\Omega_\alpha,w_{1,\alpha})$ given by eqs B2 and B4 as well as B7 and B8 conserve the individual kinetic constraints (eq 4) to within a machine accuracy at any time step size h. This has been achieved by the special decompositions (eqs 7 and 8) and analytical (i.e., exact) expressions for the single exponential propagators. Such a conservation must be considered as a very important feature of the proposed algorithm because now in principle arbitrarily large values of h can be exploited without a loss of numerical stability despite the fact that the phase trajectories are produced approximately ($\mathcal{O}(h^2) \neq 0$ in eq 13). The stability can be improved additionally by recalculating $T^{\mathrm{t}}_\alpha = \mu V^2_\alpha/2 + k_{\mathrm{B}}\mathscr{J}\tau^2_{\mathrm{t}}\nu^2_{1,\alpha}/4$ and $\mathscr{J}T^{\mathrm{r}}_\alpha \equiv J_\alpha\Omega^2_\alpha/2 + k_{\mathrm{B}}\mathscr{J}\tau^2_{\mathrm{r},\alpha}w^2_{\mathrm{r},\alpha}/4$ in eqs B2, B4, B7, and B8 before each time step rather to merely put $T^{\mathrm{t}}_\alpha = k_{\mathrm{B}}\mathscr{J}/2$ and $T^{\mathrm{r}}_\alpha = k_{\mathrm{B}}\mathscr{J}/2$. This prevents the accumulation of machine errors and provides the constraint conservation not only locally but also globally for any time $t \gg h$.

It should be pointed out also that the relatively large number of single exponentials appearing in the EINO propagation presents no numerical difficulties. The action of these exponentials on a phase space point leads to simple analytical transformations given by elementary functions. They incur practically negligible computational costs, compared to those necessary to spend on the calculation of intermolecular forces and torques. The numerical overheating can be reduced to a minimum by replacing the elementary (exponents and hyperbolic trigonometric) functions with their rational counterparts. This can be useful especially for thermostatting propagation (eqs B5 and B6) if $n_{\mathrm{t}} > 1$ and $n_\alpha{}^{\mathrm{r}} > 1$. The rational counterparts can be obtained by expanding the

dx.doi.org/10.1021/ct200157x |J. Chem. Theory Comput. 2012, 8, 6–16

functions in power series with respect to their arguments, taking into account that the latter are small. The expansion can be restricted to a finite number of terms within the required precision. For instance, for the first lines of eq B7 and B8, when $\xi = \nu_{2,\alpha}h/(2nn_t)$ or $\xi = w_{2,\alpha}h/(2nn_\alpha^r)$, we have $e^{-\xi} = (1 - \xi/2)/(1 + \xi/2) + \mathcal{O}(h^3)$ or even $e^{-\xi} = (1 - \xi/2 + \xi^2/12)/(1 + \xi/2 + \xi^2/12) + \mathcal{O}(h^5)$ that already exceeds the one-step precision $\mathcal{O}(h^3)$ of the basic integration. At the same time, in these lines, it is necessary to put $e^{-2\xi} = (e^{-\xi})^2$ rather than directly expand $e^{-2\xi}$. This maintains the exact conservation of the kinetic constraints. Similar counterparts can be used in the second lines of eqs B7 and B8 exploiting the equality $2e^{-\xi/2}\sinh(\xi/2)/\xi = (1 - e^{-\xi})/\xi = 1/(1 + (\xi/2) + \mathcal{O}(h^3)$ or $1/(1 + \xi/2 + \xi^2/12) + \mathcal{O}(h^5)$.

AUTHOR INFORMATION

Corresponding Author

*E-mail: omelyan@icmp.lviv.ua; andriy.kovalenko@nrc-cnrc.gc.ca.

ACKNOWLEDGMENT

We gratefully acknowledge the financial support by the ArboraNano—the Canadian Forest NanoProducts Network and by the National Research Council (NRC) of Canada. I.P.O. is thankful for the hospitality during his stay at the University of Alberta and the National Institute for Nanotechnology.

REFERENCES

(1) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Clarendon: Oxford, U.K., 1987.

(2) Frenkel, D.; Smit, B. *Understanding Molecular Simulation: from Algorithms to Applications*; Academic Press: New York, 1996.

(3) Theodorou, D. N.; Kotelyanski, M. *Simulation Methods for Polymers*; Marcel Dekker: New York, 2004.

(4) Leimkuhler, B.; Reich, S. *Simulating Hamiltonian Dynamics*; Cambridge University Press: Cambridge, U.K., 2005.

(5) Rojnuckarin, A.; Kim, S.; Subramaniam, S. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 4288.

(6) Hernández, G.; Jenney, F. E., Jr.; Adams, M. W. W.; LeMaster, D. M. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 3166.

(7) Karplus, M.; McCammon, J. A. *Nat. Struct. Biol.* **2002**, *9*, 646.

(8) Zhang, Y.; Peters, M. H.; Li, Y. *Proteins: Struct., Funct., Genet.* **2003**, *52*, 339.

(9) Grubmüller, H.; Heller, H.; Windemuth, A.; Schulten, K. *Mol. Simul.* **1991**, *6*, 121.

(10) Tuckerman, M. E.; Berne, B. J.; Martyna, G. J. *J. Chem. Phys.* **1992**, *97*, 1990.

(11) Stuart, S. J.; Zhou, R.; Berne, B. J. *J. Chem. Phys.* **1996**, *105*, 1426.

(12) Kopf, A.; Paul, W.; Dünweg, B. *Comput. Phys. Commun.* **1997**, *101*, 1.

(13) Zhang, G.; Schlick, T. *J. Comput. Chem.* **1993**, *14*, 1212.

(14) Zhang, G.; Schlick, T. *J. Chem. Phys.* **1994**, *101*, 4995.

(15) Schlick, T.; Barth, E.; Mandziuk, M. *Annu. Rev. Biophys. Biomol. Struct.* **1997**, *26*, 181.

(16) Barth, E.; Schlick, T. *J. Chem. Phys.* **1998**, *109*, 1617.

(17) Garcia-Archilla, B.; Sanz-Serna, J. M.; Skeel, R. D. *SIAM J. Sci. Comput.* **1998**, *20*, 930.

(18) Izaguirre, J. A.; Reich, S.; Skeel, R. D. *J. Chem. Phys.* **1999**, *110*, 9853.

(19) Ma, Q.; Izaguirre, J. A. *Multiscale Model. Simul.* **2003**, *2*, 1.

(20) Izaguirre, J. A.; Catarello, D. P.; Wozniak, J. M.; Skeel, R. D. *J. Chem. Phys.* **2001**, *114*, 2090.

(21) Skeel, R. D.; Izaguirre, J. A. *Mol. Phys.* **2002**, *100*, 3885.

(22) Melchionna, S. *J. Chem. Phys.* **2007**, *127*, 044108.

(23) Martyna, G. J.; Tuckerman, M. E.; Tobias, D. J.; Klein, M. L. *Mol. Phys.* **1996**, *87*, 1117.

(24) Cheng, A.; Merz, K. M., Jr. *J. Phys. Chem. B* **1999**, *103*, 5396.

(25) Komeiji, J. *THEOCHEM* **2000**, *530*, 237.

(26) Shinoda, W.; Mikami, M. *J. Comput. Chem.* **2003**, *24*, 920.

(27) Minary, P.; Martyna, G. J.; Tuckerman, M. E. *J. Chem. Phys.* **2003**, *118*, 2510.

(28) Minary, P.; Tuckerman, M. E.; Martyna, G. J. *Phys. Rev. Lett.* **2004**, *93*, 150201.

(29) Abrams, J. B.; Tuckerman, M. E.; Martyna, G. J. *Computer Simulations in Condensed Matter Systems: From Materials to Chemical Biology*; Springer-Verlag: Berlin, 2006; Vol. 1. [*Lect. Notes Phys.* 2006, *703*, 139.]

(30) Zhou, R.; Berne, B. J. *J. Chem. Phys.* **1995**, *103*, 9444.

(31) Watanabe, M.; Karplus, M. *J. Phys. Chem.* **1995**, *99*, 5680.

(32) Barth, E.; Schlick, T. *J. Chem. Phys.* **1998**, *109*, 1633.

(33) Mandziuk, M.; Schlick, T. *Chem. Phys. Lett.* **1995**, *237*, 525.

(34) Schlick, T.; Mandziuk, M.; Skeel, R. D.; Srinivas, K. *J. Comput. Phys.* **1998**, *140*, 1.

(35) Ma, Q.; Izaguirre, J. A.; Skeel, R. D. *SIAM J. Sci. Comput.* **2003**, *24*, 1951.

(36) Ciccotti, G.; Ryckaert, J. P.; Ferrario, M. *Mol. Phys.* **1982**, *47*, 1253.

(37) Andersen, H. C. *J. Comput. Phys.* **1983**, *52*, 24.

(38) MacKerell, A. D., Jr.; Bashford, D.; Bellott, M.; Dunbrack, R. L., Jr.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E., III; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiórkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586.

(39) Chen, B.; Martin, M. G.; Siepmann, J. I. *J. Phys. Chem. B* **1998**, *102*, 2578.

(40) Dullweber, A.; Leimkuhler, B.; McLachlan, R. *J. Chem. Phys.* **1997**, *107*, 5840.

(41) Omelyan, I. P. *Comput. Phys.* **1998**, *12*, 97.

(42) Omelyan, I. P. *Comput. Phys. Commun.* **1998**, *109*, 171.

(43) Omelyan, I. P. *Phys. Rev. E* **1998**, *58*, 1169.

(44) Matubayasi, N.; Nakahara, M. *J. Chem. Phys.* **1999**, *110*, 3291.

(45) Miller, T. F., III; Eleftheriou, M.; Pattnaik, P.; Ndirango, A.; Newns, D.; Martyna, G. J. *J. Chem. Phys.* **2002**, *116*, 8649.

(46) Omelyan, I. P. *J. Chem. Phys.* **2007**, *127*, 044102.

(47) Omelyan, I. P. *Phys. Rev. E* **2008**, *78*, 026702.

(48) Ikeguchi, M. *J. Comput. Chem.* **2004**, *25*, 529.

(49) Kamberaj, H.; Low, R. J.; Neal, M. P. *J. Chem. Phys.* **2005**, *122*, 224114.

(50) Okumura, H.; Itoh, S. G.; Okamoto, Y. *J. Chem. Phys.* **2007**, *126*, 084103.

(51) Kutteh, R.; Jones, R. B. *Phys. Rev. E* **2000**, *61*, 3186.

(52) Terada, T.; Kidera, A. *J. Chem. Phys.* **2002**, *116*, 33.

(53) Davidchack, R. L.; Handel, R.; Tretyakov, M. V. *J. Chem. Phys.* **2009**, *130*, 234101.

(54) Swope, W. C.; Andersen, H. C.; Berens, P. H.; Wilson, K. R. *J. Chem. Phys.* **1982**, *76*, 637.

(55) Omelyan, I. P.; Mryglod, I. M.; Folk, R. *Phys. Rev. E* **2002**, *65*, 056706.

(56) Creutz, M.; Gocksch, A. *Phys. Rev. Lett.* **1989**, *63*, 9.

(57) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.

(58) Omelyan, I. P. *Comput. Phys. Commun.* **1997**, *107*, 113.

(59) Zhou, R.; Harder, E.; Xu, H.; Berne, B. J. *J. Chem. Phys.* **2001**, *115*, 2348.

(60) Qian, X.; Schlick, T. *J. Chem. Phys.* **2002**, *116*, 5971.

(61) Han, G.; Deng, Y.; Glimm, J.; Martyna, G. *Comput. Phys. Commun.* **2007**, *176*, 271.

(62) Morrone, J. A.; Zhou, R.; Berne, B. J. *J. Chem. Theory Comput.* **2010**, *6*, 1798.

(63) Omelyan, I. P.; Kovalenko, A. *J. Chem. Phys.* **2011**, *135*, 114110.
(64) Omelyan, I. P.; Kovalenko, A. *J. Chem. Phys.* **2011**, to be published.
(65) Omelyan, I. P. *Mol. Phys.* **1998**, *93*, 123.
(66) Omelyan, I. P.; Mryglod, I. M.; Tokarchuk, M. V. *Condens. Matter Phys.* **2005**, *8*, 25.
(67) Kovalenko, A. In *Molecular Theory of Solvation*; Hirata, F., Ed.; Kluwer Academic Publishers: Norwell, MA, 2003; Vol. 24, Chapter 4.
(68) Miyata, T.; Hirata, F. *J. Comput. Chem.* **2008**, *29*, 871.
(69) Luchko, T.; Gusarov, S.; Roe, D. R.; Simmerling, C.; Case, D. A.; Tuszynski, J.; Kovalenko, A. *J. Chem. Theory Comput.* **2010**, *6*, 607.