Time reversible molecular dynamics algorithms with holonomic bond constraints in the *NPH* and *NPT* ensembles using molecular scaling

Trond Ingebrigtsen,[1,a)] Ole J. Heilmann,[2] Søren Toxvaerd,[2] and Jeppe C. Dyre[2]
[1]*Department of Chemistry, H. C. Ørsted Institute, University of Copenhagen, Universitetsparken 5, DK-2100 Copenhagen Ø, Denmark*
[2]*Department of Sciences, DNRF Centre "Glass and Time," IMFUFA, Roskilde University, Postbox 260, DK-4000 Roskilde, Denmark*

(Received 21 November 2009; accepted 23 February 2010; published online 19 April 2010)

A modification of the constrained equations of motion of Kalibaeva *et al.* [Mol. Phys. **101**, 765 (2003)] in the *NPH* and *NPT* ensembles is presented. The modified equations of motion are discretized using central-difference techniques, and the derived integrators are time reversible and conserve the invariant phase space measure. The constraint algorithm builds on the work of Toxvaerd *et al.* [J. Chem. Phys. **131**, 064102 (2009)] in the *NVE* and *NVT* ensembles: it thus conserves the holonomic bond constraints at the finite machine precision level in the *NPH* and *NPT* ensembles. The algorithms were tested on a system of $n=320$ *ortho*-terphenyl molecules, arriving at the target temperature and pressure in a low and high pressure state. Isobaric heat capacities in the *NPH* and *NPT* ensembles were calculated for comparison using the fluctuation formulas as well as the thermodynamic definition. The heat capacities agree within the estimated uncertainties.
 [doi:10.1063/1.3363609]

I. INTRODUCTION

Molecular dynamics (MD) solves equations of motion for a system of N particles depending on the external constraints applied to the system such as constant energy, pressure, or temperature. For realistic systems no analytical solutions exist, and numerical techniques are therefore used to solve the equations of motion. The step size (time step) of these algorithms decides how computationally demanding it is to propagate the system in time. The time step cannot be chosen arbitrarily large and is limited by the frequency of the fastest degree of freedom in the system. Typically, in molecular systems this is the fast covalent bond vibrations. Different techniques have been introduced to solve this problem,[2–4] one of them is constrained dynamics. Instead of following the fast degrees of freedom, these are eliminated by applying bond constraints to the system and hence a larger time step can be used. Constrained dynamics was originally proposed by Ryckaert *et al.*[4] and later refined by numerous groups.[5–16] The common denominator for all constrained dynamics algorithms is the application of corrections for conserving the constraints. The constraints deviate from their desired values in some of these algorithms, however, not only because of accumulating round-off errors, but also because the approximated expressions in the algorithms break the time symmetry. The broken time symmetry leads to algorithms that formally (i.e., no round-off errors) do not conserve the constraints.

It was shown recently by Toxvaerd *et al.*[14] that one can derive an algorithm to systems subject to constant energy or temperature (*NVE* or *NVT*), which conserves the bond constraints formally as well as numerically. This algorithm will be referred to as the TSCD (Time Symmetrical Central Difference) algorithm. In the present paper the results by Toxvaerd *et al.* will be used to derive an algorithm with the same properties for systems subject to constant pressure and enthalpy (*NPH*), as well as constant pressure and temperature (*NPT*).

Several equations of motion have been proposed to sample the *NPT* ensemble distribution for systems subject to constant pressure and temperature.[17–27] Kalibeva *et al.*[26] presented equations of motion using molecular scaling rather than scaling of the atomic positions. The advantage of molecular scaling is that one obtains a set of simple equations of motion which conserve the molecular bond constraints when the volume/coordinates are scaled to approach the external pressure.

The paper is organized as follows. The TSCD algorithm of Toxvaerd *et al.*[14] is reviewed in Sec. II, and the *NPT* equations of motion of Kalibaeva *et al.* are presented and modified in Sec. III. The corresponding *NPH* equations of motion and algorithms are presented in Sec. IV and extended in Sec. V to the *NPT* ensemble. Section VI presents results from simulations on a system of *ortho*-terphenyl (OTP) molecules,[28] validating that the algorithms integrate the equations of motion correctly in the *NPH* and *NPT* ensembles.

II. THE TSCD ALGORITHM

The equations of motion for N particles subject to the usual conditions in MD simulations can be derived from the

[a)]Present address: Department of Sciences, DNRF Centre "Glass and Time," IMFUFA, Roskilde University, Postbox 260, DK-4000 Roskilde, Denmark. Electronic mail: trond@ruc.dk.

modified principle of Hamilton.[29] This principle is still valid for systems where the dynamics is constrained to satisfy G holonomic relations among the coordinates

$$\psi^{\alpha}(\mathbf{r}^N,\mathbf{p}^N)=0, \tag{1}$$

$\alpha=1,\ldots,G$. Limiting the discussion to a special type of holonomic constraints, namely, bond constraints ($\psi^{\alpha}=\mathbf{r}_{ij}^2-c_{ij}^2$), the constrained equations of motion follow as[29]

$$\dot{\mathbf{r}}_i=\frac{\mathbf{p}_i}{m_i}, \tag{2}$$

$$\dot{\mathbf{p}}_i=\mathbf{F}_i+\mathbf{G}_i, \tag{3}$$

where

$$\mathbf{G}_i\equiv\sum_{\alpha=1}^{G}\lambda^{\alpha}\nabla_{\mathbf{r}_i}\psi^{\alpha}. \tag{4}$$

$\mathbf{G}_i$ denotes the constraint force and λ^{α} the Lagrange multiplier.

Toxvaerd *et al.*[14] reported that it is possible to use central difference techniques to integrate the equations of motion Eqs. (2) and (3), while maintaining the analytical properties such as time reversibility, conservation of energy, and conservation of the holonomic bond constraints.

The technique of derivation, as given by Toxvaerd *et al.*, is as follows. Applying central difference to Eqs. (2) and (3), respectively, the leap-frog (LF) algorithm appears

$$\mathbf{r}_i(t+h)=\mathbf{r}_i(t)+\frac{h}{m_i}\mathbf{p}_i(t+h/2), \tag{5}$$

$$\mathbf{p}_i(t+h/2)=\mathbf{p}_i(t-h/2)+h[\mathbf{F}_i(t)+\mathbf{G}_i(t)]. \tag{6}$$

The usual method for calculating λ^{α} is to require that the constraints are to be satisfied exactly at time $t+h$ in the discrete map generated by the integrator [Eqs. (5) and (6)]. Thus one inserts the integrator in the condition $\psi^{\alpha}(t+h)=0$ and calculates λ^{α} from this equation. This numerical algorithm is called SHAKE (Ref. 4) and its velocity Verlet counterpart, RATTLE.[6] SHAKE usually includes the iterative solution to the coupled quadratic equations in λ^{α}, while other algorithms iterate in a modified Chord-style such as MILC-SHAKE.[13,15]

Edberg *et al.*[8] followed at first sight a completely different path than Ryckaert *et al.*[4] (SHAKE) by inserting the analytical equations of motion [Eqs. (2) and (3)] in the second time derivative, $\ddot{\psi}^{\alpha}(t)=0$, and then applying a linearization to $\mathbf{v}_i^2(t)$ to be able to calculate λ^{α} by solving linear equations. The same path was also followed by Hess *et al.* in the LINCS algorithm.[11] This is analytically correct, but the linearization breaks the time symmetry and the algorithm is therefore time irreversible (the reader is referred to Ref. 14 for more details). The TSCD algorithm follows the path taken by Edberg *et al.*, but additionally adds the requirement of time reversal symmetry, for instance, by

$$\mathbf{v}_i^2(t)=\frac{\mathbf{v}_i^2(t-h/2)+\mathbf{v}_i^2(t+h/2)}{2}. \tag{7}$$

Obeying time symmetry results in coupled quadratic equations in λ^{α}, as in the SHAKE algorithm. It was proven by Toxvaerd *et al.* that the TSCD algorithm (method 0) has the following property:

$$\mathbf{r}_{ij}^2(t+h)=2\mathbf{r}_{ij}^2(t)-\mathbf{r}_{ij}^2(t-h), \tag{8}$$

which states that the bond length is *exactly* conserved in the discrete integrator map. From a numerical perspective the formula Eq. (8) is unstable and can be corrected by adding different correction terms, to the zeroth-order term, in the quadratic equations in λ^{α}. One possible way of eliminating this instability is method 1a of the TSCD algorithm, which results in the formula

$$\mathbf{r}_{ij}^2(t+h)=c_{ij}^2, \tag{9}$$

where c_{ij} is the desired bond length. This method (method 1a) is therefore *equivalent* to the SHAKE algorithm, and thus bridges the missing link between the most commonly used constraint algorithm, SHAKE, and the analytical approach introduced by Edberg *et al.*

The numerical correction terms vanish when the discrete dynamics do not contain round-off errors. Thus the present authors identify the TSCD algorithm (method 0) as the most basic algorithm, while the SHAKE algorithm is a convenient way of eliminating numerical instability. The choice of eliminating numerical instability is however not unique, and the reader is referred to method 1 or method 2 of Toxvaerd *et al.*[14] as alternatives.

The Trotter factorization can also be used to integrate Eqs. (2) and (3). For a particular factorization and neglecting that λ^{α} is a phase space function, the RATTLE algorithm appears.[6,30] This algorithm is the isochronous version of the TSCD algorithm and thus preserves the constraints formally [Eq. (8)], when calculating λ^{α} utilizing the prescription given by Toxvaerd *et al.*, i.e., obeying time symmetry. Thus it may be concluded from this kind of reasoning that it is indeed correct to consider the Lagrange multipliers as simple numbers when applying the Trotter factorization.[30]

The hidden constraint, $\dot{\psi}^{\alpha}(t)=0$, is also conserved in the TSCD algorithm, given that

$$\dot{\psi}^{\alpha}(t-h/2)=\frac{\mathbf{r}_{ij}(t)+\mathbf{r}_{ij}(t-h)}{2}\cdot\frac{\mathbf{r}_{ij}(t)-\mathbf{r}_{ij}(t-h)}{h}$$

$$=\frac{1}{2h}[\mathbf{r}_{ij}^2(t)-\mathbf{r}_{ij}^2(t-h)]=0. \tag{10}$$

The form of Eq. (10) is consistent with the philosophy of the LF algorithm. The last equality of Eq. (10) is valid for the TSCD algorithm since the bond constraints are exactly conserved.

The existence of a shadow Hamiltonian in the constrained LF dynamics was proven numerically by Toxvaerd *et al.*[14] This indicates that the TSCD algorithm (and therefore also SHAKE) is the exact time(-reversible) evolution of some extended perturbed Hamiltonian.

III. DECOUPLING THE *NPT* EQUATIONS OF MOTION

We follow the notation of Kalibaeva *et al.*,[26] according to which, $\mathbf{r}_{\mu i}$, $\mathbf{p}_{\mu i}$, and $m_{\mu i}$ denote the position, momentum,

and mass of atom i, inside molecule μ. The center of mass position, momentum, force, and mass of molecule μ are denoted by, respectively, $\mathbf{R}_\mu$, $\mathbf{P}_\mu$, $\mathbf{F}_\mu$, and M_μ. V, P_V, and M_V are, respectively, the volume, volume momentum, and the volume inertial factor. η, P_S, and M_S are the position, momenta, and inertial factor belonging to the thermostat for the N particle system. An additional thermostat has position, momenta, and inertial factor, respectively, ξ, P_ξ, and M_ξ. P and T are the instantaneous molecular pressure and kinetic atomic temperature defined in Eqs. (20) and (21). k_B is the Boltzmann constant and $g=3N-G-3$ is the number of degrees of freedom in the system.

We consider a system composed of N atoms grouped into n molecules, each molecule containing n_μ atoms ($\mu=1,\ldots,n$), i.e., $N=\Sigma_{\mu=1}^{n} n_\mu$. The system is subject to $G=\Sigma_{\mu=1}^{n} k_\mu$ holonomic molecular constraints of the type $\psi_\mu^\alpha(\mathbf{r}_{\mu 1},\ldots,\mathbf{r}_{\mu n_\mu})=0$, where $\alpha=1,\ldots,k_\mu$ and k_μ is the number of constraints applied to molecule μ. It is assumed that the following identity holds $\Sigma_{i=1}^{n_\mu}\nabla_{\mathbf{r}_{\mu i}}\psi_\mu^\alpha=0$, i.e., that the constraint functions are invariant under a translation of the coordinate system. The equations of motion of Kalibaeva *et al.*[26] for constant pressure and temperature (*NPT*) are

$$\dot{\mathbf{r}}_{\mu i}=\frac{\mathbf{p}_{\mu i}}{m_{\mu i}}+\frac{P_V}{3VM_V}\mathbf{R}_\mu, \tag{11}$$

$$\dot{\mathbf{p}}_{\mu i}=\mathbf{F}_{\mu i}+\mathbf{G}_{\mu i}-\frac{P_S}{M_S}\mathbf{p}_{\mu i}-\frac{m_{\mu i}P_V}{M_\mu 3VM_V}\mathbf{P}_\mu, \tag{12}$$

$$\dot{V}=\frac{P_V}{M_V}, \tag{13}$$

$$\dot{P}_V=[P-P_{\mathrm{EXT}}]-\frac{P_\xi}{M_\xi}P_V, \tag{14}$$

$$\dot{\eta}=\frac{P_S}{M_S}, \tag{15}$$

$$\dot{P}_S=gk_B[T-T_{\mathrm{EXT}}], \tag{16}$$

$$\dot{\xi}=\frac{P_\xi}{M_\xi}, \tag{17}$$

$$\dot{P}_\xi=\frac{P_V^2}{M_V}-k_BT_{\mathrm{EXT}}, \tag{18}$$

where

$$\mathbf{G}_{\mu i}\equiv\sum_{\alpha=1}^{k_\mu}\lambda_\mu^\alpha\nabla_{\mathbf{r}_{\mu i}}\psi_\mu^\alpha, \tag{19}$$

$$P\equiv\frac{1}{3V}\sum_{\mu=1}^{n}\left[\frac{\mathbf{P}_\mu^2}{M_\mu}+\mathbf{F}_\mu\cdot\mathbf{R}_\mu\right]=\frac{1}{3V}\left[\sum_{\mu=1}^{n}\frac{\mathbf{P}_\mu^2}{M_\mu}+\sum_{\alpha<\beta}^{n}\mathbf{F}_{\alpha\beta}\cdot\mathbf{R}_{\alpha\beta}\right], \tag{20}$$

$$T\equiv\frac{1}{gk_B}\left[\sum_{\mu=1}^{n}\sum_{i=1}^{n_\mu}\frac{\mathbf{p}_{\mu i}^2}{m_{\mu i}}\right]. \tag{21}$$

We modify these equations of motion by eliminating (ξ,P_ξ) from the extended phase space. The modified equations of motion become

$$\dot{\mathbf{r}}_{\mu i}=\frac{\mathbf{p}_{\mu i}}{m_{\mu i}}+\frac{P_V}{3VM_V}\mathbf{R}_\mu, \tag{22}$$

$$\dot{\mathbf{p}}_{\mu i}=\mathbf{F}_{\mu i}+\mathbf{G}_{\mu i}-\frac{P_S}{M_S}\mathbf{p}_{\mu i}-\frac{m_{\mu i}P_V}{M_\mu 3VM_V}\mathbf{P}_\mu, \tag{23}$$

$$\dot{V}=\frac{P_V}{M_V}, \tag{24}$$

$$\dot{P}_V=P-P_{\mathrm{EXT}}, \tag{25}$$

$$\dot{\eta}=\frac{P_S}{M_S}, \tag{26}$$

$$\dot{P}_S=gk_B[T-T_{\mathrm{EXT}}]. \tag{27}$$

These modified equations of motion are a combination of molecular scaling and one single Nosé–Hoover thermostat,[7,18,19] and as such inherit the deficiencies in the ergodic sampling from the Nosé–Hoover algorithm for stiff harmonic systems;[31] however, for many realistic and commonly used systems, this is of no concern.

In an isolated system the equations of motion demand $\mathbf{R}=\mathbf{P}=0$ in order to sample the correct *NPT* distribution.[32,33] In most MD simulations this is performed unconditionally. Equation (26) is a redundant but useful equation for calculating the conserved quantities.

The holonomic molecular constraints applied to the system are not affected by the scaling of the coordinates since

$$\dot{\psi}_\mu^\alpha=\sum_{i=1}^{n_\mu}\nabla_{\mathbf{r}_{\mu i}}\psi_\mu^\alpha\dot{\mathbf{r}}_{\mu i}=\sum_{i=1}^{n_\mu}\nabla_{\mathbf{r}_{\mu i}}\psi_\mu^\alpha\frac{\mathbf{p}_{\mu i}}{m_{\mu i}}=0. \tag{28}$$

It should be noted that the (first) definition of the instantaneous molecular pressure depends on the choice of the coordinate system used via the molecular virial. This choice can never influence how the integrating algorithm scans the phase space, and the usual expression given in the last equality of Eq. (20) should be used. The center of mass of each molecule is directly available in the algorithm, and calculating the molecular virial can easily be accomplished during the (pair) force calculation without any noticeable slowdown.

The task at hand is to integrate the equations of motion, Eqs. (22)–(27). We take the center of mass, $\mathbf{R}_\mu$ and $\mathbf{P}_\mu$, as dynamical variables of the system and thus need to derive their corresponding equations of motion. Using their definitions and $\Sigma_{i=1}^{n_\mu}\mathbf{G}_{\mu i}=\Sigma_{\alpha=1}^{k_\mu}\lambda_\mu^\alpha\Sigma_{i=1}^{n_\mu}\nabla_{\mathbf{r}_{\mu i}}\psi_\mu^\alpha=0$, it follows that

$$\dot{\mathbf{R}}_\mu=\frac{\mathbf{P}_\mu}{M_\mu}+\frac{P_V}{3VM_V}\mathbf{R}_\mu, \tag{29}$$

$$\dot{\mathbf{P}}_\mu = \mathbf{F}_\mu - \left[\frac{P_S}{M_S} + \frac{P_V}{3VM_V}\right]\mathbf{P}_\mu. \tag{30}$$

Thus we need to solve the system of coupled differential equations (22), (23), (29), and (30) and the regulator variables (24)–(27). We can ease this integration by (almost) decoupling equations (22) and (23) from the remaining differential equations. This was suggested by Marry *et al.*,[34] however their method is to derive algorithms using the Trotter factorization. In this article we derive integrators purely from central difference techniques.

The decoupled equations of motion are now derived. Two new variables are defined

$$\tilde{\mathbf{r}}_{\mu i} = \mathbf{r}_{\mu i} - \mathbf{R}_\mu, \tag{31}$$

$$\tilde{\mathbf{p}}_{\mu i} = \mathbf{p}_{\mu i} - \frac{m_{\mu i}}{M_\mu}\mathbf{P}_\mu, \tag{32}$$

$\tilde{\mathbf{r}}_{\mu i}$ and $\tilde{\mathbf{p}}_{\mu i}$ are the Galilean transformation of the position and momentum of atom i to the center of mass coordinate system of molecule μ. Substituting these definitions into Eqs. (22) and (23) and using Eqs. (29) and (30), it follows after some reduction that

$$\dot{\tilde{\mathbf{r}}}_{\mu i} = \frac{\tilde{\mathbf{p}}_{\mu i}}{m_{\mu i}}, \tag{33}$$

$$\dot{\tilde{\mathbf{p}}}_{\mu i} = \mathbf{F}_{\mu i} + \mathbf{G}_{\mu i} - \frac{m_{\mu i}}{M_\mu}\mathbf{F}_\mu - \frac{P_S}{M_S}\tilde{\mathbf{p}}_{\mu i}. \tag{34}$$

The entire system of differential equations to be solved becomes

$$\dot{\tilde{\mathbf{r}}}_{\mu i} = \frac{\tilde{\mathbf{p}}_{\mu i}}{m_{\mu i}}, \tag{35}$$

$$\dot{\tilde{\mathbf{p}}}_{\mu i} = \mathbf{F}_{\mu i} + \mathbf{G}_{\mu i} - \frac{m_{\mu i}}{M_\mu}\mathbf{F}_\mu - \frac{P_S}{M_S}\tilde{\mathbf{p}}_{\mu i}, \tag{36}$$

$$\dot{\mathbf{R}}_\mu = \frac{\mathbf{P}_\mu}{M_\mu} + \frac{P_V}{3VM_V}\mathbf{R}_\mu, \tag{37}$$

$$\dot{\mathbf{p}}_\mu = \mathbf{F}_\mu - \left[\frac{P_S}{M_S} + \frac{P_V}{3VM_V}\right]\mathbf{P}_\mu, \tag{38}$$

$$\dot{V} = \frac{P_V}{M_V}, \tag{39}$$

$$\dot{P}_V = P - P_{\mathrm{EXT}}, \tag{40}$$

$$\dot{\eta} = \frac{P_S}{M_S}, \tag{41}$$

$$\dot{P}_S = gk_B[T - T_{\mathrm{EXT}}]. \tag{42}$$

The above equations of motion are subject to two additional holonomic constraints

$$\tilde{\mathbf{R}} = 0, \tag{43}$$

$$\tilde{\mathbf{P}} = 0. \tag{44}$$

Numerical errors accumulate, and it is therefore necessary to correct for this behavior by, i.e., a transformation $\tilde{\mathbf{r}}_{\mu i} = \tilde{\mathbf{r}}_{\mu i} - \tilde{\mathbf{R}}_\mu$. The correction scales linearly and is as such consistent with the integrating algorithm.

IV. ALGORITHMS FOR THE *NPH* ENSEMBLE

Taking $(\eta, P_S) = (0,0)$ in Eqs. (35)–(42) one arrives at equations of motion for molecular constrained systems subject to constant enthalpy and pressure (*NPH*)

$$\dot{\tilde{\mathbf{r}}}_{\mu i} = \frac{\tilde{\mathbf{p}}_{\mu i}}{m_{\mu i}}, \tag{45}$$

$$\dot{\tilde{\mathbf{p}}}_{\mu i} = \mathbf{F}_{\mu i} + \mathbf{G}_{\mu i} - \frac{m_{\mu i}}{M_\mu}\mathbf{F}_\mu, \tag{46}$$

$$\dot{\mathbf{R}}_\mu = \frac{\mathbf{P}_\mu}{M_\mu} + \frac{P_V}{3VM_V}\mathbf{R}_\mu, \tag{47}$$

$$\dot{\mathbf{P}}_\mu = \mathbf{F}_\mu - \frac{P_V}{3VM_V}\mathbf{P}_\mu, \tag{48}$$

$$\dot{V} = \frac{P_V}{M_V}, \tag{49}$$

$$\dot{P}_V = P - P_{\mathrm{EXT}}. \tag{50}$$

The *NPH* ensemble has a long history going back to Andersen,[17] and Ryckaert *et al.* extended this ensemble to systems with molecular scaling and constraints.[7] The *NPH* ensemble was first investigated numerically by Haile *et al.*[35] It should be noted that the requirement of $\mathbf{R}=\mathbf{P}=0$ is still valid in an isolated system for these equations of motion.

We can use the *NPH* ensemble as another test of the validity of the integrating method, i.e., the central difference technique, and the ensemble provides an alternative way for extracting and comparing thermodynamic properties obtained at constant pressure.

We now limit the discussion to holonomic molecular constraints of the type

$$\psi_\mu^\alpha(\mathbf{r}_{\mu i}, \mathbf{r}_{\mu j}) = \mathbf{r}_{\mu ij}^2 - c_{\mu ij}^2 = 0, \tag{51}$$

i.e., holonomic molecular bond constraints. The constraint force is then $\mathbf{G}_{\mu i} = \Sigma \pm \lambda_\mu^\alpha \mathbf{r}_{\mu ij}$. Thus, before we are able to integrate Eqs. (45) and (46), we must find a suitable (discrete) expression for λ_μ^α. The constraint force can be expressed equivalently as $\mathbf{G}_{\mu i} = \Sigma \pm \lambda_\mu^\alpha \tilde{\mathbf{r}}_{\mu ij}$. This is also true for the holonomic bond constraint Eq. (51) and thus for the second derivative with respect to time.

If we now integrate Eqs. (45) and (46) using a central difference with pivot $t+h/2$ and t, respectively, i.e., the LF algorithm, then the discrete determination of λ^α in the *NPH* case can be taken to be equivalent with the (*NVE*) TSCD algorithm. The non-relative coordinates in the expressions of

Toxvaerd *et al.* are merely replaced with the internal motion coordinates. The same is valid for the *NPT* algorithm with a slight modification (see Sec. V). We therefore present no investigation of the conservation of the bond constraints in the *NPH* and *NPT* ensembles, as these will be equivalent to the *NVE* and *NVT* case, respectively, and as such conserved.

The problem is thus to solve the remaining equations of motion for the center of mass movement and the variables representing the barostat. We now define two new scaled variables

$$\mathbf{Q}_\mu \equiv \frac{1}{V^{1/3}}\mathbf{R}_\mu, \tag{52}$$

$$\mathbf{S}_\mu \equiv V^{1/3}\mathbf{P}_\mu. \tag{53}$$

By defining this transformation the equations of motion are transformed to their virtual canonical form (see Ryckaert *et al.*[7]). Substituting Eqs. (52) and (53) into Eqs. (47) and (48) and reducing, it follows that:[7]

$$\dot{\mathbf{Q}}_\mu = \frac{1}{V^{2/3}M_\mu}\mathbf{S}_\mu, \tag{54}$$

$$\dot{\mathbf{S}}_\mu = V^{1/3}\mathbf{F}_\mu. \tag{55}$$

We can solve these equations of motion using a time symmetrical central difference with the pivot $t+h/2$ and t in Eqs. (54) and (55), respectively. In that way, we arrive at the following algorithm:

$$\mathbf{Q}_\mu(t+h) = \mathbf{Q}_\mu(t) + h\frac{\mathbf{S}_\mu(t+h/2)}{M_\mu\left[\frac{V(t+h)+V(t)}{2}\right]^{2/3}}, \tag{56}$$

$$\mathbf{S}_\mu(t+h/2) = \mathbf{S}_\mu(t-h/2) + h\mathbf{F}_\mu(t)V^{1/3}(t). \tag{57}$$

Applying the same technique to V and P_V leads to

$$V(t+h) = V(t) + \frac{h}{M_V}P_V(t+h/2), \tag{58}$$

$$P_V(t+h/2) = P_V(t-h/2) + h[P(t) - P_{\mathrm{EXT}}]. \tag{59}$$

The molecular pressure $P(t)$ can be calculated from

$$P(t) = \frac{1}{3V^{2/3}(t)}\left[\sum_{\mu=1}^{n}\frac{\mathbf{S}_\mu^2(t+h/2)+\mathbf{S}_\mu^2(t-h/2)}{2M_\mu V(t)} + \sum_{\alpha<\beta}^{n}(\mathbf{F}_{\alpha\beta}(t)\cdot\mathbf{Q}_{\alpha\beta}(t))\right], \tag{60}$$

which follows by substituting the definition (52) and (53) into Eq. (20) and using a time-symmetrical mean around t.

The modified equations of motion in the virtual canonical form are symplectic (no constraints), in this context meaning that they preserve the invariant measure of Hamiltonian dynamics,[29] i.e., the phase space volume-element. As Ishida *et al.*[36] emphasize, it is the conservation of the volume-element and not the true symplectic condition that provides stability. The algorithm presented in this section (rewritten to its isochronous form) derived from the central difference technique does indeed give $|\mathbf{J}|=1$, where $\mathbf{J}$ is the Jacobian matrix of the discrete coordinate transformation. The discrete algorithm is therefore symplectic and time reversible.

V. ALGORITHMS FOR THE NPT ENSEMBLE

An algorithm for Eqs. (35)–(42) is now presented. The technique used in the Sec. IV is once again applied, and the derived algorithm degenerates to the *NPH* algorithm in the case of $(\eta, P_S)=(0,0)$.

Equations (35) and (36) are the usual Nosé–Hoover equations of motion, and the integration (and determination of λ^α) almost follows the *NVT* integration algorithm developed by Toxvaerd[37] and applied in the TSCD algorithm[14] for the constrained *NVT* ensemble. We continue from the transformation defined in Sec. IV, and the slight change in the method of Toxvaerd[37] is illustrated when integrating the similar equation of motion

$$\dot{\mathbf{S}}_\mu = V^{1/3}\mathbf{F}_\mu - \frac{P_S}{M_S}\mathbf{S}_\mu. \tag{61}$$

We rewrite Eq. (61) following the approach introduced by Jang *et al.*[38]

$$\frac{d}{dt}\left[\exp\left[\int_0^t\frac{P_S(t')}{M_S}dt'\right]\mathbf{S}_\mu\right] = \exp\left[\int_0^t\frac{P_S(t')}{M_S}dt'\right]V^{1/3}\mathbf{F}_\mu. \tag{62}$$

Jang *et al.* have modified the *NVT* algorithm developed by Toxvaerd.[37] The above transformation carries a close resemblance to the normal form procedure suggested by Legoll *et al.*,[39] i.e., rewriting the equations of motion to a form where the divergence vanishes; however, we do not in this procedure define new variables, we merely use the differential equation as a stepping stone. Applying a central difference with pivot t and using a mid/end-point approximation for the integral, it follows that[38]

$$\mathbf{S}_\mu(t+h/2) = \frac{1}{S_{\mu+}(t)}\{S_{\mu-}(t)\mathbf{S}_\mu(t-h/2) + hV^{1/3}(t)\mathbf{F}_\mu(t)\}, \tag{63}$$

where

$$S_{\mu+}(t) \equiv \exp\left[\frac{hP_S(t)}{2M_S}\right], \tag{64}$$

$$S_{\mu-}(t) \equiv \exp\left[-\frac{hP_S(t)}{2M_S}\right]. \tag{65}$$

The change in the *NVT* algorithm of Toxvaerd[37] is given by Eqs. (63)–(65), where $(1\pm x)$ is replaced with $\exp(\pm x)$, and thus has no impact on the constraint algorithm of Toxvaerd *et al.*[14] The remaining *NPT* equations of motion, η and P_S, are integrated as

$$\eta(t+h/2) = \eta(t-h/2) + \frac{h}{M_S} P_V(t), \tag{66}$$

$$P_S(t+h) = P_S(t) + hgk_B[T(t+h/2) - T_{\mathrm{EXT}}], \tag{67}$$

where the integration of $\mathbf{Q}_\mu$, V, and P_V remains the same as in Eqs. (56), (58), and (59), respectively.

The invariant phase space measure of non-Hamiltonian dynamics relates to the phase space compressibility and constitutes a modified volume-element.[33] The presented *NPT* algorithm is exactly measure preserving[40]

$$|\mathbf{J}| = \frac{e^{3(N+n)\eta(0)}}{e^{3(N+n)\eta(h)}} = \frac{d\mathbf{\Gamma}(h)}{d\mathbf{\Gamma}(0)}. \tag{68}$$

The number $(N+n)$ in Eq. (68) is due to the fact that we have taken the relative coordinates and the center of mass of the molecules as the independent variables of the system.

The *NVT* algorithm of Toxvaerd[37] is not exactly measure preserving[40,41] due to the approximation $\exp(\pm x) \approx (1 \pm x)$. This approximation only introduces accumulating errors from terms of $O(h^3)$, and we could not establish any significant difference between the two approaches. We have thus derived exactly measure preserving integrators using central difference techniques for the *NPH* and *NPT* ensembles.

Sergi *et al.*[31] proposed an extended symplectic notation for deriving equations of motion that conserve a quasi-Hamiltonian

$$\dot{x}_i = \sum_j B_{ij} \frac{\partial H}{\partial x_j}, \tag{69}$$

where B_{ij} is an antisymmetric matrix. The antisymmetry of B_{ij} ensures the conservation of the quasi-Hamiltonian H, and thus it is possible to manipulate the phase space compressibility, $\kappa = \Sigma_{i,j}(B_{ij}/\partial x_i)(\partial H/\partial x_j)$, for the desired phase space sampling. Sergi *et al.* also observed that it is possible to formulate many extended equations of motion, such as the Nosé–Hoover equations, in this notation [Eq. (69)]. The quasi-Hamiltonian can often be written as a sum of terms, i.e., $H = K + U + P_V^2/2M_V + \cdots = \Sigma_\alpha H(\alpha)$, and these terms induce a natural splitting of the Liouville operator[42]

$$\begin{aligned}\hat{L} \equiv \sum_i \dot{x}_i \frac{\partial}{\partial x_i} &= \sum_{i,j} B_{ij} \frac{\partial H}{\partial x_j} \frac{\partial}{\partial x_i} \\ &= \sum_{\alpha,i,j} B_{ij} \frac{\partial H(\alpha)}{\partial x_j} \frac{\partial}{\partial x_i} \\ &= \sum_\alpha \hat{L}(\alpha).\end{aligned} \tag{70}$$

As shown by Ezra[42] this fact ensures that any algorithm derived from the Trotter factorization technique, according to the splitting in Eq. (70), is exactly measure preserving. The splitting in Eq. (70) is however not necessarily the obvious splitting of $\hat{L} \equiv \Sigma_i \dot{x}_i(\partial/\partial x_i)$. This method thus provides another way of designing measure preserving algorithms alongside the above mentioned technique.

VI. TESTING THE ALGORITHMS BY COMPUTER SIMULATIONS

The *NPH* and *NPT* algorithms were tested on a system of $n=320$ OTP molecules with periodic boundaries in the x-, y-, and z-direction. The OTP molecule is locked as a rigid isosceles triangle using three connected holonomic bond constraints.[28] We performed simulations[43] ($h=0.0025$) at two different state points, in both ensembles with a low and high pressure: $P_{\mathrm{EXT}}=1.00$ bar and 3.50 kbar. The intermolecular interactions are defined via the truncated pair potential

$$u(r) = u_{12-6}(r) - \left(\frac{du_{12-6}(r)}{dr}\right)_{r=r_c} \cdot (r - r_c) - u_{12-6}(r_c), \tag{71}$$

with no discontinuity in the potential or the force at the truncation distance, $r_c=2.5$. The algorithms were all started by initially setting the momenta $P_V(t_0)=P_S(t_0)=0$. The inertial factors M_V and M_S were defined as

$$M_V \equiv P_{\mathrm{EXT}} \tau_V^2, \tag{72}$$

$$M_S \equiv \frac{gk_B T_{\mathrm{EXT}} \tau_S^2}{4\pi^2}, \tag{73}$$

where τ_V and τ_S are parameters to be specified.

Now consider the *NPH* ensemble and the two different state points with a low and high pressure. In the low pressure state the algorithm was unstable for $\tau_V < 0.0080$. A value of $\tau_V < 0.22$ resulted in fast volume/pressure fluctuations, with some "higher" values of τ_V producing dead zone for the barostat. The corresponding numbers for the high pressure state were $\tau_V < 0.00019$ and $\tau_V < 0.0045$. The *NVT* ensemble was used to estimate the value for $\tau_S = 0.20$.[37] The estimated values of τ_S and τ_V from the *NVT/NPH* ensemble were used in the *NPT* ensemble, where the behavior of changing the individual variables reflected the state in their sole ensemble. This is not surprising since our modified equations of motion have no coupling between the individual regulator variables. We observed fluctuating behavior of the energy/enthalpy in the *NPH* ensemble, due to the real constant of motion being the sum of the enthalpy and the kinetic energy of the "piston"—in the limit $N \to \infty$, the piston term disappears and so do the fluctuations. Nevertheless as shown below, the heat capacities calculated from the fluctuations agree, and the algorithm samples states in the *NPH* ensemble.

The momenta P_V and P_S should both be Gaussian distributed[26] in the *NPT* ensemble with $f_{NPT}(P_x) \propto \exp[-P_x^2/2k_B T_{\mathrm{EXT}} M_x]$. In the *NPH* ensemble, we cannot directly write down the probability distribution of the P_V momenta due to the δ-function expression. The *NPH* ensemble is, however, similar to the *NVE* ensemble (i.e., the inverse Laplace transform of the *NPT* ensemble[17]). We thus expect that the P_V variable is Gaussian distributed, and from the *NPT* probability distribution a reasonable estimate seems to be $f_{NPH}(P_V) \propto \exp[-P_V^2/2k_B\langle T(t)\rangle M_V]$. *QQ*-plots comparing the empirical momenta with the theoretical distributions showed superb agreement, in both ensembles, from simulations in a low pressure state. The same conclusion was

TABLE I. The mean molecular pressure $\langle P(t)\rangle$ given by Eq. (20) after n_{ts} time steps with $h=0.0025$ using single-precision accuracy. The mean value is taken over all n_{ts} time steps. The low and high pressure state had ($\tau_V=0.22, \tau_S=0.20$) and ($\tau_V=0.0045, \tau_S=0.20$), respectively.

n_{ts} time steps	*NPH*	*NPT*
P_{EXT} (1.00 bar, 300K)	$1.360\ 22\times10^{-3}$	$1.360\ 22\times10^{-3}$
100 000	$1.392\ 41\times10^{-3}$	$1.359\ 18\times10^{-3}$
250 000	$1.352\ 97\times10^{-3}$	$1.367\ 98\times10^{-3}$
1 000 000	$1.362\ 59\times10^{-3}$	$1.366\ 41\times10^{-3}$
5 000 000	$1.360\ 25\times10^{-3}$	$1.360\ 68\times10^{-3}$
P_{EXT} (3.50 kbar, 450K)	4.760 74	4.760 74
100 000	4.760 76	4.760 73
250 000	4.760 73	4.760 73
500 000	4.760 74	4.760 74

reached from a *QQ*-plot comparing the Maxwell–Boltzmann distribution with the momentum of a given molecule in a high pressure state.

The mean quantity of Eqs. (20) and (21) should equal P_{EXT} and T_{EXT}, respectively, at the used finite precision. A thermostat and a barostat share the common property, that one cannot pulse heat into the system or change the volume of the system at a faster rate than the system can come to equilibrium after a perturbation. The system relaxation time toward equilibrium, after a perturbation, depends on the investigated state point. Changes in the volume (scaling) suffer from longer relaxation times than changes in the kinetic energy, as the particles need to relocate. Thus one must carefully select the inertial factor, M_V, depending on the investigated state point. It therefore took longer simulation time to establish a correct mean pressure in our investigated low pressure state than in the high pressure state, due to a longer relaxation time.

In Table I, the mean pressure, given by Eq. (20), after n_{ts} time steps (the mean is taken over all n_{ts} time steps), is reported for the low and high pressure state, in both ensembles. It is seen that the mean molecular pressure $\langle P(t)\rangle$ in the *NPH* ensemble and the low pressure state is within (in single-precision) 0.5% of the target pressure, P_{EXT}, after just $n_{ts}=250\,000$ time steps. In the *NPT* ensemble, the corresponding number is 0.6%. The high pressure state has established the correct target pressure within the same number of time steps in the *NPH* and *NPT* ensembles. The "exact" number of time steps is sensitive to the particular choice of the factor M_V and negative mean pressure can also be experienced.[34]

An important thermodynamic property of a system is the heat capacity. As another test of our algorithms we calculated the isobaric heat capacity for the OTP system at two different state points, using different methods. In this manner we cannot only compare different ways of measuring the same property, but also between different ensembles. When comparing results between different ensembles, we must remember that we are simulating finite systems, and only in the thermodynamic limit ($N\to\infty$) all ensembles become equivalent.

TABLE II. The specific isobaric heat capacity, $C_P^*=C_P/nk_B$, calculated at a low and high pressure state in the *NPH* and *NPT* ensembles, where C_P(FLUCT) is given by Eqs. (74) and (75), respectively. C_P(SLOPE) is calculated from $C_P\equiv(\partial H/\partial T)_{N,P}$, where a number of state points surrounding the chosen statepoint were fitted to a straight line.

Ensemble	*NPH*	*NPT*
	(1.00 bar, 300 K)	(1.00 bar, 300 K)
C_P^*(FLUCT)	11.43 ± 0.06	11.51 ± 0.01
C_P^*(SLOPE)		11.52 ± 0.01
	(3.50 kbar, 450 K)	(3.50 kbar, 450 K)
C_P^*(FLUCT)	10.25 ± 0.07	10.26 ± 0.01
C_P^*(SLOPE)		10.29 ± 0.03

The isobaric heat capacity may be calculated from the thermodynamic definition $C_P\equiv(\partial H/\partial T)_{N,P}$ and from fluctuations in observable quantities. The expressions for the fluctuations are[1,44]

$$C_P^{\infty}(NPH)=k_B\left[\frac{2}{g}-\frac{\sigma^2(K)}{\langle K\rangle^2}\right]^{-1}, \tag{74}$$

$$C_P(NPT)=\sigma^2(H)[k_BT_{EXT}^2]^{-1}, \tag{75}$$

where $H=E+P_{EXT}V$ is the enthalpy and K is the kinetic energy. We estimated the heat capacities in the *NPT* ensemble from a simulation over 500×10^6 time steps. Every 5×10^6 time steps—for a total of 125×10^6 time steps—the instantaneous configuration from the *NPT* simulation was written to the disk, providing different starting configurations for measuring the heat capacity in the *NPH* ensemble. Each independent configuration was simulated 5×10^6 time steps in the *NPH* ensemble and used to estimate the heat capacity.

In Table II, the specific isobaric heat capacity, $C_P^*=C_P/nk_B$, at the two different state points, in the *NPH* and *NPT* ensembles, is reported. The heat capacities agree within the estimated uncertainties (recall the ensemble difference). We note a very weak temperature dependence of the isobaric heat capacity over the investigated temperature range, where the lower temperatures gave a slightly higher heat capacity.

VII. SUMMARY

Time reversible integrators arising from central difference techniques are widely used for solving equations of motion. Recently it was shown by Toxvaerd *et al.*[14] that one can derive central difference integrators for the *NVE* and *NVT* ensembles with holonomic bond constraints, which conserve the bond constraints formally (i.e., no round-off errors) as well as numerically. This was achieved first by obeying time symmetry and then, second, adding a correction term to the solved equations for the Lagrange multiplier, which is formally zero.

We have in this article derived central difference integrators for the constrained *NPH* and *NPT* ensembles using a modification of the equations of motion of Kalibaeva *et al.*[26] The bond constraints are satisfied at each discrete point in time, using the results of Toxvaerd *et al.* thus creating a unified theory for practical simulations of constrained dynamics in the most common ensembles. The algorithms are

time reversible and preserve the invariant phase space measure.

We performed simulations in a high pressure state in both ensembles, establishing the correct pressure and temperature within a few hundred thousand time steps. The low pressure state takes longer simulation time as expected due to a longer relaxation time. Isobaric heat capacities were compared between ensembles and using different methods. The results agreed within the estimated uncertainties and validate, that the algorithm correctly samples states in the *NPH* and *NPT* ensembles.

The equations of motion of Kalibaeva *et al.* apply a molecular scaling and a molecular description of the pressure. It is an open problem of how to extend the presented methods to the more complex equations of motion with atomiclike scaling.[23–25]

ACKNOWLEDGMENTS

The center for viscous liquid dynamics *Glass and Time* is sponsored by the Danish National Research Foundation's (DNRF). The authors gratefully acknowledge the referees for insightful comments. One of the authors (T.I.) would like to thank Jacob Louis Marott, Statistician, the Copenhagen City Heart Study.

[1] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Oxford University Press, Oxford, 2003).
[2] D. Frenkel and B. Smit, *Understanding Molecular Simulation* (Academic, New York, 2002).
[3] M. Tuckerman, B. J. Berne, and G. J. Martyna, J. Chem. Phys. **97**, 1990 (1992).
[4] J. P. Ryckaert, G. Ciccotti, and H. J. C. Berendsen, J. Comput. Phys. **23**, 327 (1977).
[5] G. Ciccotti, M. Ferrario, and J. P. Ryckaert, Mol. Phys. **47**, 1253 (1982).
[6] H. C. Andersen, J. Comput. Phys. **52**, 24 (1983).
[7] J. P. Ryckaert and G. Ciccotti, J. Chem. Phys. **78**, 7368 (1983).
[8] R. Edberg, D. J. Evans, and G. P. Morriss, J. Chem. Phys. **84**, 6933 (1986).
[9] S. G. Lambrakos, J. P. Boris, E. S. Oran, I. Chandrasekhar, and M. Nagumo, J. Comput. Phys. **85**, 473 (1989).
[10] B. J. Leimkuhler and R. D. Skeel, J. Comput. Phys. **112**, 117 (1994).
[11] B. Hess, H. Bekker, H. J. C. Berendsen, and J. G. E. M. Fraaije, J. Comput. Chem. **18**, 1463 (1997).
[12] V. Kräutler, W. F. van Gunsteren, and P. H. Hünenberger, J. Comput. Chem. **22**, 501 (2001).
[13] A. G. Bailey, C. P. Lowe, and A. P. Sutton, J. Comput. Phys. **227**, 8949 (2008).
[14] S. Toxvaerd, O. J. Heilmann, T. Ingebrigtsen, T. B. Schrøder, and J. C. Dyre, J. Chem. Phys. **131**, 064102 (2009).
[15] A. G. Bailey and C. P. Lowe, J. Comput. Chem. **30**, 2485 (2009).
[16] P. Eastman and V. S. Pande, J. Chem. Theory Comput. **6**, 434 (2010).
[17] H. C. Andersen, J. Chem. Phys. **72**, 2384 (1980).
[18] S. Nosé, J. Chem. Phys. **81**, 511 (1984).
[19] W. G. Hoover, Phys. Rev. A **31**, 1695 (1985).
[20] G. J. Martyna, D. J. Tobias, and M. L. Klein, J. Chem. Phys. **101**, 4177 (1994).
[21] S. Melchionna and G. Ciccotti, J. Chem. Phys. **106**, 195 (1997).
[22] A. Sergi, M. Ferrario, and D. Costa, Mol. Phys. **97**, 825 (1999).
[23] G. R. Kneller and T. Mülders, Phys. Rev. E **54**, 6825 (1996).
[24] T. Mülders, S. Toxvaerd, and G. R. Kneller, Phys. Rev. E **58**, 6766 (1998).
[25] G. Ciccotti, G. J. Martyna, S. Melchionna, and M. E. Tuckerman, J. Phys. Chem. B **105**, 6710 (2001).
[26] G. Kalibaeva, M. Ferrario, and G. Ciccotti, Mol. Phys. **101**, 765 (2003).
[27] H. A. Stern, J. Comput. Chem. **25**, 749 (2004).
[28] L. J. Lewis and G. Wahnström, Phys. Rev. E **50**, 3865 (1994).
[29] H. Goldstein, C. Poole, and J. Safko, *Classical Mechanics* (Addison-Wesley, Reading, MA, 2002).
[30] A. Sergi, Phys. Rev. E **69**, 021109 (2004).
[31] A. Sergi and M. Ferrario, Phys. Rev. E **64**, 056125 (2001).
[32] K. Cho, J. D. Joannopoulos, and L. Kleinman, Phys. Rev. E **47**, 3145 (1993).
[33] M. E. Tuckerman, Y. Liu, G. Ciccotti, and G. J. Martyna, J. Chem. Phys. **115**, 1678 (2001).
[34] V. Marry and G. Ciccotti, J. Comput. Phys. **222**, 428 (2007).
[35] J. M. Haile and H. W. Graben, J. Chem. Phys. **73**, 2412 (1980).
[36] H. Ishida and A. Kidera, J. Chem. Phys. **109**, 3276 (1998).
[37] S. Toxvaerd, Mol. Phys. **72**, 159 (1991).
[38] S. Jang and G. A. Voth, J. Chem. Phys. **107**, 9514 (1997).
[39] F. Legoll and R. Monneau, J. Chem. Phys. **117**, 10452 (2002).
[40] T. Yamamoto, J. Chem. Phys. **124**, 217101 (2006).
[41] S. Toxvaerd, Phys. Rev. E **47**, 343 (1993).
[42] G. S. Ezra, J. Chem. Phys. **125**, 034104 (2006).
[43] Simulations are performed with length, energy and mass chosen, respectively, in units of σ_{AA}, ϵ_{AA}. and m_A. Thus the units of time, temperature and pressure are $\sqrt{m_A\sigma_{AA}^2/\epsilon_{AA}}$, ϵ_{AA}/k_B, and $\epsilon_{AA}/\sigma_{AA}^3$. The unit for heat capacity becomes k_B.
[44] J. R. Ray, H. W. Graben, and J. M. Haile, Nuovo Cimento B **64**, 191 (1981).