

An Efficient Integrator Scheme for Sampling the (Quantum) Isobaric–Isothermal Ensemble in (Path Integral) Molecular Dynamics Simulations

Published as part of Journal of Chemical Theory and Computation *special issue* “Developments of Theoretical and Computational Chemistry Methods in Asia”.

Weihao Liang[§], Sihan Wang[§], Cong Wang[§], Weizhou Wang, Xinchen She, Chongbin Wang, Jiushu Shao, and Jian Liu^{*}



Cite This: <https://doi.org/10.1021/acs.jctc.5c00573>



Read Online

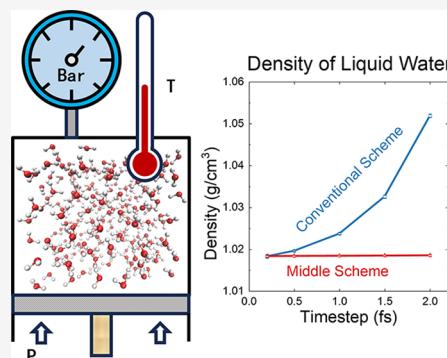
ACCESS |

Metrics & More

Article Recommendations

Supporting Information

ABSTRACT: Because most chemical or biological experiments are performed under conditions of controlled pressure and temperature, it is important to simulate the isobaric–isothermal ensemble at the atomic level to reveal the microscopic mechanism. By extending our efficient configuration sampling approach for the canonical ensemble, we propose a unified “middle” scheme to sample the coordinate (configuration) and volume distribution, which can accurately simulate either classical or quantum isobaric–isothermal processes. Various barostats and thermostats can be employed in the unified “middle” scheme for simulating real molecular systems with or without holonomic constraints. In particular, we demonstrate the recommended “middle” scheme by employing the Martyna-Tuckerman-Tobias-Klein barostat and stochastic cell-rescaling barostat, with the Langevin thermostat, in molecular simulation packages (DL_POLY, AMBER, GROMACS, and so forth). Benchmark numerical tests show that, without additional numerical effort, the “middle” scheme is competent in increasing the time interval by a factor of 5 ~ 10 to achieve the same accuracy of converged results for most thermodynamic properties in (path integral) molecular dynamics simulations.



1. INTRODUCTION

Most real experiments or processes occur under conditions of constant pressure and temperature. Recent progress on pressure-driven phenomena suggests that exotic chemical, electronic, magnetic, and structural properties can be unlocked in molecular solids and condensed matter materials,^{1–5} which include superconductivity, quantum Hall effects, abnormal phase transition, to name but a few. The isobaric–isothermal ensemble, where the number of particles (N), external pressure (P), and temperature (T) are fixed, has been extensively used for studying real complex molecular systems via (path integral) molecular dynamics (MD) in physics, chemistry, biology, astronomy, and environmental science.^{6–8} The time interval Δt is the most fundamental factor that determines both the accuracy and efficiency of the MD simulation of the isobaric–isothermal ensemble. A larger time interval produces more efficient sampling of the ensemble, but often decreases the accuracy of or even breaks down the simulation. It is important yet challenging to design robust integrators/algorithms for the isobaric–isothermal ensemble, in which larger time intervals are allowed for considerably improving the sampling efficiency while maintaining the same accuracy.

The purpose of the paper is to present a simple, efficient, and robust integrator scheme for sampling the isobaric–isothermal (constant-NPT) ensemble for real molecular systems by MD or path integral MD simulations. The outline of the paper is as follows. Section 2 begins by briefly reviewing the theory on the isobaric–isothermal ensemble in classical mechanics as well as in quantum mechanics. It then introduces the strategy for rationally constructing the MD integrator by employing the barostatting methods via MD or Monte Carlo (MC).^{7,9–30} In Section 3, the optimized “middle” scheme is numerically implemented and compared to conventional NPT algorithms used in molecular simulation packages. Benchmark test examples include the Lennard-Jones liquid, liquid *para*-Hydrogen, and liquid water. Finally, conclusion remarks are presented in Section 4.

Received: April 10, 2025

Revised: June 4, 2025

Accepted: June 5, 2025

2. THEORY AND METHODOLOGIES

We consider a general real molecular system that includes N_{atom} atoms in the three-dimensional space, which is described by the (time-independent) Hamiltonian of standard Cartesian form

$$H(\mathbf{x}, \mathbf{p}) = \mathbf{p}^T \mathbf{M}^{-1} \mathbf{p} / 2 + U(\mathbf{x}) = K + U(\mathbf{x}) \quad (1)$$

where \mathbf{M} is the diagonal ‘mass matrix’ with elements $\{m_j; 1 \leq j \leq 3N_{\text{atom}}\}$, \mathbf{x} and \mathbf{p} are the $3N_{\text{atom}}$ -dimensional coordinate and momentum vectors, respectively, $U(\mathbf{x})$ is the potential energy, and K is the kinetic energy.

2.1. NPT Ensemble in Classical Statistical Mechanics.

In classical statistical mechanics, $\tilde{H}_{\text{ins}} = PV + H$ defines the instantaneous enthalpy and the average of any physical property of the isobaric–isothermal ensemble in the thermodynamic limit reads

$$\langle B \rangle_{\text{NPT}}^{\text{CM}} = \frac{I_N}{V_0 Z_{\text{NPT}}} \int_0^\infty dV \int_{D(V)} d\mathbf{x} \int d\mathbf{p} \exp[-\beta(PV + H(\mathbf{x}, \mathbf{p}))] B(\mathbf{x}, \mathbf{p}; V) \quad (2)$$

where the partition function Z_{NPT} of the isobaric–isothermal ensemble is

$$Z_{\text{NPT}} = \frac{I_N}{V_0} \int_0^\infty dV \int_{D(V)} d\mathbf{x} \int d\mathbf{p} \exp[-\beta(PV + H(\mathbf{x}, \mathbf{p}))] \quad (3)$$

Here, $D(V)$ is the coordinate space of the volume, V , of the system, $I_N = (2\pi\hbar)^{-3N_{\text{atom}}} c_N$ is a product of the normalization factor for the (coordinate-momentum) phase space integral³¹ and that for indistinguishable particles, $\beta = 1/k_B T$ with the Boltzmann constant k_B , and V_0 is the reference volume. In eq 2, $B(\mathbf{x}, \mathbf{p}; V)$ is the estimator for the corresponding property B . For instance, the virial expression of the internal pressure in classical statistical mechanics is

$$P_{\text{int}} = \frac{1}{dV} \left(\mathbf{p}^T \mathbf{M}^{-1} \mathbf{p} - \mathbf{x}^T \frac{\partial U}{\partial \mathbf{x}} \right) - \frac{\partial U}{\partial V} \quad (4)$$

where d is the dimensionality of the space, in which the molecular system lies. (In the paper, $d = 3$ as we study the molecular system in the three-dimensional space.) The last term $\frac{\partial U}{\partial V}$ on the right hand side (RHS) of eq 4, which represents the explicit volume dependence of the potential energy function, is usually zero for molecular systems. (The last term of eq 4 should be ignored with caution when more general cases are studied.) As will be discussed in Section 2.4, the evaluation of the virial expression of the internal pressure is often requested in MD barostats. More discussion on the estimator of other physical properties is presented in Section S1 in the Supporting Information.

As the partition function of the canonical (constant-NVT) ensemble is $Z_{\text{NVT}} = I_N \int_{D(V)} d\mathbf{x} \int d\mathbf{p} \exp[-\beta H(\mathbf{x}, \mathbf{p})]$, we have the relation

$$\langle \hat{B} \rangle_{\text{NPT}}^{\text{QM}} = \frac{\lim_{L \rightarrow \infty} \int_0^\infty dV e^{-\beta PV} \int_{D(V)} d\mathbf{x}_1 \int_{D(V)} d\mathbf{x}_2 \cdots \int_{D(V)} d\mathbf{x}_L \exp\{-\beta U_{\text{eff}}(\mathbf{x}_1, \dots, \mathbf{x}_L)\} \tilde{B}(\mathbf{x}_1, \dots, \mathbf{x}_L; V)}{\lim_{L \rightarrow \infty} \int_0^\infty dV e^{-\beta PV} \int_{D(V)} d\mathbf{x}_1 \int_{D(V)} d\mathbf{x}_2 \cdots \int_{D(V)} d\mathbf{x}_L \exp\{-\beta U_{\text{eff}}(\mathbf{x}_1, \dots, \mathbf{x}_L)\}} \quad (9)$$

where L is the number of path integral beads (replica),

$$Z_{\text{NPT}} = \frac{1}{V_0} \int_0^\infty dV \exp[-\beta PV] Z_{\text{NVT}}(N, V, T) \quad (5)$$

Equation 5 can also be expressed as

$$e^{-\beta G(N, P, T)} = \int_0^\infty dV e^{-\beta(PV + F(N, V, T))} \quad (6)$$

where the Gibbs free energy is $G(N, P, T) = -\ln Z_{\text{NPT}}/\beta$ and the Helmholtz free energy is $F(N, V, T) = -\ln Z_{\text{NVT}}/\beta$. Note that physical properties such as density, isothermal compressibility, isobaric heat capacity, thermal expansion coefficient, and so forth can be estimated only in the isobaric–isothermal ensemble rather than in the canonical ensemble. The accurate sampling of the integrand of the RHS of eq 5, i.e., the partial probability distribution function of the volume, is the essential task of the barostat for controlled pressure. Equation 2 indicates that in classical statistical mechanics the accurate sampling of the joint distribution of the volume (V) and coordinate (\mathbf{x}) variables is the key for sampling the isobaric–isothermal ensemble, because most physical properties of interest depend only on the coordinate (\mathbf{x}) and volume (V). Note that the Maxwell momentum distribution is simply a Gaussian function that is independent of the coordinate and volume. It is straightforward to first use MD to obtain only the accurate joint distribution of the volume (V) and coordinate (\mathbf{x}) and then employ MC to faithfully obtain the independent momentum distribution, which produces accurate evaluation of any thermodynamic properties that involve both coordinate/volume and momentum variables in classical statistical mechanics.

2.2. NPT Ensemble in Quantum Statistical Mechanics.

Equation 5 can also be justified in quantum statistical mechanics in the thermodynamic limit, where the canonical partition function is defined as

$$Z_{\text{NVT}} = \text{Tr}[\exp(-\beta \hat{H})] = \int_{D(V)} d\mathbf{x} \langle \mathbf{x} | \exp(-\beta \hat{H}) | \mathbf{x} \rangle \quad (7)$$

and the corresponding average physical property of eq 2 becomes

$$\langle \hat{B} \rangle_{\text{NPT}}^{\text{QM}} = \frac{1}{V_0 Z_{\text{NPT}}} \int_0^\infty dV \exp[-\beta PV] \times \int_{D(V)} d\mathbf{x} \langle \mathbf{x} | \exp(-\beta \hat{H}) \hat{B} | \mathbf{x} \rangle \quad (8)$$

After decomposing $\exp(-\beta \hat{H})$ into L equal parts $[\exp(-\beta \hat{H}/L)]^L$, substituting the resolution of identity $\hat{1} \equiv \int_{D(V)} d\mathbf{x}_j |\mathbf{x}_j\rangle \langle \mathbf{x}_j|$ into each neighbor parts, and applying the symmetric Trotter factorization $\exp(-\beta \hat{H}/L) = \exp(-\beta \hat{U}/2L) \exp(-\beta \hat{K}/L) \exp(-\beta \hat{U}/2L)$, the sampling of the probability distribution function of eq 8 can often be accomplished by imaginary time path integral,^{32–35} namely,

$$U_{\text{eff}}(\mathbf{x}_1, \dots, \mathbf{x}_L) = \frac{L}{2\beta^2\hbar^2} \sum_{i=1}^L [(\mathbf{x}_{i+1} - \mathbf{x}_i)^T \mathbf{M} (\mathbf{x}_{i+1} - \mathbf{x}_i)] + \frac{1}{L} \sum_{i=1}^L U(\mathbf{x}_i) \quad (10)$$

is the effective potential of path integral beads, the estimator $\tilde{B}(\mathbf{x}_1, \dots, \mathbf{x}_L; V)$ for any operator \hat{B} is a function of only the coordinate vector of path integral beads and the volume, $I_{\text{PI}} = (L/2\pi\beta\hbar^2)^{3N_{\text{atom}}L/2} |\mathbf{M}|^{L/2}$ is a normalization factor, and the denominator of the RHS of eq 9 is related to the path integral expression of the partition function Z_{NPT} of the isobaric–isothermal ensemble.

When the staging transformation^{36,37} is used, the transformation of coordinates and that of forces read

$$\begin{aligned} \xi_1 &= \mathbf{x}_1 \\ \xi_i &= \mathbf{x}_i - \frac{(i-1)\mathbf{x}_{i+1} + \mathbf{x}_1}{i} \quad (i = 2, \dots, L) \end{aligned} \quad (11)$$

and

$$\begin{aligned} \frac{\partial \phi}{\partial \xi_1} &= \frac{1}{L} \sum_{i=1}^L U'(\mathbf{x}_i) \\ \frac{\partial \phi}{\partial \xi_i} &= \frac{\partial \phi}{\partial \mathbf{x}_i} + \frac{i-2}{i-1} \frac{\partial \phi}{\partial \xi_{i-1}} \quad (i = 2, \dots, L) \end{aligned} \quad (12)$$

$$\langle \hat{B} \rangle_{\text{NPT}}^{\text{QM}} = \frac{\lim_{L \rightarrow \infty} \int_0^\infty dV \prod_{k=1}^L (\int d\mathbf{p}_k) \prod_{j=1}^L (\int_{D(V)} d\mathbf{x}_j) e^{-\beta[PV + H_{\text{eff}}(\mathbf{p}_1, \dots, \mathbf{p}_L; \mathbf{x}_1, \dots, \mathbf{x}_L)]} \tilde{B}(\mathbf{x}_1, \dots, \mathbf{x}_L; V)}{\lim_{L \rightarrow \infty} \int_0^\infty dV \prod_{k=1}^L (\int d\mathbf{p}_k) \prod_{j=1}^L (\int_{D(V)} d\mathbf{x}_j) e^{-\beta[PV + H_{\text{eff}}(\mathbf{p}_1, \dots, \mathbf{p}_L; \mathbf{x}_1, \dots, \mathbf{x}_L)]}} \quad (15)$$

where the effective Hamiltonian reads

$$H_{\text{eff}}(\mathbf{x}_1, \dots, \mathbf{x}_L; \mathbf{p}_1, \dots, \mathbf{p}_L) = \sum_{i=1}^L \frac{1}{2} \mathbf{p}_i^T \tilde{\mathbf{M}}_i^{-1} \mathbf{p}_i + U_{\text{eff}}(\mathbf{x}_1, \dots, \mathbf{x}_L) \quad (16)$$

In eq 16, the fictitious mass $\tilde{\mathbf{M}}$ can be chosen as

$$\begin{aligned} \tilde{\mathbf{M}}_1 &= \mathbf{M} \\ \tilde{\mathbf{M}}_i &= \frac{i}{i-1} \mathbf{M} \quad (i = 2, \dots, L) \end{aligned} \quad (17)$$

(see Section S1 for more discussion). In quantum statistical mechanics, the accurate sampling of the joint distribution of the volume and the coordinate vector of path integral beads of the isobaric–isothermal ensemble (of eq 9 or eq 15) is then the only key for evaluating all thermodynamic properties.

where ϕ is defined as

$$\phi(\xi_1, \dots, \xi_L) = \frac{1}{L} \sum_{i=1}^L U[\mathbf{x}_i(\xi_1, \dots, \xi_L)]$$

The effective potential (of eq 10) after the staging transformation becomes

$$U_{\text{eff}}(\xi_1, \dots, \xi_L) = \sum_{i=1}^L \frac{1}{2} \omega_i^2 \xi_i^T \tilde{\mathbf{M}}_i \xi_i + \frac{1}{L} \sum_{i=1}^L U[\mathbf{x}_i(\xi_1, \dots, \xi_L)] \quad (13)$$

The corresponding mass matrix of the staging transformation is

$$\begin{aligned} \tilde{\mathbf{M}}_1 &= 0 \\ \tilde{\mathbf{M}}_i &= \frac{i}{i-1} \mathbf{M} \quad (i = 2, \dots, L) \end{aligned} \quad (14)$$

The integral of eq 9 can be performed by MD^{37–40} after introducing the artificial momentum and mass of the path integral bead and substituting the multidimensional Gaussian integral identity

$$(2\pi)^{-3N_{\text{atom}}/2} |\tilde{\mathbf{M}}_j|^{-1/2} \int d\mathbf{p}_j \exp\left(-\frac{1}{2} \mathbf{p}_j^T \tilde{\mathbf{M}}_j^{-1} \mathbf{p}_j\right) = 1$$

that is,

In Section S1C, we explicitly describe the estimator $B(\mathbf{x}, \mathbf{p}; V)$ of eq 2 in classical statistical mechanics as well as the estimator $\tilde{B}(\mathbf{x}_1, \dots, \mathbf{x}_L; V)$ of either eq 9 or eq 15 in quantum statistical mechanics, when the density, enthalpy, isobaric heat capacity, isothermal compressibility, and thermal expansion coefficient are simulated.

3.3. Barostat for the Volume Distribution. Either MC or MD methods^{7,9–30} can be used for the barostat for sampling the volume distribution of eq 2 or eq 15. Note that the volume distribution of eq 2 or eq 15 includes not only $e^{-\beta PV}$ but also the potential energy in the confined volume (e.g., the box with periodic boundary conditions when the bulk system is investigated).

When MD is employed for the barostat, one approach is to substitute a classical resolution of identity, i.e., replacing a unit by the one-dimensional Gaussian integral $[\beta/(2\pi W)]^{1/2} \int dp_V \exp[-\beta p_V^2/(2W)] = 1$ on the RHS of eq 2, which yields

$$\langle B \rangle_{\text{NPT}}^{\text{CM}} = \frac{\int dp_V \int_0^\infty dV \int_{D(V)} d\mathbf{x} \int d\mathbf{p} \exp\left[-\beta\left(\frac{p_V^2}{2W} + PV + H(\mathbf{x}, \mathbf{p})\right)\right] B(\mathbf{x}, \mathbf{p}; V)}{\int dp_V \int_0^\infty dV \int_{D(V)} d\mathbf{x} \int d\mathbf{p} \exp\left[-\beta\left(\frac{p_V^2}{2W} + PV + H(\mathbf{x}, \mathbf{p})\right)\right]} \quad (18)$$

where p_V and W are the “momentum” variable and “mass” of the fictitious “piston” introduced for controlling the volume

fluctuation and then sampling the volume distribution by MD trajectories. Similarly, eq 15 becomes

$$\langle \hat{B} \rangle_{\text{NPT}}^{\text{QM}} = \frac{\lim_{L \rightarrow \infty} \int dp_V \int_0^\infty dV \prod_{k=1}^L (\int dp_k) \prod_{j=1}^L (\int_{D(V)} dx_j) e^{-\beta[\frac{p_V^2}{2W} + PV + H_{\text{eff}}(p_1, \dots, p_L; x_1, \dots, x_L)]} \tilde{B}(x_1, \dots, x_L; V)}{\lim_{L \rightarrow \infty} \int dp_V \int_0^\infty dV \prod_{k=1}^L (\int dp_k) \prod_{j=1}^L (\int_{D(V)} dx_j) e^{-\beta[\frac{p_V^2}{2W} + PV + H_{\text{eff}}(p_1, \dots, p_L; x_1, \dots, x_L)]}} \quad (19)$$

In addition to the extended-system approach discussed above, another approach is to weakly couple the system to a pressure bath without the fictitious “piston” momentum, which includes the Berendsen barostat⁴¹ and stochastic cell-rescaling (SCR) barostat.⁴² During each pressure control step, the volume is scaled by a factor determined by the pressure difference between the bath and the system. While the Berendsen barostat fails to yield the correct volume distribution,⁴³ the SCR barostat is capable of faithfully producing the isobaric–isothermal ensemble^{42,44} and will be used in the paper.

Because PIMD essentially employs MD of the effective Hamiltonian system of eq 16 while using different estimators for physical properties in quantum statistical mechanics from those in classical statistical mechanics, both MD and PIMD can in principle share the same integrator for the isobaric–isothermal ensemble. In the following we only use MD integrators to demonstrate the procedure for simplicity. It is straightforward to implement the same strategy for PIMD integrators.

2.4. Molecular Dynamics Integrators. Numerical MD integrators for a finite time interval Δt often consist of a step for updating the coordinate $x(t + \Delta t) \leftarrow x(t) + M^{-1}p(t)\Delta t$, that for updating the momentum $p(t + \Delta t) \leftarrow p(t) - U'(x)\Delta t$, that for the thermostat for controlled temperature, and that for the barostat for controlled external pressure. Use $e^{\mathcal{L}_x \Delta t}$, $e^{\mathcal{L}_p \Delta t}$, $e^{\mathcal{L}_T \Delta t}$, and $e^{\mathcal{L}_{\text{Bar}} \Delta t}$ to represent the phase space propagators for the four steps, respectively. Here, \mathcal{L}_{Bar} is the Kolmogorov operator for the barostat, which is defined in the volume space via MC or on “extended” phase space (V, p_V) via MD, \mathcal{L}_T is the Kolmogorov operator for the thermostat, and \mathcal{L}_x and \mathcal{L}_p are the relevant Kolmogorov operators on coordinate-momentum phase space (x, p) related to the system Hamiltonian, $\mathcal{L}_x \rho = -p^T M^{-1} \frac{\partial \rho}{\partial x}$ and $\mathcal{L}_p \rho = \left(\frac{\partial U}{\partial x}\right)^T \frac{\partial \rho}{\partial p}$, with ρ being the corresponding probability distribution function of the ensemble.

Note that the characteristic time scale of the evolution of coordinate-momentum variables (x, p) of the molecular system is usually several orders of magnitude smaller than that of the fluctuation of the volume variable, V . It is expected that an efficient algorithm for the isobaric–isothermal (constant-NPT) ensemble should be rationally constructed from that for the canonical (constant-NVT) ensemble. Recently, we have shown that the unified middle thermostat scheme leads to efficient and robust (path integral) molecular dynamics algorithms for sampling (the coordinate space of) the canonical ensemble,^{37,45–50} regardless of whether the thermostat for controlled temperature is stochastic, deterministic, or mixed.^{37,45,49} For instance, it is straightforward in the unified middle thermostat scheme to derive the relation⁴⁶ between the algorithm on Langevin dynamics by Leimkuhler and Matthews⁵¹ and that by Grønbech-Jensen and Farago independently.⁵² Our work of ref 45 points out that only those reasonable thermostating methods, which faithfully

produce the Maxwell momentum distribution, should be considered in the unified middle thermostat scheme. For instance, special caution should be taken care of for the Nose-Hoover thermostat^{53–55} or the more advanced Nose-Hoover-Chain thermostat⁵⁶ when it is used in the unified middle thermostat scheme for sampling the canonical ensemble.⁴⁵ More importantly, the unified thermostat scheme with phase space evolution operators^{45–50} is capable of describing thermostating algorithms that lead to no corresponding differential equations in the limit of $\Delta t \rightarrow 0$.^{46,47} The middle thermostat scheme suggests $e^{\mathcal{L}_{\text{NVT}} \Delta t} = e^{\mathcal{L}_p \Delta t/2} e^{\mathcal{L}_x \Delta t/2} e^{\mathcal{L}_T \Delta t} e^{\mathcal{L}_x \Delta t/2} e^{\mathcal{L}_p \Delta t/2}$ based on the velocity-Verlet algorithm or $e^{\mathcal{L}_{\text{NVT}} \Delta t} = e^{\mathcal{L}_x \Delta t/2} e^{\mathcal{L}_T \Delta t} e^{\mathcal{L}_x \Delta t/2} e^{\mathcal{L}_p \Delta t}$ based on the leapfrog algorithm. The operations of the phase space propagators are in a right-to-left sequence throughout the paper. Because the thermostating step for a full time interval, $e^{\mathcal{L}_T \Delta t}$, is placed in the middle of the two updating steps of the coordinate for each half time interval $e^{\mathcal{L}_x \Delta t/2}$, it is denoted as the “middle” scheme.^{45–50} Similarly, conventional velocity-Verlet-based MD algorithms fall into the “side” thermostat scheme $e^{\mathcal{L}_{\text{NVT}} \Delta t} = e^{\mathcal{L}_T \Delta t/2} e^{\mathcal{L}_p \Delta t/2} e^{\mathcal{L}_x \Delta t} e^{\mathcal{L}_p \Delta t/2} e^{\mathcal{L}_T \Delta t/2}$, where a thermostating step for a half time interval $e^{\mathcal{L}_T \Delta t/2}$ is designed at both the beginning and end of the full MD integrator, or into the “end” or “beginning” thermostat scheme $e^{\mathcal{L}_{\text{NVT}} \Delta t} = e^{\mathcal{L}_T \Delta t} e^{\mathcal{L}_p \Delta t/2} e^{\mathcal{L}_x \Delta t}$ or $e^{\mathcal{L}_{\text{NVT}} \Delta t} = e^{\mathcal{L}_p \Delta t/2} e^{\mathcal{L}_x \Delta t} e^{\mathcal{L}_p \Delta t/2} e^{\mathcal{L}_T \Delta t}$, where the thermostating step for a full time interval, $e^{\mathcal{L}_T \Delta t}$, is applied at the beginning or end of the full MD integrator.^{45–50} It is trivial to show the leapfrog-based version of the “side”/“end”/“beginning” thermostat schemes that conventional MD algorithms employ.^{48,49}

Conventional MD algorithms often produce accurate sampling of the (trivial) momentum distribution but significant errors for sampling the coordinate distribution that is often much more important.^{45–49} The most prominent feature of the unified middle thermostat is that both the accuracy and efficiency of sampling are insensitive to the value of the thermostat parameter(s) in a broad range for a relatively large time interval.^{46–48} In comparison to the unified “middle” thermostat scheme, when conventional MD algorithms are used, the numerical error of sampling the coordinate distribution is significantly greater for a finite time interval, and the characteristic correlation time of sampling is rather sensitive to the choice of the thermostat parameter(s).^{46–48}

As demonstrated in refs 45 and 48, when the integrator $e^{\mathcal{L}_{\text{NVT}} \Delta t} = e^{\mathcal{L}_p \Delta t/2} e^{\mathcal{L}_x \Delta t/2} e^{\mathcal{L}_T \Delta t} e^{\mathcal{L}_x \Delta t/2} e^{\mathcal{L}_p \Delta t/2}$ based on the velocity-Verlet algorithm is used, we obtain the accurate coordinate marginal distribution as well as the accurate momentum marginal distribution at two positions: one is immediately after the first momentum-updating step for a half time interval $e^{\mathcal{L}_p \Delta t/2}$ and before $e^{\mathcal{L}_x \Delta t/2} e^{\mathcal{L}_T \Delta t} e^{\mathcal{L}_x \Delta t/2}$ is applied, and the other is immediately after $e^{\mathcal{L}_x \Delta t/2} e^{\mathcal{L}_T \Delta t} e^{\mathcal{L}_x \Delta t/2}$ and

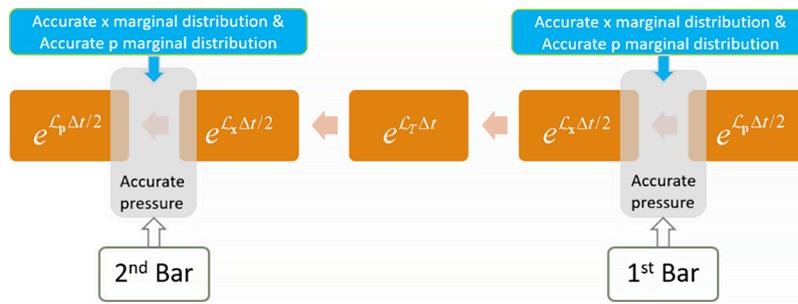


Figure 1. Schematic representation of the “middle” scheme for the isobaric–isothermal ensemble. When both the instantaneous kinetic energy and the virial term in the evaluation of the internal pressure are required, the barostat step should be applied at either of the two positions listed in the right-to-left sequence. When the evaluation of the force or potential is NOT required in the barostat step at the first position, the barostat step can also be applied at both positions to achieve a symmetric decomposition $e^{\mathcal{L}_{\text{Bar}(2)}\Delta t/2}e^{\mathcal{L}_x\Delta t/2}e^{\mathcal{L}_T\Delta t}e^{\mathcal{L}_x\Delta t/2}e^{\mathcal{L}_{\text{Bar}(1)}\Delta t/2}$ (for higher accuracy) before and after the momentum-updating step(s).

before the second momentum-updating step for a half time interval $e^{\mathcal{L}_p\Delta t/2}$. The two positions are demonstrated in Figure 1. Specifically, in the harmonic limit, both the coordinate marginal distribution and the momentum marginal distribution are *exact* at either of the two positions in Figure 1. Similarly, as discussed in ref 48, when the integrator $e^{\mathcal{L}_{\text{NVT}}\Delta t} = e^{\mathcal{L}_x\Delta t/2}e^{\mathcal{L}_T\Delta t}e^{\mathcal{L}_x\Delta t/2}e^{\mathcal{L}_p\Delta t}$ based on the leapfrog algorithm is employed, we achieve the accurate coordinate marginal distribution as well as the accurate momentum marginal distribution at two positions: one is immediately after the momentum-updating step $e^{\mathcal{L}_p\Delta t}$ and before $e^{\mathcal{L}_x\Delta t/2}e^{\mathcal{L}_T\Delta t}e^{\mathcal{L}_x\Delta t/2}$ is applied, and the other is immediately after $e^{\mathcal{L}_x\Delta t/2}e^{\mathcal{L}_T\Delta t}e^{\mathcal{L}_x\Delta t/2}$.

It is anticipated that the efficient MD integrator for a finite time interval Δt for the isobaric–isothermal ensemble should approach the “middle” thermostat scheme for the canonical ensemble, when the barostat effectively vanishes. Actually, various integrators for the isobaric–isothermal ensemble can be designed to reach the “middle” thermostat scheme in the canonical-ensemble limit. Because the barostat step updates the volume and then the potential or force, it should be applied when the coordinate (configuration) marginal distribution is accurate. That is, the phase space evolution operator for the barostat step should be applied before or after $e^{\mathcal{L}_x\Delta t/2}e^{\mathcal{L}_T\Delta t}e^{\mathcal{L}_x\Delta t/2}$. When the instantaneous kinetic energy is involved in the evaluation of the internal pressure in eq 4, the barostat step ought to be applied when the momentum marginal distribution is also accurate. This suggests that the barostat step should be designed at the first or second position of the right-to-left sequence as listed in Figure 1. The update of the force or/and the potential is only required at the second position for the barostat step before the momentum-updating step. If the evaluation of the force or potential is not requested in the barostat step at the first position, then the barostat step can also be applied at the first position such that a symmetric decomposition $e^{\mathcal{L}_{\text{Bar}(2)}\Delta t/2}e^{\mathcal{L}_x\Delta t/2}e^{\mathcal{L}_T\Delta t}e^{\mathcal{L}_x\Delta t/2}e^{\mathcal{L}_{\text{Bar}(1)}\Delta t/2}$ is achieved for higher accuracy before and after the momentum-updating step(s).

2.4.1. Integrators with the Monte Carlo Barostat. When MC is used for the barostat,^{6,9,10} the evaluation of the potential is always necessary in each move trial of the Metropolis algorithm for sampling the volume distribution. We then recommend that, either the “middle” scheme of

$$e^{\mathcal{L}_{\text{NPT}}^{\text{Middle}}\Delta t} = e^{\mathcal{L}_p\Delta t/2}e^{\mathcal{L}_{\text{Bar}}^{\text{MC}}\Delta t}e^{\mathcal{L}_x\Delta t/2}e^{\mathcal{L}_T\Delta t}e^{\mathcal{L}_x\Delta t/2}e^{\mathcal{L}_p\Delta t/2} \quad (20)$$

based on the velocity-Verlet algorithm or

$$e^{\mathcal{L}_{\text{NPT}}^{\text{Middle}}\Delta t} = e^{\mathcal{L}_{\text{Bar}}^{\text{MC}}\Delta t}e^{\mathcal{L}_x\Delta t/2}e^{\mathcal{L}_T\Delta t}e^{\mathcal{L}_x\Delta t/2}e^{\mathcal{L}_p\Delta t} \quad (21)$$

based on the leapfrog algorithm, should be used to sample the isobaric–isothermal ensemble when the MC barostat is used. Because the MC barostat does not employ the virial expression of the internal pressure (e.g., eq 4), it is convenient to employ the MC barostat for general cases, e.g., even for ab initio dynamics simulations of condensed phase molecular systems, where periodic boundary conditions are applied.

When it is straightforward to evaluate the virial expression for the internal pressure, MD barostats are more often used as they readily approach the NPT equilibrium, even when the initial density is far from the density of the isobaric–isothermal ensemble. In the following we focus on MD barostats.

2.4.2. Integrators with Molecular Dynamics Barostats. **2.4.2.1. Martyna-Tuckerman-Tobias-Klein Barostat.** When the MD barostat is used for the volume sampling under a constant external pressure condition, “extended” phase space variables (V, p_V) for the barostat and coordinate-momentum phase space variables (x, p) of the molecular system often propagate simultaneously. Equation 18 (or eq 19) indicates that the “force” for the update of p depends on both the volume V and the coordinate x during the propagation of the MD trajectory. As discussed in Section S2, the “middle” scheme of

$$e^{\mathcal{L}_{\text{NPT}}^{\text{Middle}}\Delta t} = e^{\mathcal{L}_p\Delta t/2}e^{\mathcal{L}_{\text{Bar}(2)}^{\text{MD}}\Delta t/2}e^{\mathcal{L}_x\Delta t/2}e^{\mathcal{L}_T\Delta t}e^{\mathcal{L}_x\Delta t/2}e^{\mathcal{L}_{\text{Bar}(1)}^{\text{MD}}\Delta t/2}e^{\mathcal{L}_p\Delta t/2} \quad (22)$$

based on the velocity-Verlet algorithm or

$$e^{\mathcal{L}_{\text{NPT}}^{\text{Middle}}\Delta t} = e^{\mathcal{L}_{\text{Bar}(2)}^{\text{MD}}\Delta t/2}e^{\mathcal{L}_x\Delta t/2}e^{\mathcal{L}_T\Delta t}e^{\mathcal{L}_x\Delta t/2}e^{\mathcal{L}_{\text{Bar}(1)}^{\text{MD}}\Delta t/2}e^{\mathcal{L}_p\Delta t} \quad (23)$$

based on the leapfrog algorithm is recommended for the isobaric–isothermal ensemble. The force is updated only in the step $e^{\mathcal{L}_{\text{Bar}(2)}^{\text{MD}}\Delta t/2}$ during each whole integration for a finite time interval Δt .

We use the Martyna–Tuckerman–Tobias–Klein (MTTK) barostat^{7,22–26,57} as an example. In this case, the first and second barostat steps of the right-to-left sequence in eq 22 or eq 23 are $e^{\mathcal{L}_{\text{Bar}(1)}^{\text{MD}}\Delta t} = e^{\mathcal{L}_p\Delta t}e^{\mathcal{L}_x\Delta t}e^{\mathcal{L}_V\Delta t}e^{\mathcal{L}_{p_e}\Delta t}e^{\mathcal{L}_p\Delta t}$ and $e^{\mathcal{L}_{\text{Bar}(2)}^{\text{MD}}\Delta t} = e^{\mathcal{L}_p\Delta t}e^{\mathcal{L}_{p_e}\Delta t}e^{\mathcal{L}_V\Delta t}e^{\mathcal{L}_x\Delta t}e^{\mathcal{L}_B\Delta t}$. Apparently, $e^{\mathcal{L}_{\text{Bar}(1)}^{\text{MD}}\Delta t}$ and $e^{\mathcal{L}_{\text{Bar}(2)}^{\text{MD}}\Delta t}$ are symmetric to achieve accuracy of the full integrator for the

finite time interval Δt . The Kolmogorov operator for the MTTK barostat consists of five terms:

$$\mathcal{L}_x \rho = -\frac{p_e}{W} \nabla_x \cdot (\mathbf{x} \rho) \quad (24)$$

$$\mathcal{L}_p \rho = \left(1 + \frac{d}{N_f} \right) \frac{p_e}{W} \nabla_p \cdot (\mathbf{p} \rho) \quad (25)$$

$$\mathcal{L}_V \rho = -d \frac{p_e}{W} \frac{\partial}{\partial V} (V \rho) \quad (26)$$

$$\mathcal{L}_{p_e} \rho = -dV(P_{\text{int}} - P_{\text{ext}}) \frac{\partial \rho}{\partial p_e} - \frac{d}{N_f} \mathbf{p}^T \mathbf{M}^{-1} \mathbf{p} \frac{\partial \rho}{\partial p_e} \quad (27)$$

$$\mathcal{L}_B \rho = \gamma_{\text{Lang}}^V \frac{\partial}{\partial p_e} (p_e \rho) + \frac{\gamma_{\text{Lang}}^V W}{\beta} \frac{\partial^2 \rho}{\partial p_e^2} \quad (28)$$

Here, N_f is the total number of degrees of freedom of the system ($N_f \equiv d \cdot N_{\text{atom}} = 3N_{\text{atom}}$ in this paper). Moreover, the mass W of the fictitious “piston” is sufficiently large to ensure the convergence of estimation of physical properties of interest. Typically, this value is set to

$$W = (N_f + 1) k_B T \tau_b^2 \quad (29)$$

where τ_b is the relaxation time of the barostat, which is often set to 1 ps as suggested in ref 7. In eq 27 the internal pressure (P_{int}) estimator of the system is presented in eq 4.

In the MTTK barostat, the volume-updating step $e^{\mathcal{L}_V \Delta t}$ related to eq 26 does not request the update of the virial expression of the internal pressure (or the update of the force). As shown in Figure 2, it is required to update the force and

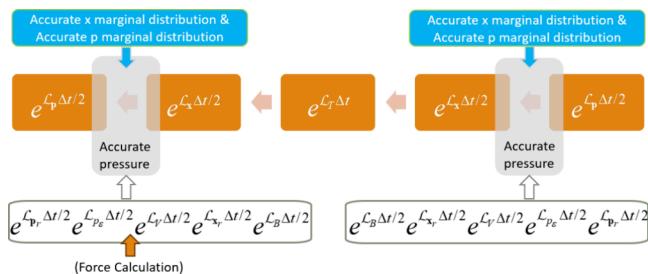


Figure 2. Schematic representation of the “middle” scheme with the MTTK barostat for the isobaric–isothermal ensemble. When the MTTK barostat step (for a half time interval $\Delta t/2$) is applied at the first position, the evaluation of the force or potential is NOT required. The MTTK barostat step should then be applied at the two positions listed in the right-to-left sequence. The update of the force or/and the potential is explicitly performed, before the piston-momentum-updating step of $e^{\mathcal{L}_{\text{Bar}}^{\text{MD}} \Delta t / 2}$, at the second position for the barostat step.

internal pressure only *once* after the operation $e^{\mathcal{L}_v \Delta t / 2} e^{\mathcal{L}_x \Delta t / 2} e^{\mathcal{L}_p \Delta t / 2}$ and before the operation $e^{\mathcal{L}_p \Delta t / 2} e^{\mathcal{L}_e \Delta t / 2}$ in the second MTTK barostat step in the whole integration for a finite time interval Δt .

It is easy to implement the MTTK barostat to PIMD.⁷ The internal pressure is an analogy of the classical form using the effective potential $U_{\text{eff}}(\mathbf{x}_1, \dots, \mathbf{x}_L)$, and the method employing the corresponding equations of motion is called the “all-mode scaled” method. Martyna et al. proposed another set of

equations of motion leading to the “reduced dynamics” method. Details of these two approaches are presented in Section S1B.

In Section S2 (which is based on our earlier investigations^{58,59}), we explicitly show that the “middle” scheme of eq 22 or eq 23 with the MTTK barostat is the optimally designed strategy by the analytical analysis and numerical investigations of a one-dimensional nanowire model and liquid water. In addition, as demonstrated in Section S2C, the recommended scheme of eq 22 or eq 23 has “equivalent” variants, of which the “middle-sinh” version employs the decomposition of Kolmogorov operators of the MTTK barostat, which is closely related to that of ref 23. Because eq 22 or eq 23 presents a clean-cut unified form for implementing other extended-system barostatting methods that have the same appropriate properties of the MTTK barostat, we use eq 22 or eq 23 throughout the paper.

24.2.2. Stochastic Cell Rescaling Barostat. In addition to extended-system barostatting methods, weak coupling methods without the fictitious “piston” momentum for the barostat can also be used in the “middle” scheme. Here, we use the SCR barostat as an example for illustration.

The update of the volume in the SCR barostat always requires the new evaluation of the virial expression of the internal pressure (or of the force), which is often the most time-consuming factor in the MD simulation. As suggested in Figure 3, the “middle” scheme with the SCR barostat can be realized with either

$$e^{\mathcal{L}_{\text{NPT}}^{\text{Middle}} \Delta t} = e^{\mathcal{L}_p \Delta t / 2} e^{\mathcal{L}_{\text{Bar}}^{\text{MD}} \Delta t} e^{\mathcal{L}_x \Delta t / 2} e^{\mathcal{L}_T \Delta t} e^{\mathcal{L}_x \Delta t / 2} e^{\mathcal{L}_p \Delta t / 2} \quad (30)$$

based on the velocity-Verlet algorithm or

$$e^{\mathcal{L}_{\text{NPT}}^{\text{Middle}} \Delta t} = e^{\mathcal{L}_{\text{Bar}}^{\text{MD}} \Delta t} e^{\mathcal{L}_x \Delta t / 2} e^{\mathcal{L}_T \Delta t} e^{\mathcal{L}_x \Delta t / 2} e^{\mathcal{L}_p \Delta t} \quad (31)$$

based on the leapfrog algorithm. The phase space evolution operator of the SCR barostat (for a finite time interval Δt) reads

$$e^{\mathcal{L}_{\text{Bar}}^{\text{MD}} \Delta t} = e^{\mathcal{L}_p \Delta t} e^{\mathcal{L}_x \Delta t} e^{\mathcal{L}_e \Delta t} \quad (32)$$

where ε is usually defined as $\varepsilon = \ln(V/V_0)$ with V_0 as the unit volume. In eq 32, the Kolmogorov operator, \mathcal{L}_e , is

$$\mathcal{L}_e \rho = -\partial_e \left[\frac{\kappa_T}{\tau_p} (P_{\text{int}} - P_{\text{ext}}) \right] \rho + \partial_e^2 \left[\frac{\kappa_T}{\tau_p \beta V} \rho \right] \quad (33)$$

where τ_p the characteristic time of the barostat, and κ_T the isothermal compressibility of the system. The value of κ_T in eq 33 is typically obtained from experimental data. When the experimental value of κ_T is not available, it can be estimated from primitive simulations or from the known value of a similar system. For instance, the value for ambient water (4.5×10^{-5} bar⁻¹) can be used for the simulation of a molecular system in water solution. The corresponding phase space evolution operator for eq 33 reads

$$e^{\mathcal{L}_e \Delta t} \rho(\mathbf{x}, \mathbf{p}; \varepsilon) = \rho(\mathbf{x}, \mathbf{p}; \varepsilon + \Delta \varepsilon) \quad (34)$$

where

$$\Delta \varepsilon \approx \frac{\kappa_T}{\tau_p} (P_{\text{int}} - P_{\text{ext}}) \Delta t + \sqrt{\frac{2\kappa_T \Delta t}{\beta \tau_p V}} \eta(t) \quad (35)$$

in which $\eta(t)$ is the standard Gaussian random number at a fixed time t (with zero mean $\langle \eta(t) \rangle = 0$ and unit deviation

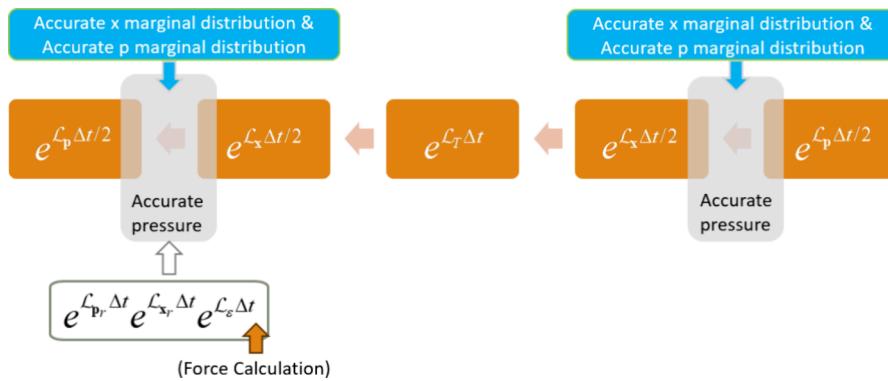


Figure 3. Schematic representation of the “middle” scheme with the SCR barostat for the isobaric–isothermal ensemble. The barostat step should be applied at only the second position in the right-to-left sequence. The update of the force or/and the potential is explicitly performed before the volume-updating step $e^{\mathcal{L}_e \Delta t}$ of $e^{\mathcal{L}_{\text{Bar}}^{\text{MD}} \Delta t}$, at the second position for the barostat step.

$\langle (\eta(t))^2 \rangle = 1$). After the operation of $e^{\mathcal{L}_e \Delta t}$, the other two phase space evolution operators are

$$e^{\mathcal{L}_{x_r} \Delta t} \rho(\mathbf{x}, \mathbf{p}; \varepsilon) = \rho(\exp(\Delta \varepsilon/d)\mathbf{x}, \mathbf{p}; \varepsilon) \quad (36)$$

$$e^{\mathcal{L}_{p_r} \Delta t} \rho(\mathbf{x}, \mathbf{p}; \varepsilon) = \rho(\mathbf{x}, \exp(-\Delta \varepsilon/d)\mathbf{p}; \varepsilon) \quad (37)$$

respectively. The time scale of the operation $e^{\mathcal{L}_{x_r} \Delta t}$ is same as that of the operation $e^{\mathcal{L}_e \Delta t}$, although they are well-separable from the time scale of the atomic motion (where the operations $e^{\mathcal{L}_x \Delta t}$ and $e^{\mathcal{L}_p \Delta t}$ are applied). The volume-updating step $e^{\mathcal{L}_e \Delta t}$ of the SCR barostat (related to eq 33) always directly involves the internal pressure (P_{int}) and the force. Under such a circumstance, the SCR barostat should be applied only once in the whole integration for each finite time interval Δt , as suggested by Figure 3. More discussion and numerical results are available in Section S3.

We apply the proposed MD/PIMD integrators of the “middle” scheme (for the isobaric–isothermal ensemble) to different benchmark molecular systems to test their accuracy and efficiency. Section S5 presents the explicit numerical algorithms of the “middle” scheme with the MTTK or SCR barostat for the MD/PIMD simulation of the isobaric–isothermal ensemble.

When the MTTK barostat is used, the first conventional scheme employed for comparison reads

$$\begin{aligned} e^{\mathcal{L}_{\text{NPT}}^{\text{Side}-1} \Delta t} = & e^{\mathcal{L}_B \Delta t / 2} e^{\mathcal{L}_T \Delta t / 2} e^{\mathcal{L}_{p_e} \Delta t / 2} e^{\mathcal{L}_p \Delta t / 2} e^{\mathcal{L}_V \Delta t} e^{\mathcal{L}_x \Delta t} e^{\mathcal{L}_p \Delta t / 2} \\ & \times e^{\mathcal{L}_{p_e} \Delta t / 2} e^{\mathcal{L}_T \Delta t / 2} e^{\mathcal{L}_B \Delta t / 2} \end{aligned} \quad (38)$$

and the second one is

$$\begin{aligned} e^{\mathcal{L}_{\text{NPT}}^{\text{Side}-2} \Delta t} = & e^{\mathcal{L}_B \Delta t / 2} e^{\mathcal{L}_{p_e} \Delta t / 2} e^{\mathcal{L}_p \Delta t / 2} e^{\mathcal{L}_T \Delta t / 2} e^{\mathcal{L}_V \Delta t} e^{\mathcal{L}_x \Delta t} e^{\mathcal{L}_T \Delta t / 2} \\ & \times e^{\mathcal{L}_p \Delta t / 2} e^{\mathcal{L}_{p_e} \Delta t / 2} e^{\mathcal{L}_B \Delta t / 2} \end{aligned} \quad (39)$$

Both schemes are implemented in the open-source and commercial software GROMACS⁶⁰ and are referred to as “vv” and “vv-avek”. When the SCR barostat is employed, the Euler integrator described in ref 42,

$$e^{\mathcal{L}_{\text{NPT}}^{\text{Euler}} \Delta t} = e^{\mathcal{L}_T \Delta t / 2} e^{\mathcal{L}_{\text{Bar}}^{\text{MD}} \Delta t} e^{\mathcal{L}_p \Delta t / 2} e^{\mathcal{L}_x \Delta t} e^{\mathcal{L}_p \Delta t / 2} e^{\mathcal{L}_T \Delta t / 2} \quad (40)$$

is also of the conventional “side” scheme. It is compared to the “middle” scheme of eq 30 with the SCR barostat.

In Section S5 of the Supporting Information, we present the details of the MD/PIMD algorithms of the recommended

“middle” scheme with the MTTK/SCR barostat for sampling the isobaric–isothermal ensemble.

3. NUMERICAL EXAMPLES

We first show the coordinate marginal distribution, volume marginal distribution, and momentum marginal distribution for a standard one-dimensional nanowire model, where both the recommended “middle” scheme and conventional “side” scheme are tested. We then test the Lennard-Jones (LJ) fluid and liquid para-hydrogen near its triple point with the Silvera-Goldman potential.⁶¹ These two systems only include nonelectrostatic intermolecular interactions. We further consider liquid water where all intramolecular, polar/nonpolar and electrostatic interactions are included. We test the q-SPC/fw flexible water model,⁶² SPC/E model where both intramolecular O–H bonds and H–O–H bond angles are rigidly fixed,⁶³ and accurate flexible and polarizable MB-pol model.^{64–66}

In Sections S1 and S2, we show that, in classical statistical mechanics, when the evaluation of the physical property includes the components that are functions of only the momentum, it is useful to take advantage of the known Maxwell momentum distribution to yield more accurate estimators. In classical MD simulations, while we employ eqs S36 and S44 to compute the density and isothermal compressibility, respectively, we use eqs S49, S51, and S52 for the evaluation of the enthalpy, isobaric heat capacity, and thermal expansion coefficient, respectively. When PIMD is used for quantum statistical mechanics, the evaluation of any physical property depends on only the path integral bead coordinate and the volume. In PIMD simulations, we utilize eqs S34, S36, S42, S44, and S48 as the estimators for calculating the enthalpy, density, isobaric heat capacity, isothermal compressibility, and thermal expansion coefficient, respectively.

3.1. One-Dimensional Nanowire Model. The one-dimensional nanowire model involves a single particle moving within a one-dimensional cosine potential

$$U(x, V) = \frac{m\omega^2 V^2}{4\pi^2} \left[1 - \cos\left(\frac{2\pi x}{V}\right) \right] \quad (41)$$

where $m = 1$, $\omega = 1$. Classical MD simulations of the isobaric–isothermal ensemble are performed at $k_B T = 0.01$, and $P_{\text{ext}} = 0.01$, and the “piston mass” is set to $W = 1000$, where all parameters and variables for this model are in reduced units.

We use the MTTK barostat for demonstration. For the middle scheme, all marginal distributions are sampled immediately after the barostatting step $e^{\mathcal{L}_{\text{Bar}}^{\text{MD}} \Delta t}$ in eq 31 or eq 30. When the side scheme is used, its marginal distributions are obtained after a full MD step in either eq 38 or eq 39, where the barostatting step ends. Both algorithms of eqs 38 and 39 of the side scheme are tested.

Figure 4 compares different schemes for the numerical performance of simulating the coordinate marginal distribu-

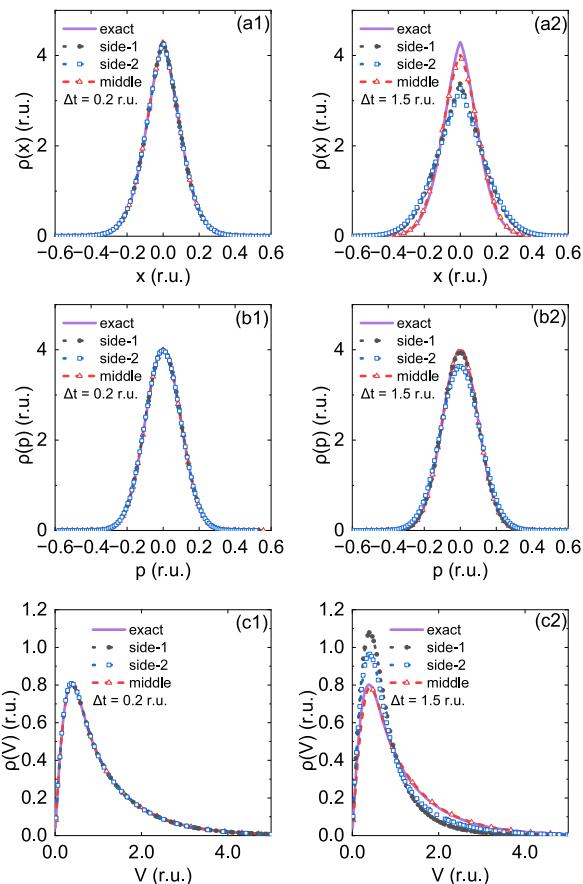


Figure 4. Results of the NPT MD simulations for the one-dimensional nanowire model. The MTTK barostat is used. Panels (a1,a2) show the coordinate marginal distribution using different schemes with a time interval of $\Delta t = 0.2$ r.u. and that with $\Delta t = 1.5$ r.u., respectively, while Panels (b1,b2) demonstrate the momentum distributions and Panels (c1,c2) illustrate the volume distributions. Red lines: the velocity-verlet-based “middle” scheme. Black and blue lines: two conventional “side” schemes. Purple lines: exact results. Twenty trajectories for each scheme are used. Each trajectory is equilibrated for 2×10^7 r.u., then propagated for additional 8×10^7 r.u. for evaluating physical properties. The friction parameter γ_{Lang} for the Langevin thermostat and γ_{Lang}^V for the barostat are 1.0 r.u. and 0.001 r.u., respectively. All MD simulations are performed by our independently developed MD/PIMD program.

tion, momentum marginal distribution and volume marginal distribution. As demonstrated in Panels (a1), (b1), and (c1), all schemes produce converged accurate marginal distributions with a time interval of $\Delta t = 0.2$ r.u. It is consistent with the expectation that all MD integrators should yield the same converged results as the time interval approaches zero. When a much larger time interval $\Delta t = 1.5$ r.u. is used, Panels (a2), (b2) and (c2) show that the middle scheme leads to results

that are in good agreement with exact marginal distributions, while two conventional side schemes exhibit pronounced deviations in both volume marginal distribution and coordinate marginal distribution. It verifies that, when the time interval Δt is relatively large, the middle scheme offers a much more accurate marginal distribution of the coordinate and volume variables (for the isothermal–isobaric ensemble) than the conventional MD algorithms (of the side scheme) do.

3.2. Lennard-Jones Molecular Liquid.

The Lennard-Jones (LJ) potential reads $u(r) = 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right]$, where r is the distance between two particles. The LJ potential is commonly used to describe the two-body interaction between nonpolar atoms/molecules in various force fields. For instance, the LJ parameters for liquid Argon are $\epsilon/k_B = 119.8$ K, $\sigma = 340.5$ pm, and $m = 39.962 A_r$ (relative atomic mass). Since the LJ liquid involves only the van der Waals intermolecular interaction, it serves as a clean benchmark system in which intramolecular and electrostatic forces play no role. We study the state point at external pressure $P = 1.706\epsilon/\sigma^3$ and inverse temperature $\beta = 0.4\epsilon$. The reduced unit (r.u.), where $\epsilon = \sigma = m = 1$, is used for convenience.

Classical MD simulations are performed with 256 particles in a cubic box with periodic boundary conditions applied using the minimum image convention. The time interval Δt ranges from 10^{-3} r.u. to 10^{-2} r.u. After the system is equilibrated for 2.5×10^5 r.u., 64 trajectories with each propagated up to 2.5×10^5 r.u. are employed for estimating thermodynamic properties. We test the three integrators with the MTTK barostat, namely, eq 23 of the recommended “middle” scheme, and eqs 38 and 39 of the conventional scheme.

Figure 5 demonstrates that the three integrators approach the same converged results (within statistical error bars) as the time interval is decreased, which is consistent with the expectation that all numerical integrators are in principle equivalent as the time interval becomes zero. Fully converged results are obtained at $\Delta t = 10^{-3}$ r.u. As shown in Panels a–b of **Figure 5**, the absolute deviation of the density is only 2×10^{-4} r.u. and that of the enthalpy is only 2×10^{-3} r.u. when comparing the converged value to value at $\Delta t = 10^{-2}$ r.u. for the “middle” scheme. In comparison, the same quantities for the “side” schemes are an order of magnitude larger, 1.6×10^{-3} r.u. for the density and 2.7×10^{-2} r.u. for the enthalpy per atom, respectively. Panels c–e of **Figure 5** display the results for isobaric heat capacity per atom, isothermal compressibility, and thermal expansion coefficient. Because the estimation of these properties involves second-order derivatives of the partition function, numerical simulations are more difficult to converge. Panels c–e of **Figure 5** also show that the time interval Δt for achieving the same accuracy is increased by a factor of 4–10 by the “middle” scheme over the conventional algorithms.

3.3. Liquid para-Hydrogen. Consider liquid *para*-hydrogen at low temperature, a benchmark system that includes only nonpolar intermolecular interactions, but significant nuclear quantum effects. Because its rotational state is at the ground energy level $J = 0$, the *para*-hydrogen molecule can effectively be treated as a spherical particle. The three integrators (eqs 23, 38, and 39) are tested for the PIMD simulation with 125 *para*-hydrogen molecules in a cubic box with periodic boundary conditions applied using the minimum image convention. In the PIMD simulation, 72 beads are used for the state point at

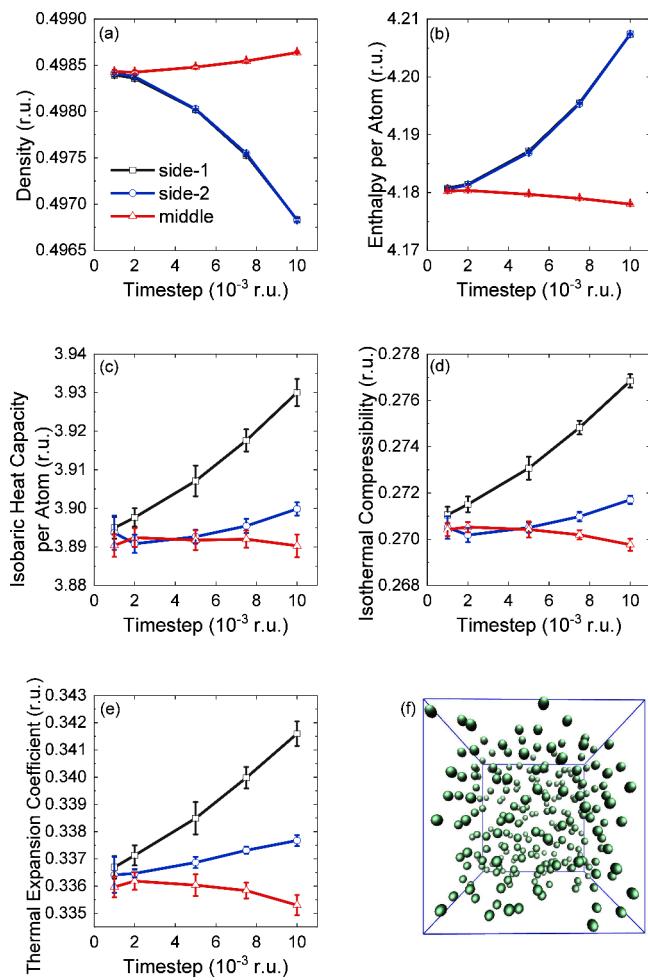


Figure 5. Results of NPT MD simulations for liquid argon using Lennard-Jones potential with increasing time intervals. The parameters of LJ potential are $\epsilon = k_B \times 119.8$ K, $\sigma = 340.5$ pm and $m = 39.962$ Å. The state point is at $\beta = 0.4 \epsilon$ and $P_{\text{ext}} = 1.706 \epsilon / \sigma^3$. The MTTK barostat is used for the simulation. The calculated thermodynamic properties include (a) density, (b) enthalpy per atom, (c) isobaric heat capacity per atom, (d) isothermal compressibility and (e) thermal expansion coefficient. Red lines: the velocity-verlet-based “middle” scheme. Black and blue lines: two conventional “side” schemes. 64 trajectories are equilibrated for 2.5×10^5 r.u. and then propagated for 2.5×10^5 r.u., using a cubic box with periodic boundary conditions containing 256 particles. All data are divided into 16 groups to calculate the standard error. The friction parameter for the Langevin thermostat (γ_{Lang}) and that for the barostat (γ_{Lang}^V) are 5.0 r.u. and 0.5 r.u., respectively. The “piston” mass W is set to 1000 r.u. A smooth cutoff is chosen, i.e., the potential is truncated at $r_c = 3.0 \sigma$ and is shifted to zero from $r_s = 2.5 \sigma$. For liquid argon, the reduced unit of time is 1.000×10^{-3} r.u. = $1.000 \times 10^{-3} \sqrt{m\sigma^2/\epsilon} = 2.1564$ fs. All MD simulations are performed by our independently developed MD/PIMD program. The data are verified by a modified DL_POLY_2 package.⁶⁷ Panel (f) is a snapshot of 256 Lennard-Jones particles in a cubic box with periodic boundary conditions.

zero external pressure and constant temperature $T = 14$ K, which is nearly the critical point of the system. After the system approaches equilibrium, 64 PIMD trajectories with each propagated up to ~ 40 ns are employed for estimation of physical properties of the isobaric–isothermal ensemble. Both the “all-mode scaled” method and “reduced dynamics” method are used for the MTTK barostat for the PIMD simulation of this system.

Figure 6 shows that the “middle” scheme (eq 23) is much more accurate and robust than the two conventional

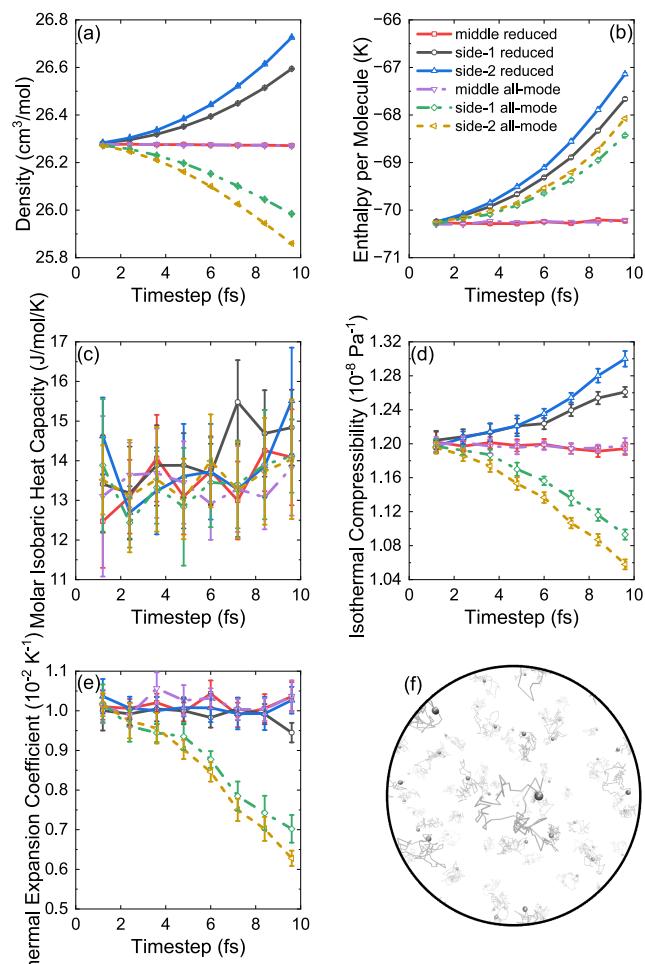


Figure 6. Results of NPT PIMD simulations for liquid para hydrogen at $T = 14$ K and $P_{\text{ext}} = 0$ using the Silvera-Goldman model⁶¹ with increasing time intervals. The MTTK barostat is used for the simulation. The calculated thermodynamic properties include (a) density, (b) enthalpy per molecule using the virial form kinetic energy, (c) molar isobaric heat capacity, (d) isothermal compressibility and (e) thermal expansion coefficient. Red lines: the velocity-verlet-based “middle” scheme for the “reduced dynamics” method. Black and blue lines: conventional “side” schemes for the “reduced dynamics” method. Purple dashed lines: the velocity-verlet-based “middle” scheme for the “all-mode scaled” method. Green and yellow dashed lines: “side” schemes for the “all-mode scaled” method. 72 beads are employed in the PIMD simulation. 64 trajectories are equilibrated for ~ 40 ns and propagated for ~ 40 ns using a cubic box with periodic boundary conditions containing 125 *para*-hydrogen molecules. All data are divided into 16 groups to calculate the standard error. The friction parameter for the Langevin thermostat (γ_{Lang}) is 1×10^{-3} a.u.⁻¹ for the first staging mode and $\sqrt{L}/\beta\hbar$ for other staging modes of PIMD as suggested in ref 37. The friction parameter for the barostat (γ_{Lang}^V) is 1×10^{-5} a.u.⁻¹. The “piston” mass W is set to 1×10^{10} a.u. The cutoff value for the Silvera-Goldman model is 8.466 Å. The atomic unit of time is 1 a.u. = 0.024189 fs. All PIMD simulations are performed by our independently developed MD/PIMD program. The data are verified by a modified DL_POLY_2 package.⁶⁷ Panel (f) is an illustration of the closed imaginary-time path obtained by 72 beads when simulating 125 *para*-hydrogen molecules in a cubic box with periodic boundary conditions.

algorithms of the “side” scheme (eqs 38 and 39). The “side” scheme fails at the time interval $\Delta t = 9.6$ fs (~ 400 a.u.) and produces significant deviations from the converged results, while the “middle” scheme with $\Delta t = 9.6$ fs still leads to converged results. When the “all-mode scaled” method is used for estimation of the internal pressure of the MTTK barostat in PIMD, the absolute difference from the converged value of the molar volume is as large as ~ 0.42 cm³/mol for the result produced by the “side” scheme at $\Delta t = 9.6$ fs. In comparison, the absolute difference of the molar volume is only $\sim 7 \times 10^{-3}$ cm³/mol for $\Delta t = 9.6$ fs when the “middle” scheme is utilized. When the “reduced dynamics” method is employed to estimate the internal pressure of the MTTK barostat in PIMD, the two conventional algorithms of the “side” scheme also lead to considerable deviation—the absolute difference from the converged value of the molar volume is as large as ~ 0.40 cm³/mol at $\Delta t = 9.6$ fs, while the same quantity produced by the “middle” scheme is only $\sim 6 \times 10^{-3}$ cm³/mol for $\Delta t = 9.6$ fs. The “middle” scheme also significantly outperforms the conventional “side” schemes for estimation of other properties, e.g., enthalpy per molecule, isothermal compressibility, thermal expansion coefficient, and so forth. Since the state point is near the critical point of the system, it is rather demanding to converge the value for the molar isobaric heat capacity. It is expected that the performance trend is similar when fully converged results (for this property) are obtained for each time interval.

3.4. Liquid Water with the Flexible q-SPC/fw Model.

We then consider liquid water at ambient conditions, where intramolecular, electrostatic, and van der Waals intermolecular interactions are involved. We use the flexible q-SPC/fw model for the demonstration. MD simulations are performed at constant temperature $T = 298.15$ K and constant external pressure $P = 1$ atm for a system of 216 molecules in a cubic box with periodic boundary conditions applied using the minimum image convention.

After the system reaches the equilibrium, 20 MD trajectories, each evolving up to 10 ns, are employed for evaluating thermodynamic properties. The time interval size Δt ranges from 0.2 to 2.0 fs. All integrators lead to the same converged results (within the statistical error) at $\Delta t = 0.2$ fs for all tested properties. As shown in Panel a in Figure 7, with a time interval of $\Delta t = 2.0$ fs, two “side” schemes produce significant absolute deviation from the converged value of the density, which is $\sim 3.4 \times 10^{-2}$ g·cm⁻³. In contrast, the same quantity produced by the proposed “middle” scheme is less than 1×10^{-3} g·cm⁻³. As shown in Panels b–e in Figure 7, the “middle” scheme significantly outperforms the “side” schemes for all other properties as well.

In addition to MD, we also perform PIMD for the same system. The “reduced dynamics” method (described in Section S1B) is used for the barostat for the PIMD simulation of liquid water. Here, 24 path integral beads are employed in the simulation. Five independent PIMD trajectories are propagated for 10 ns after equilibration for 1 ns. Figure 8 demonstrates that the advantage of the “middle” scheme over the conventional “side” scheme is even more pronounced for PIMD in comparison to MD. The absolute deviation at $\Delta t = 1.5$ fs from the converged value of the density is less than 1×10^{-3} g·cm⁻³ for the “middle” scheme of PIMD. In contrast, the corresponding deviation for the “side” scheme of PIMD is as large as $\sim 6.3 \times 10^{-2}$ g·cm⁻³, which is more than 60 times larger than the same quantity yield by the “middle” scheme of

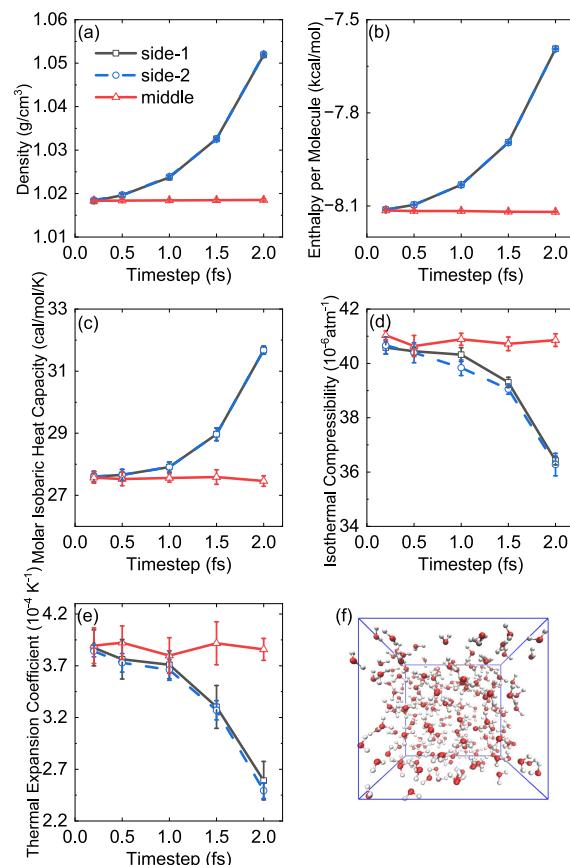


Figure 7. Results of NPT MD simulations for liquid water at $T = 298.15$ K and $P_{\text{ext}} = 1$ atm using q-SPC/fw model with increasing time intervals. The MTTK barostat is used for the MD simulation. The calculated thermodynamic properties include (a) density, (b) enthalpy per molecule, (c) molar isobaric heat capacity, (d) isothermal compressibility and (e) thermal expansion coefficient. Red lines: the velocity-verlet-based “middle” scheme. Black and blue lines: two conventional algorithms of the “side” scheme. Twenty trajectories are equilibrated for 1 ns and propagated for 10 ns, using a cubic box with periodic boundary conditions containing 216 molecules. All data are divided into 20 groups to calculate the standard error. The friction parameter for the Langevin thermostat (γ_{Lang}) and that for the barostat (γ_{Lang}^V) are all 5.0 ps⁻¹. The “piston” mass W is set to the recommended value as shown in eq 29. The cutoff for the short-range interactions is 9 Å. All MD simulations are performed by our modified DL_POLY_2 package.⁶⁷ Panel (f) is a snapshot of 216 water molecules in a cubic box with periodic boundary conditions.

PIMD. When the enthalpy or isobaric heat capacity is considered, the absolute deviation for the “side” scheme increases significantly as the time interval increases. When the time interval $\Delta t = 1.5$ fs is used, the deviation of the molar isobaric heat capacity for the “side” scheme can reach 130–180 cal/(mol·K), while that for the “middle” scheme is only 0.6 cal/(mol·K). Figure 8 shows that the “middle” scheme always outperforms the “side” scheme for the PIMD simulation of any physicochemical properties. When we compare the PIMD results of Figure 8 to the corresponding MD results of Figure 7, it is evident that the superiority of the “middle” scheme to the conventional algorithms of the “side” scheme is even more prominent in the PIMD simulation than in the MD simulation.

3.5. Liquid Water Model with Holonomic Constraints.

Constraints are widely used in the molecular simulation of

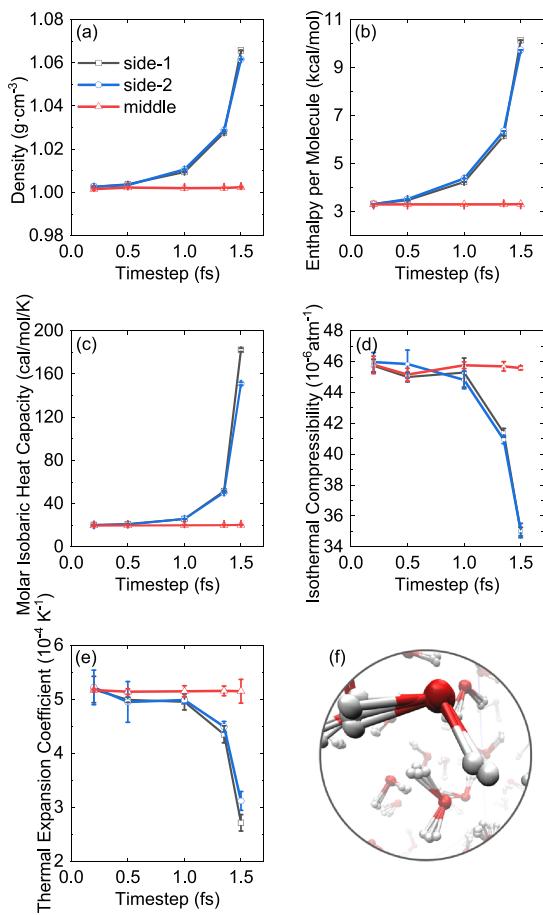


Figure 8. Results of NPT PIMD simulations for liquid water at $T = 298.15\text{ K}$ and $P_{\text{ext}} = 1\text{ atm}$ using q-SPC/fw model with increasing time intervals. The MTTK barostat is used for the PIMD simulation. The calculated thermodynamic properties include (a) density, (b) enthalpy per molecule using the virial form kinetic energy, (c) molar isobaric heat capacity, (d) isothermal compressibility and (e) thermal expansion coefficient. Red lines: the velocity-verlet-based “middle” scheme. Black and blue lines: conventional “side” schemes. Twenty-four beads are employed and 5 trajectories are equilibrated for 1 ns and propagated for 10 ns using a cubic box with periodic boundary conditions containing 216 molecules. The friction parameter for the Langevin thermostat (γ_{Lang}) is 5.0 ps^{-1} for the first staging mode and $\sqrt{L}/\beta\hbar$ for other staging modes of PIMD. The friction parameter for the barostat (γ_{Lang}) is 5.0 ps^{-1} . The “piston” mass W is set to the recommended value as shown in eq 29. The cutoff for the short-range interactions is 9 \AA . All PIMD simulations are performed by our modified DL_POLY_2 package.⁶⁷ Panel (f) is an illustration of path integral beads when simulating 216 water molecules in a cubic box with periodic boundary conditions.

biological systems to tackle multitime-scale problems. By applying holonomic constraints to high-frequency modes, it is possible to significantly increase the time interval in the MD simulation. In light of our investigation on the canonical ensemble in ref 49, we propose the “middle” scheme with holonomic constraints for the isobaric–isothermal ensemble. The details of the decomposition and of the algorithm are shown in Section S4. Here, we use the SCR barostat as an example for a liquid water model with fixed O–H bond lengths and H–O–H bond angles, where all intramolecular modes are frozen.

MD simulations with the SPC/E water model are performed at the state point with constant temperature $T = 298.15\text{ K}$ and constant external pressure $P = 1\text{ atm}$. We implement the “middle” scheme for the isobaric–isothermal ensemble in the AMBER software package.^{68,69} While a system of 324 water molecules in a cubic box with periodic boundary conditions is simulated with pmemd.MPI for CPU-based parallel computing, a system of 1296 water molecules in a cubic box with periodic boundary conditions is investigated with pmemd.cuda_SPFP for GPU acceleration. In the MD simulations, 16 trajectories are equilibrated for 1 ns and propagated for additional 10 ns for estimating thermodynamic properties. All data are divided into 16 groups to calculate the standard error. The friction parameter for the Langevin thermostat (γ_{Lang}) is 5.0 ps^{-1} . The characteristic time τ_p for the SCR barostat is set to 2.0 ps and the isothermal compressibility κ_T is $4.5 \times 10^{-5}\text{ bar}^{-1}$. The cutoff for the short-range interactions is 9 \AA .

The results are shown in Figure 9. When the time interval $\Delta t = 6.0\text{ fs}$ is used, the absolute deviation from the converged value of the density for the “middle” scheme is only $5 \times 10^{-4}\text{ g}\cdot\text{cm}^{-3}$, while the same quantity for the “side” scheme is as large as $6 \times 10^{-3}\text{ g}\cdot\text{cm}^{-3}$. Similarly, the “middle” scheme significantly outperforms the conventional “side” scheme in AMBER for calculating the enthalpy per molecule, molar isobaric heat capacity, isothermal compressibility, and thermal expansion coefficient. Figure 9 indicates that the “middle” scheme is superior to the conventional “side” scheme for simulating the isobaric–isothermal ensemble even for molecular systems with holonomic constraints.

3.6. Liquid Water with the Flexible and Polarizable MB-pol Force Field. MB-pol is an accurate, flexible and polarizable water model developed by Paesani and co-workers.^{64–66} Reference 71 shows that the density anomaly of water is indicated by the simulation results of liquid water at various temperatures computed by MB-pol. The MD simulations with the conventional “side” scheme reported in ref 71 was performed by the authors’ in-house DL_POLY_2 software package, where a cubic box of 256 molecules with periodic boundary conditions was used and the total time length of the MD trajectory ranged from 1.9 to 8.0 ns with a time interval of $\Delta t = 0.2\text{ fs}$. The Nose-Hoover chain thermostat and the MTTK barostat were applied. To obtain converged values, longer trajectories were required for simulations at lower temperatures.

Since ref 71 reported MD simulation results for the density and isobaric heat capacity produced by the conventional “side” scheme, we focus on the two thermodynamic properties for the comparison between the “middle” and “side” schemes. We perform MD simulations with MB-pol for liquid water under the condition of constant external pressure $P = 1\text{ atm}$ for a system of 256 molecules in a cubic box with periodic boundary conditions. The Langevin thermostat and the MTTK barostat are employed. After each trajectory is propagated to 0.2 ns for approaching the equilibrium, it is further propagated for 1 ns for evaluating thermodynamic properties. Ten MD trajectories are employed for $T = 278\text{ K}$ or lower temperatures, while five MD trajectories are used for temperatures above $T = 278\text{ K}$. The friction parameter for the Langevin thermostat (γ_{Lang}) and that for the barostat (γ_{Lang}^V) are 5.0 ps^{-1} . The “piston” mass W is set to the recommended value as described in eq 29. The cutoff for the short-range interactions is 9 \AA .

Figure 10 shows that, when the “middle” scheme is used, a time interval of $\Delta t = 1.5\text{ fs}$ is sufficient to achieve the same

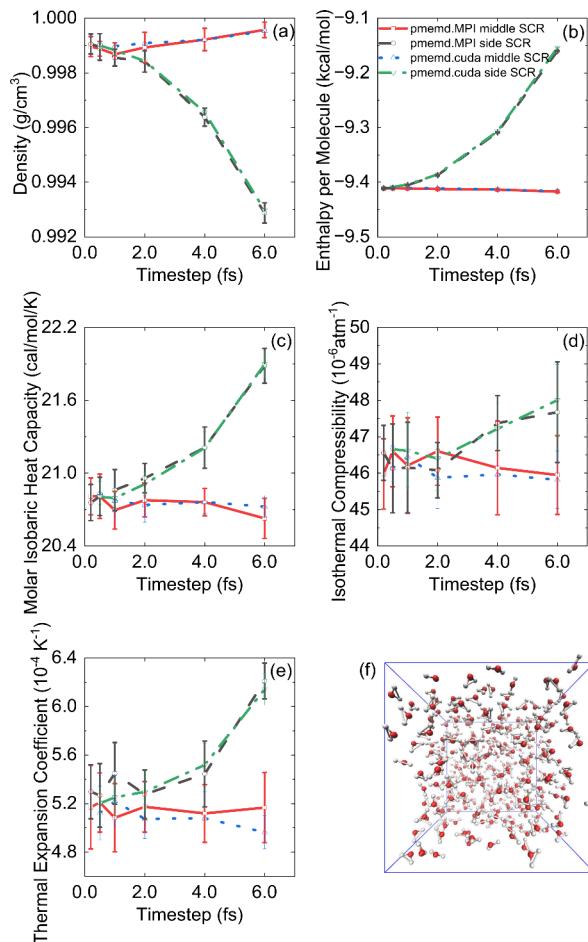


Figure 9. Results of liquid water with the SPC/E model at $T = 298.15\text{ K}$ and $P_{\text{ext}} = 1\text{ atm}$ with increasing time intervals. Intramolecular O–H bonds length and H–H bonds length corresponding to the fixed bond angle are constrained. The MD simulations are performed with AMBER. The SETTLE algorithm for constraining triatomic molecules is employed in the AMBER simulation.⁷⁰ The SCR barostat is implemented and used in AMBER. The calculated thermodynamic properties include (a) density, (b) enthalpy per molecule, (c) molar isobaric heat capacity, (d) isothermal compressibility and (e) thermal expansion coefficient. Red and green lines: the leapfrog-based “middle” scheme with CPU-based parallelization and with GPU-accelerated computing, respectively. Black and blue lines: the conventional “side” scheme with CPU-based parallelization and with GPU-accelerated computing, respectively. All MD simulations are performed using the modified AMBER software package.^{68,69} Panel (f) is a snapshot of 324 rigid water molecules in a cubic box with periodic boundary conditions.

converged results yielded by the conventional “side” scheme with $\Delta t = 0.2\text{ fs}$. This is consistent with the conclusion in Figure 7 with a simpler water force field. When $\Delta t = 1.5\text{ fs}$ is used with the “side” scheme instead, the results significantly deviate from the converged results. The “middle” scheme enhances the sampling efficiency for the isobaric–isothermal ensemble by a factor of 5–10 even for the accurate, but more computationally expensive MB-pol force field. It is expected that the “middle” scheme for the isobaric–isothermal ensemble will be generally useful to accelerate developing accurate (machine learning) force fields and to improve the efficiency of ab initio simulations for water/ice and other important chemical, materials, and biological molecular systems in solution or at interfaces/surfaces.^{64–66,72–84}

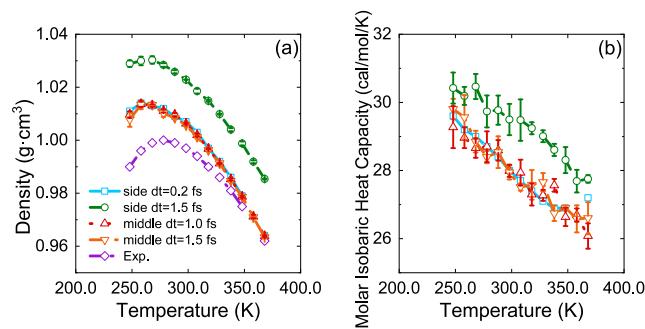


Figure 10. Simulation results for the temperature dependence of the density and that of the isobaric heat capacity for liquid water at $P_{\text{ext}} = 1\text{ atm}$ using the MB-pol model. The MTTK barostat is used for the MD simulation. Red lines and orange lines: the velocity-verlet-based “middle” scheme with two different time intervals. Green line: the conventional “side” scheme with the time interval of $\Delta t = 1.5\text{ fs}$. Cyan line: data from ref 71 produced by the conventional “side” scheme with $\Delta t = 0.2\text{ fs}$. Purple line: experiment data. All MD simulations are performed by our modified DL_POLY_2 software package.⁶⁷

Finally, we note that $\frac{1}{V_0}$ should, in principle, be replaced by $\frac{1}{V}$ and moved into the integrand of the volume variable for all related equations for the isobaric–isothermal ensemble throughout the paper. The modification leads to a more faithful description of the isobaric–isothermal ensemble when the number of particles of the system, N , is small.^{85,86} For instance, eqs 2 and 3 become

$$\langle B \rangle_{\text{NPT}}^{\text{CM}} = \frac{I_N}{Z_{\text{NPT}}} \int_0^\infty dV \int_{D(V)} d\mathbf{x} \int d\mathbf{p} \frac{1}{V} \exp[-\beta(PV + H(\mathbf{x}, \mathbf{p}))] B(\mathbf{x}, \mathbf{p}; V) \quad (42)$$

and

$$Z_{\text{NPT}} = I_N \int_0^\infty dV \int_{D(V)} d\mathbf{x} \int d\mathbf{p} \frac{1}{V} \exp[-\beta(PV + H(\mathbf{x}, \mathbf{p}))] \quad (43)$$

In the ideal gas limit, eqs 42 and 43 produce the correct equation of state $P\langle V \rangle = Nk_B T$ while eqs 2 and 3 yield $P\langle V \rangle = (N+1)k_B T$. It makes no difference in the thermodynamic limit (where the number of particles of the system approaches infinity). The difference due to the finite size effect is of the order of $1/N$, which is often smaller than the statistical error bar of the data for real molecular systems (e.g., those studied in the paper). It is trivial to apply the recommended “middle” scheme to cases where the modification is necessary.

4. CONCLUDING REMARKS

In this paper we propose the unified “middle” scheme for MD and PIMD to sample the isobaric–isothermal (constant-NPT) ensemble, regardless of which barostatting or thermostatting method is used. As described in ref 45, a reasonable thermostat always produces the correct Maxwell momentum distribution after the operation of $e^{\mathcal{L}_T \Delta t}$. Similarly, a reasonable barostat yields the correct volume distribution after the operation of $e^{\mathcal{L}_{\text{Bar}} \Delta t}$ at least in the limit $\Delta t \rightarrow 0$. Any reasonable barostats and thermostats can be employed in the unified “middle” scheme to produce simple, robust, accurate and efficient algorithms for MD or PIMD simulations of general molecular systems under conditions of constant pressure and temper-

ature. In the limit case that the barostat is effectively withdrew, the unified “middle” scheme for the constant-NPT ensemble is reduced to that for the constant-NVT ensemble proposed in refs 37, 45, 48, and 49. For demonstration in this work, we employ the MTTK barostat and the SCR barostat, with the Langevin thermostat, for studying isotropic systems. It is straightforward to implement the recommended “middle” scheme when other barostats (including the Parrinello–Rahman (PR) barostat³⁸) or thermostats (Andersen thermostat,¹⁴ Nosé–Hoover thermostat and Nosé–Hoover Chains,^{53,54,56,87} and so on) are used, as well as when anisotropic systems are investigated. For instance, in Section S6, we present the “middle” scheme with the MTTK barostat for simulating anisotropic systems under conditions of constant pressure and temperature.

As demonstrated in Section S2, the recommended “middle” scheme leads to a more accurate distribution of the volume and coordinate (configuration), when the same finite time interval Δt is used. The conclusion applies to general molecular systems where the van der Waals or/and electrostatic/polar intermolecular interactions are involved. MD or PIMD simulations of the isobaric–isothermal ensemble are performed for the LJ liquid, liquid *para*-hydrogen, liquid water with several typical force fields (SPC/E rigid water model, flexible q-SPC/fw model, and flexible and polarizable MB-pol model). We test a few typical thermodynamic properties of the isobaric–isothermal ensemble, which include the density, enthalpy, isobaric heat capacity, isothermal compressibility, and thermal expansion coefficient. In comparison to the two conventional algorithms of the “side” scheme (as implemented in GROMACS), the “middle” scheme demonstrates superior performance and increases the time interval by a factor of 5–10 to achieve the same accuracy of converged results. Because the “middle” scheme does not cost any additional numerical effort beyond what is required by the conventional NPT MD algorithms, it is capable of improving the sampling efficiency by a factor of 5–10 for the isobaric–isothermal ensemble.

We have integrated the recommended “middle” scheme for the isobaric–isothermal ensemble in prevailing molecular simulation packages, e.g., AMBER, GROMACS, and DL-POLY. For instance, based on our earlier contribution (on implementation of the “middle” thermostat scheme for sampling the canonical ensemble) to AmberTools and AMBER,^{49,88,89} the “middle” scheme with the SCR barostat for sampling the isobaric–isothermal ensemble has been coded in “pmemd” of AMBER for both serial and parallel versions of either of CPU and GPU computing, and in “sander” of AmberTools for CPU-based serial as well as parallel computing. It will be available in the 2025 version and later versions of AMBER and AmberTools. In the previous versions of GROMACS, a bug exists in the evaluation of the instantaneous internal pressure of the implementation of the MTTK barostat, which counts twice the tail corrections from the contribution of the dispersion interactions between the atoms beyond the cutoff. Such a bug leads to the higher converged density in MD simulations. We have reported this issue to GROMACS developers,⁹⁰ and have rectified the bug in the 2025 version. More importantly, we have integrated the efficient “middle” scheme with the MTTK barostat in GROMACS, which will be released soon. The recommended “middle” scheme for the isobaric–isothermal ensemble should also be implemented in CHARMM,⁹¹ LAMMPS,⁹² and other molecular simulation packages.

The “middle” scheme can be combined with other enhanced sampling techniques (e.g., thermodynamic integration,⁹³ umbrella sampling,^{94,95} metadynamics,⁹⁶ integrated tempering/Hamiltonian methods,^{97–100} reweighting techniques¹⁰¹). Because most nonadiabatic transition processes of real systems in condensed phase take place under conditions of constant pressure and temperature, the “middle” scheme is also useful to prepare the corresponding initial condition for trajectory-based nonadiabatic molecular dynamics methods.^{102–104} It is expected that the “middle” scheme for the isobaric–isothermal ensemble will be widely used for simulations for real complex systems, and will also help develop accurate (machine learning) force fields for important systems to understand or predict experimental processes/phenomena.^{64–66,72–84}

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jctc.5c00573>.

Evaluation of thermodynamic properties in MD and PIMD simulations; comparison between the “Middle” schemes with the MTTK barostat; derivation of the “Middle” scheme with the SCR barostat; “Middle” scheme for the isobaric–isothermal ensemble with holonomic constraints; numerical algorithms for the “Middle” scheme; and “Middle” scheme in the isobaric–isothermal ensemble for anisotropic systems ([PDF](#))

AUTHOR INFORMATION

Corresponding Author

Jian Liu – Beijing National Laboratory for Molecular Sciences, Institute of Theoretical and Computational Chemistry, College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, China; orcid.org/0000-0002-2906-5858; Email: jianliupku@pku.edu.cn

Authors

Weihao Liang – Beijing National Laboratory for Molecular Sciences, Institute of Theoretical and Computational Chemistry, College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, China

Sihan Wang – Beijing National Laboratory for Molecular Sciences, Institute of Theoretical and Computational Chemistry, College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, China

Cong Wang – Beijing National Laboratory for Molecular Sciences, Institute of Theoretical and Computational Chemistry, College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, China; College of Chemistry and Center for Advanced Quantum Studies, Key Laboratory of Theoretical and Computational Photochemistry, Ministry of Education, Beijing Normal University, Beijing 100875, China

Weizhou Wang – Beijing National Laboratory for Molecular Sciences, Institute of Theoretical and Computational Chemistry, College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, China

Xinchen She – Beijing National Laboratory for Molecular Sciences, Institute of Theoretical and Computational Chemistry, College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, China

Chongbin Wang — Beijing National Laboratory for Molecular Sciences, Institute of Theoretical and Computational Chemistry, College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, China

Jiushu Shao — College of Chemistry and Center for Advanced Quantum Studies, Key Laboratory of Theoretical and Computational Photochemistry, Ministry of Education, Beijing Normal University, Beijing 100875, China

Complete contact information is available at:
<https://pubs.acs.org/10.1021/acs.jctc.5c00573>

Author Contributions

§W.L., S.W., and C.W. contributed equally to this work.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank Xiangsong Cheng for helpful discussions. This work was supported by the National Science Fund for Distinguished Young Scholars Grant 22225304 and by the National Natural Science Foundation of China (NSFC) Grant 22450002. We acknowledge the High-Performance Computing Platform of Peking University, Beijing Beilong Super Cloud Computing Co., Ltd., Beijing PARATERA Tech Co., Ltd., the Beijing Super Cloud Computing Center (BSCC), the Computility Network of CAS, IKKEM Intelligent Computing Center and Guangzhou Supercomputer Center for providing computational resources.

REFERENCES

- (1) Drummond, N. D.; Monserrat, B.; Lloyd-Williams, J. H.; Ríos, P. L.; Pickard, C. J.; Needs, R. J. Quantum Monte Carlo Study of the Phase Diagram of Solid Molecular Hydrogen at Extreme Pressures. *Nat. Commun.* **2015**, *6*, 7794.
- (2) Bi, W.; Culverhouse, T.; Nix, Z.; Xie, W.; Tien, H.-J.; Chang, T.-R.; Dutta, U.; Zhao, J.; Lavina, B.; Alp, E. E.; Zhang, D.; Xu, J.; Xiao, Y.; Vohra, Y. K. Drastic Enhancement of Magnetic Critical Temperature and Amorphization in Topological Magnet EuSn₂p₂ under Pressure. *npj Quantum Mater.* **2022**, *7*, 43.
- (3) Eremets, M. I.; Minkov, V. S.; Drozdov, A. P.; Kong, P. P. The Characterization of Superconductivity under High Pressure. *Nat. Mater.* **2024**, *23*, 26–27.
- (4) Tagawa, S.; Sakamoto, N.; Hirose, K.; Yokoo, S.; Hernlund, J.; Ohishi, Y.; Yurimoto, H. Experimental Evidence for Hydrogen Incorporation into Earth's Core. *Nat. Commun.* **2021**, *12*, 2588.
- (5) Gleason, A. E.; Rittman, D. R.; Bolme, C. A.; Galtier, E.; Lee, H. J.; Granados, E.; Ali, S.; Lazicki, A.; Swift, D.; Celliers, P.; Militzer, B.; Stanley, S.; Mao, W. L. Dynamic Compression of Water to Conditions in Ice Giant Interiors. *Sci. Rep.* **2022**, *12*, 715.
- (6) Frenkel, D.; Smit, B. *Understanding Molecular Simulation: From Algorithms to Applications*, 3rd ed.; Academic Press, an imprint of Elsevier: London San Diego, CA Cambridge, MA Kidlington, 2023; p 728.
- (7) Martyna, G. J.; Hughes, A.; Tuckerman, M. E. Molecular Dynamics Algorithms for Path Integrals at Constant Pressure. *J. Chem. Phys.* **1999**, *110*, 3275–3290.
- (8) Tuckerman, M. E. *Statistical Mechanics: Theory and Molecular Simulation*; Oxford University Press: Oxford, New York, 2010.
- (9) McDonald, I. R. NPT-Ensemble Monte Carlo Calculations for Binary Liquid Mixtures. *Mol. Phys.* **1972**, *23*, 41–58.
- (10) Wood, W. W. Monte Carlo Calculations for Hard Disks in the Isothermal-Isobaric Ensemble. *J. Chem. Phys.* **1968**, *48*, 415–434.
- (11) Chow, K.-H.; Ferguson, D. M. Isothermal-Isobaric Molecular Dynamics Simulations with Monte Carlo Volume Sampling. *Comput. Phys. Commun.* **1995**, *91*, 283–289.
- (12) Faller, R.; De Pablo, J. J. Constant Pressure Hybrid Molecular Dynamics—Monte Carlo Simulations. *J. Chem. Phys.* **2002**, *116*, 55–59.
- (13) Åqvist, J.; Wennerström, P.; Nervall, M.; Bjelic, S.; Brandsdal, B. O. Molecular Dynamics Simulations of Water and Biomolecules with a Monte Carlo Constant Pressure Algorithm. *Chem. Phys. Lett.* **2004**, *384*, 288–294.
- (14) Andersen, H. C. Molecular Dynamics Simulations at Constant Pressure and/or Temperature. *J. Chem. Phys.* **1980**, *72*, 2384–2393.
- (15) Parrinello, M.; Rahman, A. Crystal Structure and Pair Potentials: A Molecular-Dynamics Study. *Phys. Rev. Lett.* **1980**, *45*, 1196–1199.
- (16) Parrinello, M.; Rahman, A. Polymorphic Transitions in Single Crystals: A New Molecular Dynamics Method. *J. Appl. Phys.* **1981**, *S2*, 7182–7190.
- (17) Nosé, S.; Klein, M. L. Constant Pressure Molecular Dynamics for Molecular Systems. *Mol. Phys.* **1983**, *50*, 1055–1076.
- (18) Di Pierro, M.; Elber, R.; Leimkuhler, B. A Stochastic Algorithm for the Isobaric–Isothermal Ensemble with Ewald Summations for All Long Range Forces. *J. Chem. Theory Comput.* **2015**, *11*, 5624–5637.
- (19) Ke, Q.; Gong, X.; Liao, S.; Duan, C.; Li, L. Effects of Thermostats/Barostats on Physical Properties of Liquids by Molecular Dynamics Simulations. *J. Mol. Liq.* **2022**, *365*, No. 120116.
- (20) Hoover, W. G. Constant-Pressure Equations of Motion. *Phys. Rev. A* **1986**, *34*, 2499–2500.
- (21) Melchionna, S.; Ciccotti, G.; Lee Holian, B. Hoover Npt Dynamics for Systems Varying in Shape and Size. *Mol. Phys.* **1993**, *78*, 533–544.
- (22) Martyna, G. J.; Tobias, D. J.; Klein, M. L. Constant Pressure Molecular Dynamics Algorithms. *J. Chem. Phys.* **1994**, *101*, 4177–4189.
- (23) Martyna, G. J.; Tuckerman, M. E.; Tobias, D. J.; Klein, M. L. Explicit Reversible Integrators for Extended Systems Dynamics. *Mol. Phys.* **1996**, *87*, 1117–1157.
- (24) Tuckerman, M. E.; Alejandre, J.; López-Rendón, R.; Jochim, A. L.; Martyna, G. J. A Liouville-Operator Derived Measure-Preserving Integrator for Molecular Dynamics Simulations in the Isothermal–Isobaric Ensemble. *J. Phys. A-Math. Gen.* **2006**, *39*, 5629–5651.
- (25) Yu, T.-Q.; Alejandre, J.; López-Rendón, R.; Martyna, G. J.; Tuckerman, M. E. Measure-Preserving Integrators for Molecular Dynamics in the Isothermal–Isobaric Ensemble Derived from the Liouville Operator. *Chem. Phys.* **2010**, *370*, 294–305.
- (26) Lippert, R. A.; Predescu, C.; Ierardi, D. J.; Mackenzie, K. M.; Eastwood, M. P.; Dror, R. O.; Shaw, D. E. Accurate and Efficient Integration for Molecular Dynamics Simulations at Constant Temperature and Pressure. *J. Chem. Phys.* **2013**, *139*, 164106.
- (27) Feller, S. E.; Zhang, Y.; Pastor, R. W.; Brooks, B. R. Constant Pressure Molecular Dynamics Simulation: The Langevin Piston Method. *J. Chem. Phys.* **1995**, *103*, 4613–4621.
- (28) Kolb, A.; Dünweg, B. Optimized Constant Pressure Stochastic Dynamics. *J. Chem. Phys.* **1999**, *111*, 4453–4459.
- (29) Grønbech-Jensen, N.; Farago, O. Constant Pressure and Temperature Discrete-Time Langevin Molecular Dynamics. *J. Chem. Phys.* **2014**, *141*, 194108.
- (30) Gao, X.; Fang, J.; Wang, H. Sampling the Isothermal-Isobaric Ensemble by Langevin Dynamics. *J. Chem. Phys.* **2016**, *144*, 124113.
- (31) He, X.; Wu, B.; Shang, Y.; Li, B.; Cheng, X.; Liu, J. New Phase Space Formulations and Quantum Dynamics Approaches. *Wiley Interdiscip. Rev.-Comput. Mol. Sci.* **2022**, *12*, No. e1619.
- (32) Feynman, R. P. Atomic Theory of the Lambda-Transition in Helium. *Phys. Rev.* **1953**, *91*, 1291–1301.
- (33) Feynman, R. P.; Hibbs, A. R. *Quantum Mechanics and Path Integrals*; McGraw-Hill: New York, 1965.
- (34) Chandler, D.; Wolynes, P. G. Exploiting the Isomorphism between Quantum Theory and Classical Statistical Mechanics of Polyatomic Fluids. *J. Chem. Phys.* **1981**, *74*, 4078–4095.
- (35) Ceperley, D. M. Path Integrals in the Theory of Condensed Helium. *Rev. Mod. Phys.* **1995**, *67*, 279–355.

- (36) Tuckerman, M. E.; Berne, B. J.; Martyna, G. J.; Klein, M. L. Efficient Molecular Dynamics and Hybrid Monte Carlo Algorithms for Path Integrals. *J. Chem. Phys.* **1993**, *99*, 2796–2808.
- (37) Liu, J.; Li, D.; Liu, X. A Simple and Accurate Algorithm for Path Integral Molecular Dynamics with the Langevin Thermostat. *J. Chem. Phys.* **2016**, *145*, No. 024103.
- (38) Parrinello, M.; Rahman, A. Study of an F Center in Molten KCl. *J. Chem. Phys.* **1984**, *80*, 860–867.
- (39) Berne, B. J.; Thirumalai, D. On the Simulation of Quantum Systems: Path Integral Methods. *Annu. Rev. Phys. Chem.* **1986**, *37*, 401–424.
- (40) Tuckerman, M. E.; Marx, D.; Klein, M. L.; Parrinello, M. Efficient and General Algorithms for Path Integral Car–Parrinello Molecular Dynamics. *J. Chem. Phys.* **1996**, *104*, 5579–5588.
- (41) Berendsen, H. J. C.; Postma, J. P. M.; Van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81*, 3684–3690.
- (42) Bernetti, M.; Bussi, G. Pressure Control Using Stochastic Cell Rescaling. *J. Chem. Phys.* **2020**, *153*, 114107.
- (43) Rogge, S. M. J.; Vanduyfhuys, L.; Ghysels, A.; Waroquier, M.; Verstraelen, T.; Maurin, G.; Van Speybroeck, V. A Comparison of Barostats for the Mechanical Characterization of Metal–Organic Frameworks. *J. Chem. Theory Comput.* **2015**, *11*, 5583–5597.
- (44) Del Tutto, V.; Raiteri, P.; Bernetti, M.; Bussi, G. Molecular Dynamics of Solids at Constant Pressure and Stress Using Anisotropic Stochastic Cell Rescaling. *Appl. Sci.* **2022**, *12*, 1139.
- (45) Zhang, Z.; Liu, X.; Chen, Z.; Zheng, H.; Yan, K.; Liu, J. A Unified Thermostat Scheme for Efficient Configurational Sampling for Classical/Quantum Canonical Ensembles Via Molecular Dynamics. *J. Chem. Phys.* **2017**, *147*, No. 034109.
- (46) Li, D.; Han, X.; Chai, Y.; Wang, C.; Zhang, Z.; Chen, Z.; Liu, J.; Shao, J. Stationary State Distribution and Efficiency Analysis of the Langevin Equation Via Real or Virtual Dynamics. *J. Chem. Phys.* **2017**, *147*, 184104.
- (47) Li, D.; Chen, Z.; Zhang, Z.; Liu, J. Understanding Molecular Dynamics with Stochastic Processes Via Real or Virtual Dynamics. *Chin. J. Chem. Phys.* **2017**, *30*, 735–760.
- (48) Zhang, Z.; Yan, K.; Liu, X.; Liu, J. A Leap-Frog Algorithm-Based Efficient Unified Thermostat Scheme for Molecular Dynamics. *Chin. Sci. Bull.* **2018**, *63*, 3467–3483.
- (49) Zhang, Z.; Liu, X.; Yan, K.; Tuckerman, M. E.; Liu, J. Unified Efficient Thermostat Scheme for the Canonical Ensemble with Holonomic or Isokinetic Constraints Via Molecular Dynamics. *J. Phys. Chem. A* **2019**, *123*, 6056–6079.
- (50) Liu, X.; Yan, K.; Liu, J. In Efficient “Middle” Thermostat Scheme for the Quantum/Classical Canonical Ensemble Via Molecular Dynamics, Advances in Quantum Systems in Chemistry, Physics, and Biology: Selected Proceedings of QSCP-XXIII (Kruger Park, South Africa, September 2018) 23; Springer, 2020; pp 257–281.
- (51) Leimkuhler, B.; Matthews, C. Rational Construction of Stochastic Numerical Methods for Molecular Sampling. *Appl. Math. Res. eXpress* **2012**, *2013*, 34–56.
- (52) Grønbech-Jensen, N.; Farago, O. A Simple and Effective Verlet-Type Algorithm for Simulating Langevin Dynamics. *Mol. Phys.* **2013**, *111*, 983–991.
- (53) Nosé, S. A Molecular Dynamics Method for Simulations in the Canonical Ensemble. *Mol. Phys.* **1984**, *52*, 255–268.
- (54) Hoover, W. G. Canonical Dynamics: Equilibrium Phase-Space Distributions. *Phys. Rev. A* **1985**, *31*, 1695–1697.
- (55) Itoh, S. G.; Morishita, T.; Okumura, H. Decomposition-Order Effects of Time Integrator on Ensemble Averages for the Nosé–Hoover Thermostat. *J. Chem. Phys.* **2013**, *139*, No. 064103.
- (56) Martyna, G. J.; Klein, M. L.; Tuckerman, M. Nosé–Hoover Chains: The Canonical Ensemble Via Continuous Dynamics. *J. Chem. Phys.* **1992**, *97*, 2635–2643.
- (57) Quigley, D.; Probert, M. I. J. Langevin Dynamics in Constant Pressure Extended Systems. *J. Chem. Phys.* **2004**, *120*, 11432–11441.
- (58) Wang, C.; Liu, J.; Shao, J. Unpublished.
- (59) Wang, W.; She, X.; Liu, J. Unpublished.
- (60) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. Gromacs: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* **2015**, *1–2*, 19–25.
- (61) Silveira, I. F.; Goldman, V. V. The Isotropic Intermolecular Potential for H₂ and D₂ in the Solid and Gas Phases. *J. Chem. Phys.* **1978**, *69*, 4209–4213.
- (62) Paesani, F.; Zhang, W.; Case, D. A.; Cheatham, T. E.; Voth, G. A. An Accurate and Simple Quantum Model for Liquid Water. *J. Chem. Phys.* **2006**, *125*, 184507.
- (63) Berendsen, H. J.; Grigera, J. R.; Straatsma, T. P. The Missing Term in Effective Pair Potentials. *J. Phys. Chem.* **1987**, *91*, 6269–6271.
- (64) Medders, G. R.; Babin, V.; Paesani, F. Development of a “First-Principles” Water Potential with Flexible Monomers. III. Liquid Phase Properties. *J. Chem. Theory Comput.* **2014**, *10*, 2906–2910.
- (65) Babin, V.; Medders, G. R.; Paesani, F. Development of a “First Principles” Water Potential with Flexible Monomers. II: Trimer Potential Energy Surface, Third Virial Coefficient, and Small Clusters. *J. Chem. Theory Comput.* **2014**, *10*, 1599–1607.
- (66) Babin, V.; Leforestier, C.; Paesani, F. Development of a “First Principles” Water Potential with Flexible Monomers: Dimer Potential Energy Surface, Vrt Spectrum, and Second Virial Coefficient. *J. Chem. Theory Comput.* **2013**, *9*, 5395–5403.
- (67) Smith, W.; Forester, T. DL_Poly_2. 0: A General-Purpose Parallel Molecular Dynamics Simulation Package. *J. Mol. Graphics* **1996**, *14*, 136–141.
- (68) Case, D. A.; Aktulga, H. M.; Belfon, K.; Cerutti, D. S.; Cisneros, G. A.; Cruzeiro, V. W. D.; Forouzesh, N.; Giese, T.; Götz, A. W.; Gohlke, H.; Izadi, S.; Kasavajhala, K.; Kaymak, M. C.; King, E.; Kurtzman, T.; Lee, T.; Li, P.; Liu, J.; Luchko, T.; Luo, R.; Manathunga, M.; Machado, M. R.; Nguyen, H.; O’Hearn, K. A.; Onufriev, A.; Pan, F.; Pantano, S.; Qi, R.; Rahnamoun, A.; Risheh, A.; Schott-Verdugo, S.; Shajan, A.; Swails, J.; Wang, J.; Wei, H.; Wu, X.; Wu, Y.; Zhang, S.; Zhao, S.; Zhu, Q.; Cheatham, T. E., III; Roe, D. R.; Roitberg, A.; Simmerling, C.; York, D. M.; Nagan, M. C.; Merz, K. M. Ambertools. *J. Chem. Inf. Model.* **2023**, *63*, 6183–6191.
- (69) Case, D. A.; Aktulga, H. M.; Belfon, K.; Ben-Shalom, I. Y.; Berryman, J. T.; Brozell, S. R.; Cerutti, D. S.; Cheatham, III, T. E.; Cisneros, G. A.; Cruzeiro, V. W. D.; Darden, T. A.; Forouzesh, N.; Giambas, G.; Giese, T.; Gilson, M. K.; Gohlke, H.; Götz, A. W.; Harris, J.; Izadi, S.; Izmailov, S. A.; Kasavajhala, K.; Kaymak, M. C.; King, E.; Kovalenko, A.; Kurtzman, T.; Lee, T.; Li, P.; Lin, C.; Liu, J.; Luchko, T.; Luo, R.; Machado, M.; Man, V.; Manathunga, M.; Merz, K. M.; Miao, Y.; Mikhailovskii, O.; Monard, G.; Nguyen, H.; Hearn, K. A. O.; Onufriev, A.; Pan, F.; Pantano, S.; Qi, R.; Rahnamoun, A.; Roe, D. R.; Roitberg, A.; Sagui, C.; Schott-Verdugo, S.; Shajan, A.; Shen, J.; Simmerling, C. L.; Skrynnikov, N. R.; Smith, J.; Swails, J.; Walker, R. C.; Wang, J.; Wang, J.; Wei, H.; Wu, X.; Xiong, Y.; Xue, Y.; York, D. M.; Zhao, S.; Zhu, Q.; Kollman, P. A. Amber 2023; University of California: San Francisco, 2023.
- (70) Miyamoto, S.; Kollman, P. A. Settle: An Analytical Version of the Shake and Rattle Algorithm for Rigid Water Models. *J. Comput. Chem.* **1992**, *13*, 952–962.
- (71) Reddy, S. K.; Straight, S. C.; Bajaj, P.; Huy Pham, C.; Riera, M.; Moberg, D. R.; Morales, M. A.; Knight, C.; Götz, A. W.; Paesani, F. On the Accuracy of the Mb-Pol Many-Body Potential for Water: Interaction Energies, Vibrational Frequencies, and Classical Thermodynamic and Dynamical Properties from Clusters to Liquid Water and Ice. *J. Chem. Phys.* **2016**, *145*, 194504.
- (72) Qu, C.; Yu, Q.; Houston, P. L.; Conte, R.; Nandi, A.; Bowman, J. M. Interfacing Q-Aqua with a Polarizable Force Field: The Best of Both Worlds. *J. Chem. Theory Comput.* **2023**, *19*, 3446–3459.
- (73) Hoffman, M. P.; Xantheas, S. S. The Many-Body Expansion for Aqueous Systems Revisited: IV. Stabilization of Halide–Anion Pairs in Small Water Clusters. *J. Phys. Chem. A* **2024**, *128*, 9876–9892.
- (74) Heindel, J. P.; Herman, K. M.; Xantheas, S. S. Many-Body Effects in Aqueous Systems: Synergies between Interaction Analysis

- Techniques and Force Field Development. *Annu. Rev. Phys. Chem.* **2023**, *74*, 337–360.
- (75) Heindel, J. P.; Sami, S.; Head-Gordon, T. Completely Multipolar Model as a General Framework for Many-Body Interactions as Illustrated for Water. *J. Chem. Theory Comput.* **2024**, *20*, 8594–8608.
- (76) Weldon, R.; Wang, F. Water Potential from Adaptive Force Matching for Ice and Liquid with Revised Dispersion Predicts Supercooled Liquid Anomalies in Good Agreement with Two Independent Experimental Fits. *J. Phys. Chem. B* **2024**, *128*, 3398–3407.
- (77) Feng, Y.; Cong, Y.; Zhao, Y.; Zhang, C.; Song, H.; Fang, B.; Yang, F.; Zhang, H.; Zhang, J. Z. H.; Zhang, L. “Blade of Polarized Water Molecule” Is the Key to Hydrolase Catalysis Regulation. *J. Chem. Inf. Model.* **2024**, *64*, 7987–7997.
- (78) Fu, B.; Zhang, D. H. Accurate Fundamental Invariant-Neural Network Representation of Ab Initio Potential Energy Surfaces. *Natl. Sci. Rev.* **2023**, *10*, No. nwad321.
- (79) Zhang, L.; Han, J.; Wang, H.; Car, R.; E, W. Deep Potential Molecular Dynamics: A Scalable Model with the Accuracy of Quantum Mechanics. *Phys. Rev. Lett.* **2018**, *120*, No. 143001.
- (80) Wang, H.; Zhang, L.; Han, J.; E, W. DeepMD-Kit: A Deep Learning Package for Many-Body Potential Energy Representation and Molecular Dynamics. *Comput. Phys. Commun.* **2018**, *228*, 178–184.
- (81) Jiang, F.; Zhou, C.-Y.; Wu, Y.-D. Residue-Specific Force Field Based on the Protein Coil Library. Rssf1: Modification of Opls-Aa/L. *J. Phys. Chem. B* **2014**, *118*, 6983–6998.
- (82) Zhou, C.-Y.; Jiang, F.; Wu, Y.-D. Residue-Specific Force Field Based on Protein Coil Library. Rssf2: Modification of Amber Ff99sb. *J. Phys. Chem. B* **2015**, *119*, 1035–1047.
- (83) Ji, C.; Mei, Y.; Zhang, J. Z. H. Developing Polarized Protein-Specific Charges for Protein Dynamics: MD Free Energy Calculation of Pka Shifts for Asp26/Asp20 in Thioredoxin. *Biophys. J.* **2008**, *95*, 1080–1088.
- (84) Chen, K.-W.; Sun, T.-Y.; Wu, Y.-D. New Insights into the Cooperativity and Dynamics of Dimeric Enzymes. *Chem. Rev.* **2023**, *123*, 9940–9981.
- (85) Han, K.-K.; Son, H. S. On the Isothermal-Isobaric Ensemble Partition Function. *J. Chem. Phys.* **2001**, *115*, 7793–7794.
- (86) Corti, D. S. Isothermal-Isobaric Ensemble for Small Systems. *Phys. Rev. E* **2001**, *64*, No. 016128.
- (87) Nosé, S. A Unified Formulation of the Constant Temperature Molecular Dynamics Methods. *J. Chem. Phys.* **1984**, *81*, 511–519.
- (88) Zhang, Z. Methods and Applications of Quantum and Classical Dynamics for the Complex Molecular Systems. Postdoctoral Report (Advisor: Liu, J.), Peking University: China, 2019.
- (89) Zheng, H. Development of “Middle-Scheme” Thermostat Molecular Dynamics Method and Its Implementation and Applications on Software. Ph.D Thesis (Advisor: Liu, J.), Peking University: China, 2022.
- (90) GitLab Results of MTTK Barostat Diffs from Other Barostats. <https://gitlab.com/gromacs/gromacs/-/issues/5072> (accessed Sept 01, 2024).
- (91) Brooks, B. R.; Brooks, C. L.; Mackerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. CHARMM: The Biomolecular Simulation Program. *J. Comput. Chem.* **2009**, *30*, 1545–1614.
- (92) Thompson, A. P.; Aktulga, H. M.; Berger, R.; Bolintineanu, D. S.; Brown, W. M.; Crozier, P. S.; In 'T Veld, P. J.; Kohlmeyer, A.; Moore, S. G.; Nguyen, T. D.; Shan, R.; Stevens, M. J.; Tranchida, J.; Trott, C.; Plimpton, S. J. LAMMPS - a Flexible Simulation Tool for Particle-Based Materials Modeling at the Atomic, Meso, and Continuum Scales. *Comput. Phys. Commun.* **2022**, *271*, No. 108171.
- (93) Kirkwood, J. G. Statistical Mechanics of Fluid Mixtures. *J. Chem. Phys.* **1935**, *3*, 300–313.
- (94) Kästner, J. Umbrella Sampling. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2011**, *1*, 932–942.
- (95) Liu, C.; Gao, J.; Liu, M. Tutorial on Umbrella Sampling Simulation with a Combined Qm/Mm Potential: The Potential of Mean Force for an S_N2 Reaction in Water. *J. Phys. Chem. B* **2024**, *128*, 10506–10514.
- (96) Laio, A.; Parrinello, M. Escaping Free-Energy Minima. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 12562–12566.
- (97) Yang, L.; Gao, Y. Q. A Selective Integrated Tempering Method. *J. Chem. Phys.* **2009**, *131*, 214109.
- (98) Han, X.; Lei, Y.-K.; Li, M.; Gao, Y. Q. A Brief Review of Integrated Tempering Sampling Molecular Simulation. *Chemical Physics Reviews* **2024**, *5*, 5.
- (99) Mori, T.; Hamers, R. J.; Pedersen, J. A.; Cui, Q. Integrated Hamiltonian Sampling: A Simple and Versatile Method for Free Energy Simulations and Conformational Sampling. *J. Phys. Chem. B* **2014**, *118*, 8210–8220.
- (100) Yang, Y. I.; Niu, H.; Parrinello, M. Combining Metadynamics and Integrated Tempering Sampling. *J. Phys. Chem. Lett.* **2018**, *9*, 6426–6430.
- (101) Chen, J.-N.; Dai, B.; Wu, Y.-D. Probability Density Reweighting of High-Temperature Molecular Dynamics. *J. Chem. Theory Comput.* **2024**, *20*, 4977–4985.
- (102) Wu, B.; He, X.; Liu, J. Nonadiabatic Field on Quantum Phase Space: A Century after Ehrenfest. *J. Phys. Chem. Lett.* **2024**, *15*, 644–658.
- (103) He, X.; Cheng, X.; Wu, B.; Liu, J. Nonadiabatic Field with Triangle Window Functions on Quantum Phase Space. *J. Phys. Chem. Lett.* **2024**, *15*, 5452–5466.
- (104) Wu, B.; Li, B.; He, X.; Cheng, X.; Ren, J.; Liu, J. Nonadiabatic Field: A Conceptually Novel Approach for Nonadiabatic Quantum Molecular Dynamics. *J. Chem. Theory Comput.* **2025**, *21*, 3775–3813.