Efficient "Middle" Thermostat Scheme for the Quantum/Classical Canonical Ensemble via Molecular Dynamics

Xinzijian Liu, Kangyu Yan and Jian Liu

Abstract We have recently developed a unified "middle" thermostat scheme for rationally designing molecular dynamics (MD)/path integral molecular dynamics (PIMD) algorithms for efficiently sampling the configuration space for the classical/quantum canonical ensemble with or without constraints. It significantly improves the time interval by a factor of 4–10 to achieve the same accuracy for structural and configuration-dependent thermodynamic properties for MD (as well as for any thermodynamic properties for PIMD). It has been implemented in AMBER (2018/2019 version), which is available for MD/PIMD simulations with force fields, QM/MM, or ab initio methods.

Keywords Canonical ensemble · Molecular dynamics · Path integral molecular dynamics · Thermostat algorithms · "Middle" thermostat scheme · Quantum statistics · Classical statistics · Holonomic constraints · Multi-time-step techniques · Isokinetic constraints · Sampling efficiency

1 Introduction

Molecular dynamics (MD) has been widely used in chemistry, biology, materials, environmental science, and other scientific fields [1, 2]. When nuclear quantum effects are important, MD can be implemented to perform imaginary time path integral (PIMD) by virtue of the ring-polymer isomorphism [3, 4], which in principle offers a numerically exact tool for practically studying quantum statistical properties in molecular systems [4–8] where quantum exchange effects are negligible. Many cases of interest involve the canonical ensemble, of which the number of particles (N), volume (V) and temperature (T) are constant.

Xinzijian Liu and Kangyu Yan: Both authors contributed equally to the work.

X. Liu · K. Yan · J. Liu (✉)
Beijing National Laboratory for Molecular Sciences, Institute of Theoretical and Computational Chemistry, College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, China
e-mail: jianliupku@pku.edu.cn

L. Mammino et al. (eds.), *Advances in Quantum Systems in Chemistry, Physics, and Biology*, Progress in Theoretical Chemistry and Physics 32,
https://doi.org/10.1007/978-3-030-34941-7_13

In this review we focus on recent progress on the efficient "middle" thermostat [9–12] that has been established in a series of papers [7, 9, 11–16] and has been implemented in AMBER [17] (the 2018 and 2019 versions). In 2016 we first suggested the Langevin thermostat in the "middle" scheme to construct efficient sampling algorithms for PIMD [7, 13, 14]. In 2017 and 2018 we developed a unified framework [9, 11] based on either velocity-Verlet or leap-frog algorithms to apply various thermostat algorithms for MD [1, 15, 18–32] and those for PIMD [7, 13, 14, 33–39] proposed for the canonical (NVT) ensemble. In our unified theoretical framework, we can describe most conventional algorithms [1, 15, 18–27, 33–36, 38, 39] in the "side" or "end" scheme. The unified "middle" scheme we proposed in Refs. [9, 11] provides an efficient tool for configurational sampling with either stochastic or deterministic thermostats. We have further extended the "middle" thermostat scheme as an efficient configurational sampling tool for systems with holonomic or isokinetic constraints [12]. Even when multiple-time-step (MTS) techniques are employed, the "middle" scheme leads to new algorithms that outperform original ones [12]. We have recently extended PIMD to offer an exact tool for quantum statistical properties in coupled multi-electronic-state systems [40, 41] where the Born-Oppenheimer approximation breaks down, in which the "middle" thermostat scheme still offers an efficient sampling tool.

2 "Middle" scheme

The "middle" scheme for MD or PIMD allows for larger time intervals to obtain the same convergence value, which significantly reduces the computational cost. The "middle" scheme is a powerful tool for configurational sampling, thus is helpful for evaluating thermodynamic properties that depend on coordinates in classical MD simulation For PIMD, all the structural and thermodynamic properties in quantum mechanics are functions of only the configurations of the path integral beads, the "middle" scheme is particularly useful because the time interval can be increased by 4 to 10 times without loss of accuracy.

We briefly review three typical thermostats, including stochastic thermostats (the Andersen thermostat, Langevin dynamics) as well as deterministic ones (the Nosé-Hoover thermostat and Nosé-Hoover chain). The integration step with thermostats can be described as three parts, the step for updating coordinates, that for updating momenta and that for thermostatting, denoted "x", "p" and "T", respectively. Then the equations of motion are

$$\begin{pmatrix} d\mathbf{x} \\ d\mathbf{p} \end{pmatrix} = \underbrace{\begin{pmatrix} \mathbf{M}^{-1}\mathbf{p} \\ 0 \end{pmatrix} dt}_{\mathbf{x}} + \underbrace{\begin{pmatrix} 0 \\ -\frac{\partial U(\mathbf{x})}{\partial \mathbf{x}} \end{pmatrix} dt}_{\mathbf{p}} + \underbrace{\text{Thermostat}}_{T} \tag{1}$$

Here, $\mathbf{x}$ and $\mathbf{p}$ are coordinates and momentum in vector forms, $U(\mathbf{x})$ is the physical potential energy, $\mathbf{M}$ the diagonal mass matrix.

The evolution of density distribution for the phase space can be described by the forward Kolmogorov equation. Define the translational propagator of coordinates as $\mathcal{L}_{\mathbf{x}}$ and that of momenta as $\mathcal{L}_{\mathbf{p}}$.

$$\mathcal{L}_{\mathbf{x}}\rho = -\mathbf{p}^T\mathbf{M}^{-1}\frac{\partial\rho}{\partial\mathbf{x}} \tag{2}$$

$$\mathcal{L}_{\mathbf{p}}\rho = \left(\frac{\partial U}{\partial\mathbf{x}}\right)^T\frac{\partial\rho}{\partial\mathbf{p}} \tag{3}$$

The definition of $\mathcal{L}_T$ depends on the thermostat employed. For a time step Δt, the phase space propagators are $e^{\mathcal{L}_{\mathbf{x}}\Delta t}$, $e^{\mathcal{L}_{\mathbf{p}}\Delta t}$ and $e^{\mathcal{L}_T\Delta t}$. The phase space propagation of velocity-Verlet "middle" scheme we propose is written as

$$e^{\mathcal{L}\Delta t} \approx e^{\mathcal{L}_{VV}^{\text{Middle}}\Delta t} = e^{\mathcal{L}_{\mathbf{p}}\Delta t/2}e^{\mathcal{L}_{\mathbf{x}}\Delta t/2}e^{\mathcal{L}_T\Delta t}e^{\mathcal{L}_{\mathbf{x}}\Delta t/2}e^{\mathcal{L}_{\mathbf{p}}\Delta t/2} \tag{4}$$

We denote it as "p-x-T-x-p", where the operations are performed from right to left. Similarly, the phase space propagation of the leap frog "middle" scheme reads

$$e^{\mathcal{L}\Delta t} \approx e^{\mathcal{L}_{\text{LF}}^{\text{Middle}}\Delta t} = e^{\mathcal{L}_{\mathbf{x}}\Delta t/2}e^{\mathcal{L}_T\Delta t}e^{\mathcal{L}_{\mathbf{x}}\Delta t/2}e^{\mathcal{L}_{\mathbf{p}}\Delta t}, \tag{5}$$

which is donated as "p-x-T-x-p", with the operations in a right-to-left sequence.

In the "middle" scheme, the thermostat process for a full time step is performed immediately after the coordinate translation for half a time step, then the coordinate translation for another half time step is followed. In contrast, traditional thermostat algorithms can be included in the "side" scheme

$$e^{\mathcal{L}\Delta t} \approx e^{\mathcal{L}_{VV}^{\text{Side}}\Delta t} = e^{\mathcal{L}_T\Delta t/2}e^{\mathcal{L}_{\mathbf{p}}\Delta t/2}e^{\mathcal{L}_{\mathbf{x}}\Delta t}e^{\mathcal{L}_{\mathbf{p}}\Delta t/2}e^{\mathcal{L}_T\Delta t/2} \tag{6}$$

$$e^{\mathcal{L}\Delta t} \approx e^{\mathcal{L}_{\text{LF}}^{\text{Side}}\Delta t} = e^{\mathcal{L}_{\mathbf{x}}\Delta t}e^{\mathcal{L}_T\Delta t/2}e^{\mathcal{L}_{\mathbf{p}}\Delta t}e^{\mathcal{L}_T\Delta t/2} \tag{7}$$

or the "end" scheme

$$e^{\mathcal{L}\Delta t} \approx e^{\mathcal{L}_{VV}^{\text{End}}\Delta t} = e^{\mathcal{L}_T\Delta t}e^{\mathcal{L}_{\mathbf{p}}\Delta t/2}e^{\mathcal{L}_{\mathbf{x}}\Delta t}e^{\mathcal{L}_{\mathbf{p}}\Delta t/2}. \tag{8}$$

2.1 Typical thermostats

Various thermostat methods have been developed for NVT simulations, e.g. Andersen thermostat, Nosé-Hoover chain (NHC), Langevin dynamics, etc. The "middle" scheme naturally includes the two efficient Langevin thermostat algorithms

developed for MD in Refs. [29, 30, 32] and in Ref. [28] as shown in Ref. [15]. Below we briefly review the "middle" scheme with the Andersen thermostat and Nosé-Hoover chain.

2.1.1 "Middle" scheme with the Andersen thermostat

The Andersen thermostat was proposed by Andersen in 1980 for the canonical ensemble [18]. The temperature of the system is controlled by stochastic collisions with a heat bath. The time between the collisions is exponentially distributed. When a particle j undergoes a collision, its momentum is reassigned from a Maxwell momentum distribution of the target temperature T, while momenta of other particles remain unchanged. The evolution of the Andersen thermostat step is

$$\mathbf{p}^{(j)} \leftarrow \sqrt{\frac{1}{\beta}}\mathbf{M}_j^{1/2}\boldsymbol{\theta}_j, \quad \text{if } \mu_j < \nu\Delta t \big(\text{or more precisely } \mu_j < 1 - e^{-\nu\Delta t}\big) \quad \big(j = \overline{1, N}\big) \tag{9}$$

Here, ν is the collision frequency, $\boldsymbol{\theta}_j$ is a 3-dimensional Gaussian distributed random vector with zero mean and unit variance, which is independent for each particle as well as each invocation.

The phase space density propagator for the Andersen thermostat in a time interval Δt is

$$e^{\mathcal{L}_T\Delta t}\rho = \big(1 - e^{-\nu\Delta t}\big)\rho_{\mathrm{MB}}(\mathbf{p}) \int_{-\infty}^{\infty} \rho(\mathbf{x}, \mathbf{p})d\mathbf{p} + \ e^{-\nu\Delta t}\rho(\mathbf{x}, \mathbf{p}). \tag{10}$$

Consider a harmonic system $V(\mathbf{x}) = \frac{1}{2}\big(\mathbf{x} - \mathbf{x}_{eq}\big)^T\mathbf{A}\big(\mathbf{x} - \mathbf{x}_{eq}\big)$. It is easy to prove the stationary state distribution for the "middle" scheme with the Andersen thermostat is

$$\begin{aligned}\rho^{\mathrm{Middle}} = \frac{1}{Z_N}\exp\Bigg[-\beta\Bigg(&\frac{1}{2}\mathbf{p}^T(\mathbf{M} - \mathbf{A}\frac{\Delta t^2}{4})^{-1}\mathbf{p} \\ &+\frac{1}{2}\big(\mathbf{x} - \mathbf{x}_{eq}\big)^T\mathbf{A}\big(\mathbf{x} - \mathbf{x}_{eq}\big)\Bigg)\Bigg],\end{aligned} \tag{11}$$

where Z_N is the normalization constant, while that for the traditional schemes (either "side" or "end") reads

$$\rho^{\mathrm{Side}} = \rho^{\mathrm{End}} = \frac{1}{\overline{Z}_N}\exp\Bigg[-\beta\Bigg(\frac{1}{2}\mathbf{p}^T\mathbf{M}^{-1}\mathbf{p} + \frac{1}{2}\big(\mathbf{x} - \mathbf{x}_{eq}\big)^T(\mathbf{1} - \mathbf{A}\mathbf{M}^{-1}\frac{\Delta t^2}{4})\mathbf{A}\big(\mathbf{x} - \mathbf{x}_{eq}\big)\Bigg)\Bigg]. \tag{12}$$

That is, in the harmonic limit, either the "side" or "end" schemes with a finite time interval leads to the exact momentum distribution but an approximate configurational distribution, and in contrast, the "middle" scheme produces the exact configurational distribution but an approximate momentum distribution. Since evaluation of most properties in MD and that of all properties in PIMD depend on only the configurational distribution, the "middle" scheme is a better choice for MD/PIMD simulations.

2.1.2 "Middle" scheme with the Nosé-Hoover chain thermostat

Nosé-Hoover chain (NHC) [21] performs deterministic MD by adding extended degrees of freedom to control the temperature of the system. The equations of motion for NHC read

$$\left.\begin{aligned}
\dot{x}_i &= \frac{p_i}{m_i} \\
\dot{p}_i &= -\frac{\partial U}{\partial x_i} - \frac{p_{\eta_1^{(i)}}}{Q_1} p_i \\
\dot{\eta}_j^{(i)} &= \frac{p_{\eta_j^{(i)}}}{Q_j} \\
\dot{p}_{\eta_1^{(i)}} &= \frac{p_i^2}{m_i} - k_B T - \frac{p_{\eta_2^{(i)}}}{Q_2} p_{\eta_1^{(i)}} \\
\dot{p}_{\eta_j^{(i)}} &= \frac{p^2_{\eta_{j-1}^{(i)}}}{Q_{j-1}} - k_B T - \frac{p_{\eta_{j+1}^{(i)}}}{Q_{j+1}} p_{\eta_j^{(i)}} \left(j = \overline{2, M_{\mathrm{NHC}} - 1}\right) \\
\dot{p}_{\eta_{M_{\mathrm{NHC}}}^{(i)}} &= \frac{p^2_{\eta_{M_{\mathrm{NHC}}-1}^{(i)}}}{Q_{M_{\mathrm{NHC}}-1}} - k_B T
\end{aligned}\right\} \left(i = \overline{1, 3N}\right) \tag{13}$$

where M_{NHC} means the number of pairs for the additional variables $\left\{\eta_j^{(i)}, p_{\eta_j^{(i)}}\right\}$ $\left(j = \overline{1, M_{\mathrm{NHC}}}\right)$ coupled to each degree of freedom $\left(i = \overline{1, 3N}\right)$, the parameters $Q_1, \ldots, Q_{M_{\mathrm{NHC}}}$ are the NHC thermal masses. For achieving the accuracy, one often uses the multiple time-scale techniques including the reference system propagator algorithm (RESPA) [23] and higher-order factorizations such as the Suzuki-Yoshida decomposition framework [42–44]. When the phase space density propagator for the thermostat part $e^{\mathcal{L}_T \Delta t}$ is effectively accurate, the thermostat propagation keeps the Maxwell-Boltzmann distribution for the momentum effectively unchanged, i.e.,

$$e^{\mathcal{L}_T \Delta t} \exp\left\{-\beta\left[\frac{1}{2}\mathbf{p}^T \mathbf{M}^{-1}\mathbf{p} + \sum_{i=1}^{3N}\sum_{j=1}^{M_{\mathrm{NHC}}} \frac{p^2_{\eta_j^{(i)}}}{2Q_j}\right]\right\} = \exp\left\{-\beta\left[\frac{1}{2}\mathbf{p}^T \mathbf{M}^{-1}\mathbf{p} + \sum_{i=1}^{3N}\sum_{j=1}^{M_{\mathrm{NHC}}} \frac{p^2_{\eta_j^{(i)}}}{2Q_j}\right]\right\}. \tag{14}$$

The stationary state distribution of the variables $\left(\mathbf{x}, \mathbf{p}, \mathbf{p}_{\boldsymbol{\eta}}\right)$ for the harmonic system propagated with "middle-NHC" scheme is

$$\rho^{\text{Middle - NHC}} = \frac{1}{Z'_N} \exp\Bigg[-\beta\Bigg(\frac{1}{2}\mathbf{p}^T(\mathbf{M} - \mathbf{A}\frac{\Delta t^2}{4})^{-1}\mathbf{p} + \frac{1}{2}\big(\mathbf{x} - \mathbf{x}_{eq}\big)^T\mathbf{A}\big(\mathbf{x} - \mathbf{x}_{eq}\big) + \sum_{i=1}^{3N}\sum_{j=1}^{M_{\text{NHC}}}\frac{p^2_{\eta_j^{(i)}}}{2Q_j}\Bigg)\Bigg], \tag{15}$$

where Z'_N is the normalization constant. The stationary marginal distribution is obtained by integration over p_η in Eq. (15) for the phase space variables $(\mathbf{x}, \mathbf{p})$, which is the same as Eq. (11). Similarly, the stationary state marginal distribution of the physical phase space variables $(\mathbf{x}, \mathbf{p})$ for the harmonic system obtained by either "side-NHC" or "end-NHC" is the same as Eq. (12).

2.2 *Simulation results*

2.2.1 Quartic potential

We test the quartic potential $U(x) = x^4/4$ as a typical example in the anharmonic region, where the mass is 1 au and the temperature parameter is $\beta \equiv \frac{1}{k_B T} = 8$ au.

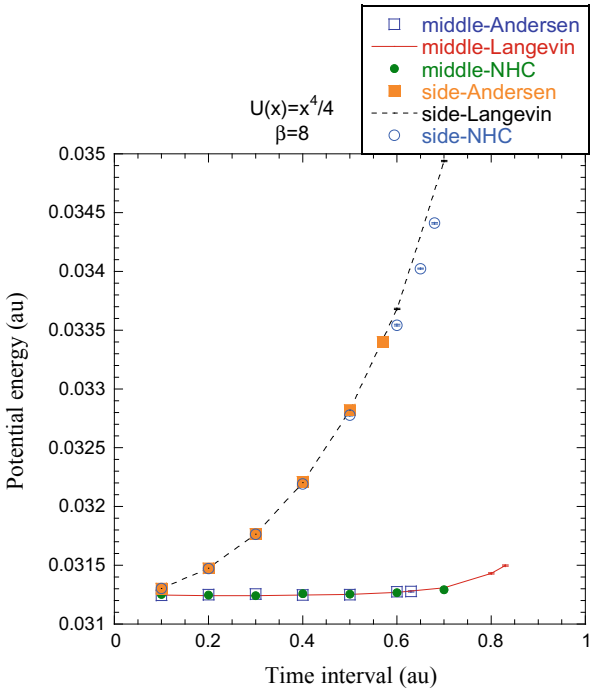

Fig. 1 MD results for the average potential energy (of the quartic potential system) with various time intervals (Reproduced with permission from Ref. [9])

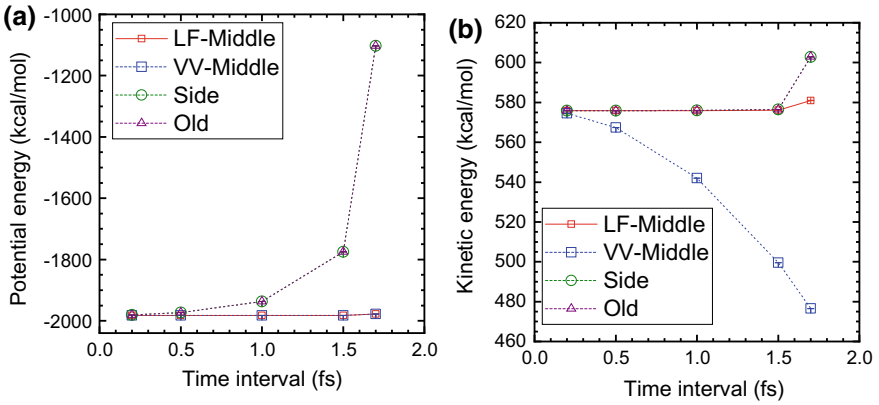

Fig. 2 MD results for liquid water at $T = 298.15$ K using different schemes, **a** average potential energy per atom $\langle U(\mathbf{x})\rangle/(N_{\text{atom}}k_B)$ (unit: Kelvin) **b** average kinetic energy per atom $\left\langle \mathbf{p}^T\mathbf{M}^{-1}\mathbf{p}\right\rangle/(2N_{\text{atom}}k_B)$ (unit: Kelvin) (Reproduced with permission from Ref. [11])

Figure 1 shows that the "middle" scheme is more efficient as well as more stable than the conventional "side" scheme, irrespective of the type of thermostat employed. To obtain the same accuracy, the "middle" scheme can employ a (much) larger time interval.

2.2.2 Liquid water

We have implemented the "middle" scheme in the AMBER2018 package [17]. A liquid water system containing a box of 216 water molecules using the TIP3P water model is used as a test case.

Figure 2 demonstrates that all algorithms lead to the same converged results as the time interval decreases to zero. As the time interval increases, both "VV-Middle" and "LF-Middle" schemes perform better than the "side" scheme or the conventional algorithm of AMBER (denoted "Old" in Fig. 2) for estimating the average potential energy. In addition, "LF-Middle" is more accurate than "VV-Middle" for evaluating the kinetic energy when the time interval becomes large.

3 Path integral molecular dynamics

Imaginary time path integral maps a quantum system into a classical ring polymer. Each bead of the ring polymer is a replica of the system, connected with adjacent beads by harmonic springs [3, 45, 46]. By assigning fictitious momenta and masses to the beads, one can employ MD to perform imaginary time path integral for the quantum system [4]. This approach is denoted path integral molecular dynamics

(PIMD), which has become a powerful tool for estimating quantum thermodynamic properties of realistic molecular systems [5, 7, 9, 38–40, 47].

3.1 Thermodynamic properties

One can write any thermodynamic property of the canonical ensemble in quantum mechanics in the general form

$$\left\langle \hat{B} \right\rangle = \frac{1}{Z} \mathrm{Tr}\left(e^{-\beta \hat{H}} \hat{B} \right), \tag{16}$$

where $Z = \mathrm{Tr}\left[e^{-\beta \hat{H}} \right]$ is the partition function and $\hat{B}$ represents a relevant operator of our interest. Eq. (16) can be expressed in the configuration space $\mathbf{x}$, i.e.,

$$\left\langle \hat{B} \right\rangle = \frac{\int d\mathbf{x} \langle \mathbf{x} | e^{-\beta \hat{H}} \hat{B} | \mathbf{x} \rangle}{\int d\mathbf{x} \langle \mathbf{x} | e^{-\beta \hat{H}} | \mathbf{x} \rangle}. \tag{17}$$

Substituting the expression of the identity operator in the configuration space $\hat{\mathbf{1}} = \int d\mathbf{x}_i |\mathbf{x}_i\rangle \langle \mathbf{x}_i|$, one can express the partition function as

$$\begin{aligned} Z &= \int\limits_{\mathbf{x}_1 \equiv \mathbf{x}} d\mathbf{x} \langle \mathbf{x} | e^{-\beta \hat{H}} | \mathbf{x} \rangle \\ &= \lim_{P \to \infty} \int d\mathbf{x}_1 \int d\mathbf{x}_2 \ldots \int d\mathbf{x}_P \left(\frac{P}{2\pi\beta\hbar^2} \right)^{3NP/2} |\mathbf{M}|^{P/2} \\ &\quad \times \exp\left\{ -\frac{P}{2\beta\hbar^2} \sum_{i=1}^{P} \left[(\mathbf{x}_{i+1} - \mathbf{x}_i)^T \mathbf{M} (\mathbf{x}_{i+1} - \mathbf{x}_i) \right] - \frac{\beta}{P} \sum_{i=1}^{P} U(\mathbf{x}_i) \right\} \end{aligned}, \tag{18}$$

where $\mathbf{x}_{P+1} \equiv \mathbf{x}_1$ and P is the total number of path integral beads. Similarly, the numerator of Eq. (17) becomes

$$\begin{aligned} \int d\mathbf{x} \langle \mathbf{x} | e^{-\beta \hat{H}} \hat{B} | \mathbf{x} \rangle &\overset{\mathbf{x}_1 \equiv \mathbf{x}}{=} \lim_{P \to \infty} \int d\mathbf{x}_1 \int d\mathbf{x}_2 \ldots \int d\mathbf{x}_P \left(\frac{P}{2\pi\beta\hbar^2} \right)^{3NP/2} |\mathbf{M}|^{P/2} \\ &\quad \times \exp\left\{ -\frac{P}{2\beta\hbar^2} \sum_{i=1}^{P} \left[(\mathbf{x}_{i+1} - \mathbf{x}_i)^T \mathbf{M} (\mathbf{x}_{i+1} - \mathbf{x}_i) \right] - \frac{\beta}{P} \sum_{i=1}^{P} U(\mathbf{x}_i) \right\} \\ &\quad \times \tilde{B}(\mathbf{x}_1, \ldots, \mathbf{x}_P) \end{aligned} \tag{19}$$

The estimator $\tilde{B}(\mathbf{x}_1, \ldots, \mathbf{x}_P)$ for the operator $\hat{B}(\hat{\mathbf{x}})$ is

$$\tilde{B}(\mathbf{x}_1, \ldots, \mathbf{x}_P) = \frac{1}{P} \sum_{j=1}^{P} B(\mathbf{x}_j). \tag{20}$$

When operator $\hat{B}$ involves the momentum, the estimator $\tilde{B}(\mathbf{x}_1, \ldots, \mathbf{x}_P)$ can also be expressed. For instance, the primitive version of the estimator for the kinetic energy operator $\hat{B} = \frac{1}{2}\hat{\mathbf{p}}^T\mathbf{M}^{-1}\hat{\mathbf{p}}$ is

$$\tilde{B}(\mathbf{x}_1, \ldots, \mathbf{x}_P) = \frac{3NP}{2\beta} - \sum_{j=1}^{P} \frac{P}{2\beta^2\hbar^2}\Big[(\mathbf{x}_{j+1} - \mathbf{x}_j)^T\mathbf{M}(\mathbf{x}_{j+1} - \mathbf{x}_j)\Big], \tag{21}$$

and the virial version is

$$\tilde{B}(\mathbf{x}_1, \ldots, \mathbf{x}_P) = \frac{3N}{2\beta} + \frac{1}{2P}\sum_{j=1}^{P}\left[(\mathbf{x}_j - \mathbf{x}^*)^T\frac{\partial U(\mathbf{x}_j)}{\partial \mathbf{x}_j}\right], \tag{22}$$

where $\mathbf{x}^*$ can be selected as the centroid of the path integral beads [48]

$$\mathbf{x}^* = \mathbf{x}_c \equiv \frac{1}{P}\sum_{j=1}^{P}\mathbf{x}_j \tag{23}$$

or any specific bead.

3.2 Staging Path Integral Molecular Dynamics

Consider the staging transformation [6, 38, 47, 49]

$$\begin{aligned}\boldsymbol{\xi}_1 &= \mathbf{x}_1\\ \boldsymbol{\xi}_j &= \mathbf{x}_j - \frac{(j-1)\mathbf{x}_{j+1} + \mathbf{x}_1}{j} \quad \big(j = \overline{2, P}\big).\end{aligned} \tag{24}$$

Define

$$\omega_P = \frac{1}{\beta\hbar}. \tag{25}$$

Eq. (18) becomes

$$\begin{aligned}Z \overset{\boldsymbol{\xi}_1 \equiv \mathbf{x}_1}{=} \lim_{\mathrm{P}\to\infty}\left(\frac{P}{2\pi\beta\hbar^2}\right)^{3N\mathrm{P}/2} |\mathbf{M}|^{P/2}\int d\boldsymbol{\xi}_1 \int d\boldsymbol{\xi}_2 \ldots \int d\boldsymbol{\xi}_P\\ \times \exp\left\{-\beta\sum_{j=1}^{P}\left[\frac{1}{2}\omega_{\mathrm{P}}^2\boldsymbol{\xi}_j^T\overline{\mathbf{M}}_{\mathrm{j}}\boldsymbol{\xi}_j + \frac{1}{P}U\big(\mathbf{x}_j(\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_P)\big)\right]\right\},\end{aligned} \tag{26}$$

The diagonal mass matrices are given by

$$\begin{aligned}\overline{\mathbf{M}}_1 &= 0\\ \overline{\mathbf{M}}_j &= \frac{j}{j-1}P\mathbf{M} \quad \left(j=\overline{2,P}\right).\end{aligned} \tag{27}$$

Define

$$\phi\left(\boldsymbol{\xi}_1,\ldots,\boldsymbol{\xi}_P\right)=\frac{1}{P}\sum_{j=1}^{P}U\left(\mathbf{x}_j\left(\boldsymbol{\xi}_1,\ldots,\boldsymbol{\xi}_P\right)\right). \tag{28}$$

Its derivatives satisfy the chain rule

$$\begin{aligned}\frac{\partial\phi}{\partial\boldsymbol{\xi}_1}&=\sum_{i=1}^{P}\frac{\partial\phi}{\partial\mathbf{x}_i}=\frac{1}{P}\sum_{i=1}^{P}U'(\mathbf{x}_i)\\ \frac{\partial\phi}{\partial\boldsymbol{\xi}_j}&=\frac{\partial\phi}{\partial\mathbf{x}_j}+\frac{j-2}{j-1}\frac{\partial\phi}{\partial\boldsymbol{\xi}_{j-1}} \quad \left(j=\overline{2,P}\right)\end{aligned}. \tag{29}$$

Adding fictitious momenta $(\mathbf{p}_1,\ldots,\mathbf{p}_P)$ into Eq. (26) leads to

$$\begin{aligned}Z\overset{\xi_1\equiv\mathbf{x}_1}{=}&\lim_{P\to\infty}\left(\frac{P}{4\pi^2\hbar^2}\right)^{3NP/2}|\mathbf{M}|^{P/2}\left(\prod_{j=1}^{P}\left|\tilde{\mathbf{M}}_j\right|\right)^{-1/2}\int\left(\prod_{j=1}^{P}d\boldsymbol{\xi}_j d\mathbf{p}_j\right)\\ &\times\exp\left[-\beta H_{\mathrm{eff}}\left(\boldsymbol{\xi}_1,\ldots,\boldsymbol{\xi}_P;\mathbf{p}_1,\ldots,\mathbf{p}_P\right)\right]\end{aligned} \tag{30}$$

with the effective Hamiltonian of the form

$$H_{\mathrm{eff}}\left(\boldsymbol{\xi}_1,\ldots,\boldsymbol{\xi}_P;\mathbf{p}_1,\ldots,\mathbf{p}_P\right)=\sum_{j=1}^{P}\frac{1}{2}\mathbf{p}_j^T\tilde{\mathbf{M}}_j^{-1}\mathbf{p}_j+U_{\mathrm{eff}}\left(\boldsymbol{\xi}_1,\ldots,\boldsymbol{\xi}_P\right), \tag{31}$$

where

$$U_{\mathrm{eff}}\left(\boldsymbol{\xi}_1,\ldots,\boldsymbol{\xi}_P\right)=\sum_{j=1}^{P}\frac{1}{2}\omega_{\mathrm{P}}^2\boldsymbol{\xi}_j^T\overline{\mathbf{M}}_{\mathrm{j}}\boldsymbol{\xi}_j+\phi\left(\boldsymbol{\xi}_1,\ldots,\boldsymbol{\xi}_P\right). \tag{32}$$

The fictitious masses are defined as

$$\begin{aligned}\tilde{\mathbf{M}}_1&=\mathbf{M}\\ \tilde{\mathbf{M}}_j&=\overline{\mathbf{M}}_j=\frac{j}{j-1}P\mathbf{M} \quad \left(j=\overline{2,P}\right)\end{aligned} \tag{33}$$

such that all staging modes $\left(\boldsymbol{\xi}_2,\ldots,\boldsymbol{\xi}_P\right)$ move on the same frequency. The estimator of the thermodynamic property in Eq. (17) is

$$\left\langle \hat{B} \right\rangle = \lim_{P\to\infty} \frac{\int \left(\prod_{j=1}^{P} d\boldsymbol{\xi}_j d\mathbf{p}_j \right) \exp\left\{ -\beta H_{\text{eff}}\left(\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_P; \mathbf{p}_1, \ldots, \mathbf{p}_P\right) \right\} \tilde{B}(\mathbf{x}_1, \ldots, \mathbf{x}_P)}{\int \left(\prod_{j=1}^{P} d\boldsymbol{\xi}_j d\mathbf{p}_j \right) \exp\left\{ -\beta H_{\text{eff}}\left(\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_P; \mathbf{p}_1, \ldots, \mathbf{p}_P\right) \right\}}. \tag{34}$$

Equations (31) and (34) suggest the Hamilton equations of motion

$$\begin{aligned} \dot{\boldsymbol{\xi}}_j &= \tilde{\mathbf{M}}_j^{-1}\, \mathbf{p}_j \\ \dot{\mathbf{p}}_j &= -\omega_P^2 \overline{\mathbf{M}}_j \boldsymbol{\xi}_j - \tfrac{\partial \phi}{\partial \boldsymbol{\xi}_j} \quad \left(j = \overline{1, P}\right) \end{aligned}. \tag{35}$$

One should couple it with a thermostat to ensure the canonical distribution for $\left(\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_P, \mathbf{p}_1, \ldots, \mathbf{p}_P\right)$. Note that the estimator of any thermodynamic properties in Eq. (34) depends on only the configurational distribution of the beads in PIMD.

The choice of the fictitious masses in Eqs. (25), (27), and (33) is different from that in Refs. [7, 9]. The procedure in Eqs. (24)–(35) for PIMD makes it possible to use the same time interval to obtain converged results, regardless of the value of P, the total number of path integral beads. This has been suggested earlier in Ref. [6].

The conventional wisdom often employs the decomposition of the equations of motion in PIMD algorithms

$$\begin{pmatrix} \dot{\boldsymbol{\xi}}_j \\ \dot{\mathbf{p}}_j \end{pmatrix} = \underbrace{\begin{pmatrix} \tilde{\mathbf{M}}_j^{-1}\, \mathbf{p}_j \\ -\omega_P^2 \overline{\mathbf{M}}_j \boldsymbol{\xi}_j \end{pmatrix}}_{\text{x}} + \underbrace{\begin{pmatrix} 0 \\ -\frac{\partial \phi}{\partial \boldsymbol{\xi}_j} \end{pmatrix}}_{\text{p}} + \underbrace{(\text{Thermostat})}_{T} \left(j = \overline{1, P}\right) \tag{36}$$

because the fictitious ring polymer force term $-\omega_P^2 \overline{\mathbf{M}}_j \boldsymbol{\xi}_j$ varies more frequently than the physical force term $-\frac{\partial \phi}{\partial \boldsymbol{\xi}_j}$, and the exact solution to the first term of Eq. (36) is available [39]. E. g., Eq. (36) leads to exact results in the free particle limit. Our recent work [7], however, shows that

$$\begin{pmatrix} \dot{\boldsymbol{\xi}}_j \\ \dot{\mathbf{p}}_j \end{pmatrix} = \underbrace{\begin{pmatrix} \tilde{\mathbf{M}}_j^{-1}\, \mathbf{p}_j \\ 0 \end{pmatrix}}_{\text{x}} + \underbrace{\begin{pmatrix} 0 \\ -\frac{\partial U_{\text{eff}}}{\partial \boldsymbol{\xi}_j} \end{pmatrix}}_{\text{p}} + \underbrace{(\text{Thermostat})}_{T} \quad \left(j = \overline{1, P}\right) \tag{37}$$

is a more accurate and efficient decomposition scheme for developing PIMD algorithms in the "middle" thermostat scheme.

E.g., when one apply the Langevin dynamics as thermostat, it has been clarified in Appendix C of Ref. [7] (and its Supplementary Material [50]) that Eq. (37) guarantees the exact marginal configuration distribution of the path integral beads in the harmonic limit, irrespective to the time interval, while Eq. (36) does not. The conclusion holds for any thermostat as long as the thermostat faithfully maintains

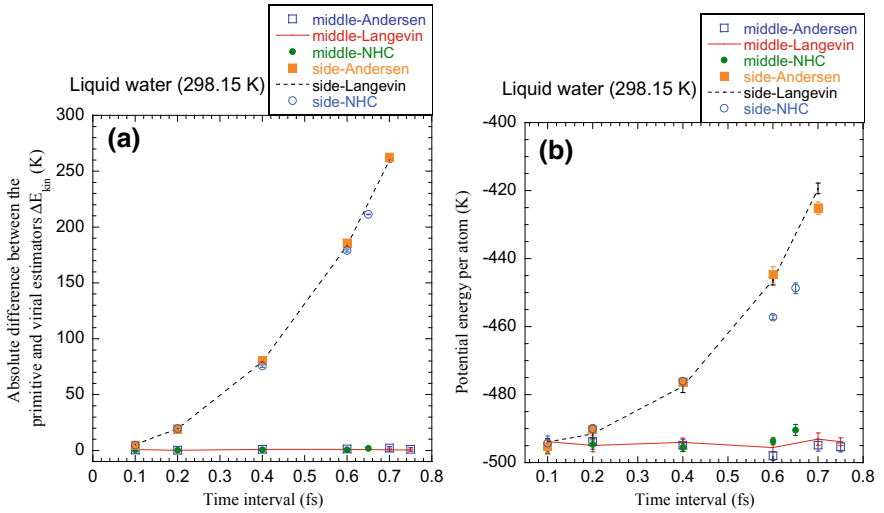

Fig. 3 PIMD results using different time intervals for liquid water at $T = 298.15$ K. **a** Absolute difference between the primitive and virial estimators for averaged kinetic energy per atom (unit: Kelvin). **b** The averaged potential energy per atom $\langle U(\mathbf{x})\rangle/(N_{\text{atom}}k_{\text{B}})$ (unit: Kelvin). Statistical error bars are included. (Reproduced with permission from Ref. [9])

the Maxwell momentum distribution. This is verified by the simulation results for PIMD in Ref. [9] when the Andersen thermostat and NHC are used.

We apply the "middle" and conventional "side" schemes with PIMD algorithms for liquid water system to study the canonical ensemble at constant temperature $T = 298.15$ K and $\rho_l = 0.997\ \text{g}\cdot\text{cm}^{-3}$. As presented in Fig. 3, in comparison to the "side" scheme, the "middle" scheme reduces the error by about an order of magnitude with the same time interval.

3.3 Normal-mode Path Integral Molecular Dynamics

Consider the normal mode transformation of path integral beads [48, 51, 52]

$$\mathbf{x} = \mathbf{C}^{norm}\mathbf{q} \tag{38}$$

with $\mathbf{q} = \left(\mathbf{q}_0 \ldots \mathbf{q}_{P-1}\right)^T$ and $\mathbf{x} = \left(\mathbf{x}_1 \ldots \mathbf{x}_P\right)^T$, where the element of the orthogonal transformation matrix $\mathbf{C}^{norm}$ is

$$C_{jk}^{norm} = \begin{cases} \sqrt{1/P}, & k = 0 \\ \sqrt{2/P}\cos(2\pi jk/P), & 1 \le k \le P/2 - 1 \\ \sqrt{1/P}(-1)^j, & k = P/2 \\ \sqrt{2/P}\sin(2\pi jk/P), & P/2 + 1 \le k \le P - 1 \end{cases} \quad \left(j = \overline{1, P}\right), \tag{39}$$

for even P and

$$C_{jk}^{norm} = \begin{cases} \sqrt{1/P}, & k = 0 \\ \sqrt{2/P}\cos(2\pi jk/P), & 1 \le k \le (P-1)/2 \\ \sqrt{2/P}\sin(2\pi jk/P), & (P+1)/2 \le k \le P - 1 \end{cases} \quad \left(j = \overline{1, P}\right), \tag{40}$$

for odd P. When one employs the normal mode transformation, the partition function becomes

$$Z = \lim_{P\to\infty}\left(\frac{P}{2\pi\beta\hbar^2}\right)^{NP/2} |\mathbf{M}|^{P/2} \int d\mathbf{q}_0 \int d\mathbf{q}_1 \ldots \int d\mathbf{q}_{P-1} \times \exp\left\{-\beta\left[\sum_{k=0}^{P-1}\frac{1}{2}\omega_k^2 \mathbf{q}_k^T \overline{\mathbf{M}}_k^{norm}\mathbf{q}_k + \frac{1}{P}\sum_{j=1}^{P} V\left(\mathbf{x}_j(\mathbf{q}_0, \ldots, \mathbf{q}_{P-1})\right)\right]\right\}. \tag{41}$$

The mass matrices are defined as $\overline{\mathbf{M}}_0^{norm} = \mathbf{0}$ and $\overline{\mathbf{M}}_k^{norm} = P\mathbf{M} \quad \left(k = \overline{1, P-1}\right)$, the frequency for each mode is given by

$$\omega_k = 2\omega_P \sin(k\pi/P) \quad \left(k = \overline{0, P-1}\right), \tag{42}$$

respectively. Eq. (28) then becomes

$$\phi(\mathbf{q}_0, \ldots, \mathbf{q}_{P-1}) = \frac{1}{P}\sum_{j=1}^{P} V\left(\mathbf{x}_j(\mathbf{q}_0, \ldots, \mathbf{q}_{P-1})\right), \tag{43}$$

and the derivatives $\partial\phi/\partial\mathbf{q}_k$ is obtained from

$$\frac{\partial\phi}{\partial\mathbf{q}} = \left(\frac{\partial\mathbf{x}}{\partial\mathbf{q}}\right)^T \frac{\partial\phi}{\partial\mathbf{x}} = \left(\mathbf{C}^{norm}\right)^T \frac{\partial\phi}{\partial\mathbf{x}}. \tag{44}$$

Employing fictitious momenta $(\mathbf{p}_0, \ldots, \mathbf{p}_{P-1})$ into Eq. (41) produces

$$Z = \lim_{P\to\infty}\left(\frac{P}{4\pi^2\hbar^2}\right)^{NP/2} |\mathbf{M}|^{P/2}\left(\prod_{k=0}^{P-1}\left|\tilde{\mathbf{M}}_k^{norm}\right|\right)^{-1/2} \int\left(\prod_{k=0}^{P-1} d\mathbf{q}_k d\mathbf{p}_k\right) \times \exp\left[-\beta H_{\text{eff}}^{norm}(\mathbf{q}_0, \ldots, \mathbf{q}_{P-1}; \mathbf{p}_0, \ldots, \mathbf{p}_{P-1})\right], \tag{45}$$

where the effective Hamiltonian can be expressed by the normal mode variables

$$H_{\text{eff}}^{norm}(\mathbf{q};\mathbf{p}) = \sum_{k=0}^{P-1} \frac{1}{2}\mathbf{p}_k^T \tilde{\mathbf{M}}_{norm,k}^{-1}\mathbf{p}_k + U_{\text{eff}}^{norm}(\mathbf{q}). \tag{46}$$

In Eq. (46) the effective potential is

$$U_{\text{eff}}^{norm}(\mathbf{q}) = \sum_{k=0}^{P-1} \frac{1}{2}\omega_k^2 \mathbf{q}_k^T \overline{\mathbf{M}}_k^{norm}\mathbf{q}_k + \phi(\mathbf{q}), \tag{47}$$

and the fictitious masses $\left\{\tilde{\mathbf{M}}_{norm,k}\right\}$ can be arbitrary, which can be chosen as $\tilde{\mathbf{M}}_{norm,k=0} = \mathbf{M}$ and $\tilde{\mathbf{M}}_{norm,k} = P\mathbf{M}\left(k = \overline{1, P-1}\right)$ such that all non-zeroth normal modes $\left(k = \overline{1, P-1}\right)$ move on the same frequency.

The thermodynamic property Eq. (17) can be evaluated by

$$\left\langle \hat{B} \right\rangle = \lim_{P\to\infty} \frac{\int \left(\prod_{k=0}^{P-1} d\mathbf{q}_k d\mathbf{p}_k\right) \exp\left\{-\beta H_{\text{eff}}^{norm}(\mathbf{q}_0,\ldots,\mathbf{q}_{P-1};\mathbf{p}_0,\ldots,\mathbf{p}_{P-1})\right\} \tilde{B}(\mathbf{x}_1,\ldots,\mathbf{x}_P)}{\int \left(\prod_{k=0}^{P-1} d\mathbf{q}_k d\mathbf{p}_k\right) \exp\left\{-\beta H_{\text{eff}}^{norm}(\mathbf{q}_0,\ldots,\mathbf{q}_{P-1};\mathbf{p}_0,\ldots,\mathbf{p}_{P-1})\right\}}. \tag{48}$$

It then suggests a MD scheme to sample $(\mathbf{q}_0,\ldots,\mathbf{q}_{P-1},\mathbf{p}_0,\ldots,\mathbf{p}_{P-1})$. The equations of motion derived from the effective Hamiltonian (Eq. 46) are

$$\begin{aligned} \dot{\mathbf{q}}_k &= \tilde{\mathbf{M}}_{norm,k}^{-1}\mathbf{p}_k \\ \dot{\mathbf{p}}_k &= -\omega_k^2 \overline{\mathbf{M}}_k^{norm}\mathbf{q}_k - \frac{\partial \phi}{\partial \mathbf{q}_k} \quad \left(k = \overline{0, P-1}\right), \end{aligned} \tag{49}$$

which must be coupled with a thermostat to ensure the correct canonical distribution for $(\mathbf{q}_0,\ldots,\mathbf{q}_{P-1},\mathbf{p}_0,\ldots,\mathbf{p}_{P-1})$. This is denoted normal mode PIMD (NM-PIMD).

Similar to Eq. (37), the equations of motion of NM-PIMD should be decomposed into three parts

$$\begin{pmatrix} \dot{\mathbf{q}}_k \\ \dot{\mathbf{p}}_k \end{pmatrix} = \underbrace{\begin{pmatrix} \tilde{\mathbf{M}}_{norm,k}^{-1}\mathbf{p}_k \\ 0 \end{pmatrix}}_{\text{x}} + \underbrace{\begin{pmatrix} 0 \\ -\frac{\partial U_{\text{eff}}}{\partial \mathbf{q}_k} \end{pmatrix}}_{\text{p}} + \underbrace{(\text{thermostat})}_{\text{T}} \quad \left(k = \overline{0, P-1}\right) \tag{50}$$

in the "middle" thermostat scheme for designing efficient NM-PIMD algorithms.

3.4 Multi-electronic-state PIMD

Consider a Hamiltonian with N_s electronic states of the form $\hat{\mathbf{H}} = \hat{\mathbf{T}} + \hat{\mathbf{V}}$, where $\hat{\mathbf{V}} = \mathbf{V}(\hat{\mathbf{R}})$ is a $N_s \times N_s$ symmetric diabatic potential energy matrix and $\hat{\mathbf{T}}$ is the kinetic energy matrix. The canonical partition function is defined as

$$Z = \mathrm{Tr}_{ne}[e^{-\beta\hat{\mathbf{H}}}] \tag{51}$$

and a specific physical property of interest is evaluated by

$$\left\langle \hat{\mathbf{B}} \right\rangle = \frac{1}{Z}\mathrm{Tr}_{ne}[e^{-\beta\hat{\mathbf{H}}}\hat{\mathbf{B}}] \tag{52}$$

In Eqs. (51) and (52) the trace is integrated over both the nuclear and electronic degrees of freedom, i.e.

$$\mathrm{Tr}_e[\hat{\mathbf{O}}] = \sum_{n=1}^{N_s} \langle n|\hat{\mathbf{O}}|n\rangle \tag{53}$$

and

$$\mathrm{Tr}_n[\hat{\mathbf{O}}] = \int d\mathbf{R}\langle \mathbf{R}|\hat{\mathbf{O}}|\mathbf{R}\rangle. \tag{54}$$

Substituting the resolution of the identity into Eq. (51) yields

$$Z = \lim_{P\to\infty}\left|\frac{P\mathbf{M}}{2\pi\beta\hbar^2}\right|^{P/2}\int d\mathbf{R}_1\ldots d\mathbf{R}_P \exp\left[-\frac{\beta}{2}\omega_P^2\sum_{i=1}^{P}(\mathbf{R}_i-\mathbf{R}_{i+1})^T\mathbf{M}(\mathbf{R}_i-\mathbf{R}_{i+1})\right]$$
$$\times\,\mathrm{Tr}_e\left[\prod_{i=1}^{P}\mathbf{O}^T(\mathbf{R}_i)\mathbf{O}(\mathbf{R}_i)\right] \tag{55}$$

where $\mathbf{O}(\mathbf{R}_i)$ is related to the splitting scheme. We have studied three typical decomposition schemes, namely, the "diagonalization", "first-order expansion", and "hyperbolic function" methods in Ref. [40].

Because $\mathrm{Tr}_e\left[\prod_{i=1}^{P}\mathbf{O}^T(\mathbf{R}_i)\mathbf{O}(\mathbf{R}_i)\right]$ is not always positive-definite for general multi-electronic-state (MES) systems, regardless of which decomposition scheme is employed, it is not recommended to use either $\mathrm{Tr}_e\left[\prod_{i=1}^{P}\mathbf{O}^T(\mathbf{R}_i)\mathbf{O}(\mathbf{R}_i)\right]$ or its absolute value to define an effective potential function $\phi(\mathbf{R}_1,\ldots,\mathbf{R}_P)$ for performing PIMD. A reasonable effective potential is defined by

$$e^{-\beta\phi^{(\mathrm{dia})}(\mathbf{R}_1,\ldots,\mathbf{R}_P)} = \mathrm{Tr}_e\left[\prod_{i=1}^{P} e^{-\beta\mathbf{V}_{\mathrm{diag}}(\mathbf{R}_i)/P}\right]. \tag{56}$$

Note that the right-hand side of Eq. (56) is always positive-definite. Here $\mathbf{V}_{\mathrm{diag}}(\mathbf{R}_i)$ is a diagonal matrix, whose elements are the diagonal elements of $\mathbf{V}(\mathbf{R})$. The partition function [Eq.(55)] may then be expressed as

$$Z = \lim_{P\to\infty}\left|\frac{P\mathbf{M}}{2\pi\beta\hbar^2}\right|^{P/2}\int d\mathbf{R}_1\ldots d\mathbf{R}_P \exp\left[-\beta U_{\mathrm{eff}}^{(\mathrm{dia})}(\mathbf{R}_1,\ldots,\mathbf{R}_P)\right]\tilde{Z}^{(\mathrm{dia})}(\mathbf{R}_1,\ldots,\mathbf{R}_P), \tag{57}$$

of which the estimator is

$$\tilde{Z}^{(\mathrm{dia})}(\mathbf{R}_1,\ldots,\mathbf{R}_P) = \frac{\mathrm{Tr}_e\left[\prod_{i=1}^{P}\mathbf{O}^T(\mathbf{R}_i)\mathbf{O}(\mathbf{R}_i)\right]}{\mathrm{Tr}_e\left[\prod_{i=1}^{P} e^{-\beta\mathbf{V}_{\mathrm{diag}}(\mathbf{R}_i)/P}\right]} \tag{58}$$

and

$$U_{\mathrm{eff}}^{(\mathrm{dia})}(\mathbf{R}_1,\ldots,\mathbf{R}_P) = \frac{1}{2}\omega_P^2\sum_{i=1}^{P}(\mathbf{R}_i-\mathbf{R}_{i+1})^T P\mathbf{M}(\mathbf{R}_i-\mathbf{R}_{i+1}) + \phi^{(\mathrm{dia})}(\mathbf{R}_1,\ldots,\mathbf{R}_P). \tag{59}$$

Then one can obtain any specific physical property of interest in Eq. (52) from

$$\left\langle\hat{\mathbf{B}}\right\rangle = \frac{\int d\mathbf{R}_1\ldots d\mathbf{R}_P \exp\left[-\beta U_{\mathrm{eff}}^{(\mathrm{dia})}(\mathbf{R}_1,\ldots,\mathbf{R}_P)\right]\tilde{B}^{(\mathrm{dia})}(\mathbf{R}_1,\ldots,\mathbf{R}_P)}{\int d\mathbf{R}_1\ldots d\mathbf{R}_P \exp\left[-\beta U_{\mathrm{eff}}^{(\mathrm{dia})}(\mathbf{R}_1,\ldots,\mathbf{R}_P)\right]\tilde{Z}^{(\mathrm{dia})}(\mathbf{R}_1,\ldots,\mathbf{R}_P)} \tag{60}$$

The estimators $\tilde{B}^{(\mathrm{dia})}(\mathbf{R}_1,\ldots,\mathbf{R}_P)$ in the diabatic representation for some typical operators are described in Ref. [40].

Define an effective Hamiltonian

$$H_{\mathrm{eff}}^{(\mathrm{dia})}(\mathbf{R}_1,\ldots,\mathbf{R}_P;\mathbf{p}_1,\ldots,\mathbf{p}_P) = \sum_{i=1}^{P}\frac{1}{2}\mathbf{p}_i^T\tilde{\mathbf{M}}_i^{-1}\mathbf{p}_i + U_{\mathrm{eff}}^{(\mathrm{dia})}(\mathbf{R}_1,\ldots,\mathbf{R}_P) \tag{61}$$

with the fictitious masses $\tilde{\mathbf{M}}_i$ and momenta $\mathbf{p}_i$. It leads to the MES-PIMD equations of motion

$$\begin{aligned}\dot{\mathbf{R}}_i &= \tilde{\mathbf{M}}_i^{-1}\mathbf{p}_i\\ \dot{\mathbf{p}}_i &= -\tfrac{\partial}{\partial\mathbf{R}_i}U_{\mathrm{eff}}^{(\mathrm{dia})}(\mathbf{R}_1,\ldots,\mathbf{R}_P)\end{aligned}\quad\left(i=\overline{1,P}\right). \tag{62}$$

Their coupling to a thermostat produces a proper canonical distribution for $(\mathbf{R}_1, \ldots, \mathbf{R}_P; \mathbf{p}_1, \ldots, \mathbf{p}_P)$, which changes Eq. (60) into

$$\left\langle \hat{\mathbf{B}} \right\rangle = \lim_{P\to\infty} \frac{\int \left(\prod_{i=1}^{P} d\mathbf{R}_i d\mathbf{p}_i \right) \exp\left\{ -\beta H_{\text{eff}}^{(\text{dia})}(\mathbf{R}_1, \ldots, \mathbf{R}_P; \mathbf{p}_1, \ldots, \mathbf{p}_P) \right\} \tilde{B}^{(\text{dia})}(\mathbf{R}_1, \ldots, \mathbf{R}_P)}{\int \left(\prod_{i=1}^{P} d\mathbf{R}_i d\mathbf{p}_i \right) \exp\left\{ -\beta H_{\text{eff}}^{(\text{dia})}(\mathbf{R}_1, \ldots, \mathbf{R}_P; \mathbf{p}_1, \ldots, \mathbf{p}_P) \right\} \tilde{Z}^{(\text{dia})}(\mathbf{R}_1, \ldots, \mathbf{R}_P)}. \quad (63)$$

Similar to Eq. (37), the MES-PIMD equations of motion are decomposed into three parts

$$\begin{pmatrix} \dot{\mathbf{R}}_i \\ \dot{\mathbf{p}}_i \end{pmatrix} = \underbrace{\begin{pmatrix} \tilde{\mathbf{M}}_i^{-1}\mathbf{p}_i \\ 0 \end{pmatrix}}_{\mathbf{x}} + \underbrace{\begin{pmatrix} 0 \\ -\frac{\partial U_{\text{eff}}^{(\text{dia})}}{\partial \mathbf{R}_i} \end{pmatrix}}_{\mathbf{p}} + \underbrace{(\text{Thermostat})}_{T} \qquad \left(i = \overline{1, P}\right). \quad (64)$$

The "middle" scheme then yields efficient MES-PIMD algorithms. As discussed in Sects. 3.2 and 3.3, the staging or normal-mode transformation of path integral beads can be used in Eq. (64).

We have investigated a seven-state system in Ref. [40]. It is shown in Fig. 4 that the "middle" thermostat scheme greatly improves the efficiency over the conventional "side" scheme.

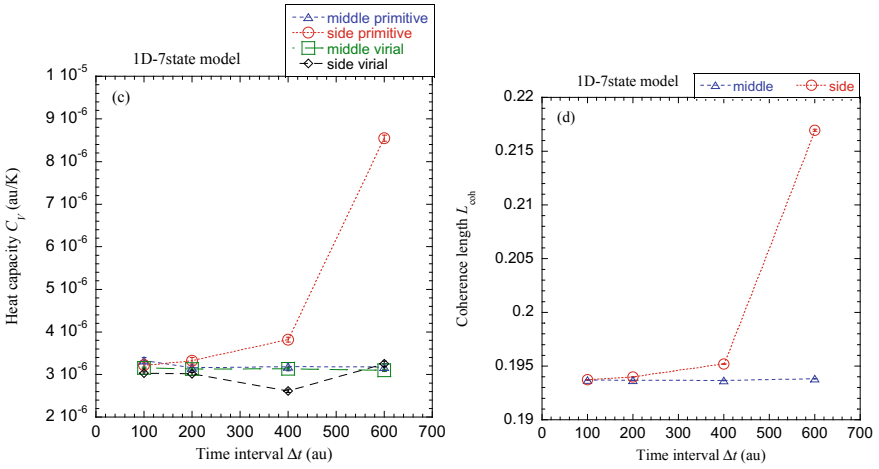

Fig. 4 Results for the "middle" scheme in comparison to those for conventional thermostat schemes for MES-PIMD for a 1-D seven-state system. Results for the heat capacity and coherence length are plotted as functions of the time interval Δt. Atomic units (au) are used (Reproduced with permission from Ref. [40])

4 "Middle" scheme with constraints

4.1 Holonomic constraint

Define the holonomic constraint

$$\boldsymbol{\sigma}(\mathbf{x}) = \mathbf{0}, \tag{65}$$

where $\boldsymbol{\sigma}(\mathbf{x})$ is n_c-dimensional vector function of the configuration $\mathbf{x}$. Its derivative yields the constraint for the momentum $\mathbf{p}$

$$\frac{d}{dt}\boldsymbol{\sigma}(\mathbf{x}) = \left(\frac{\partial \boldsymbol{\sigma}}{\partial \mathbf{x}}\right)^T \mathbf{M}^{-1}\mathbf{p} = \mathbf{0}. \tag{66}$$

SHAKE [53] and RATTLE [54] are two typical algorithms for applying constraints to molecular systems. While SHAKE guarantees only the position constraint, RATTLE satisfies both the position and momentum constraints. It is straightforward to employ the "middle" scheme with the SHAKE or RATTLE algorithm for sampling the canonical (NVT) ensemble. Various versions can be constructed to guarantee that the position and momentum satisfy Eqs. (65)–(66) at the end of a time step. Our recommended "VV-Middle" scheme with holonomic constraints is

$$\tilde{\mathbf{p}}\left(\frac{\Delta t}{2}\right) \leftarrow \mathbf{p}(0) - \frac{\partial U}{\partial \mathbf{x}(0)}\frac{\Delta t}{2}$$

$$\mathrm{C}_2: \begin{cases} \text{Solve } \boldsymbol{\mu}: \left(\frac{\partial \boldsymbol{\sigma}}{\partial \mathbf{x}(0)}\right)^T \mathbf{M}^{-1}\left(\tilde{\mathbf{p}}\left(\frac{\Delta t}{2}\right) + \frac{\partial \boldsymbol{\sigma}}{\partial \mathbf{x}(0)}\boldsymbol{\mu}\right) = \mathbf{0} \\ \tilde{\tilde{\mathbf{p}}}\left(\frac{\Delta t}{2}\right) \leftarrow \tilde{\mathbf{p}}\left(\frac{\Delta t}{2}\right) + \frac{\partial \boldsymbol{\sigma}}{\partial \mathbf{x}(0)}\boldsymbol{\mu} \end{cases}$$

$$\tilde{\mathbf{x}}\left(\frac{\Delta t}{2}\right) \leftarrow \mathbf{x}(0) + \mathbf{M}^{-1}\tilde{\tilde{\mathbf{p}}}\left(\frac{\Delta t}{2}\right)\frac{\Delta t}{2}$$

Thermostat for a full time step Δt *(in which* $\tilde{\tilde{\mathbf{p}}}\left(\frac{\Delta t}{2}\right)$ *is updated)*

$$\tilde{\mathbf{x}}(\Delta t) \leftarrow \tilde{\mathbf{x}}\left(\frac{\Delta t}{2}\right) + \mathbf{M}^{-1}\tilde{\tilde{\mathbf{p}}}\left(\frac{\Delta t}{2}\right)\frac{\Delta t}{2}$$

$$\mathrm{C}_1: \begin{cases} \text{Solve } \boldsymbol{\lambda}: \boldsymbol{\sigma}\left(\tilde{\mathbf{x}}(\Delta t) + \mathbf{M}^{-1}\frac{\partial \boldsymbol{\sigma}}{\partial \mathbf{x}(0)}\boldsymbol{\lambda}\right) = \mathbf{0} \\ \mathbf{x}(\Delta t) \leftarrow \tilde{\mathbf{x}}(\Delta t) + \mathbf{M}^{-1}\frac{\partial \boldsymbol{\sigma}}{\partial \mathbf{x}(0)}\boldsymbol{\lambda} \\ \mathbf{p}\left(\frac{\Delta t}{2}\right) \leftarrow \tilde{\tilde{\mathbf{p}}}\left(\frac{\Delta t}{2}\right) + \frac{1}{\Delta t}\frac{\partial \boldsymbol{\sigma}}{\partial \mathbf{x}(0)}\boldsymbol{\lambda} \end{cases}$$

$$\tilde{\mathbf{p}}(\Delta t) \leftarrow \mathbf{p}\left(\frac{\Delta t}{2}\right) - \frac{\partial U}{\partial \mathbf{x}(\Delta t)}\frac{\Delta t}{2}$$

$$\mathrm{C}_2: \begin{cases} \text{Solve } \boldsymbol{\mu}: \left(\frac{\partial \boldsymbol{\sigma}}{\partial \mathbf{x}(\Delta t)}\right)^T \mathbf{M}^{-1}\left(\tilde{\mathbf{p}}(\Delta t) + \frac{\partial \boldsymbol{\sigma}}{\partial \mathbf{x}(\Delta t)}\boldsymbol{\mu}\right) = \mathbf{0} \\ \mathbf{p}(\Delta t) \leftarrow \tilde{\mathbf{p}}(\Delta t) + \frac{\partial \boldsymbol{\sigma}}{\partial \mathbf{x}(\Delta t)}\boldsymbol{\mu} \end{cases}, \tag{67}$$

which is denoted "C2-p-C1-x-T-x-C2-p", where the operations are performed from right to left. When "LF-Middle" is used, our recommended version is

$$
\begin{aligned}
&\tilde{\mathbf{p}}\left(\frac{\Delta t}{2}\right) \leftarrow \mathbf{p}\left(-\frac{\Delta t}{2}\right) - \frac{\partial U}{\partial \mathbf{x}(0)}\Delta t \\
&\mathrm{C}_2 : \begin{cases} \text{Solve } \boldsymbol{\mu} : \left(\frac{\partial \boldsymbol{\sigma}}{\partial \mathbf{x}(0)}\right)^T \mathbf{M}^{-1}\left(\tilde{\mathbf{p}}\left(\frac{\Delta t}{2}\right) + \frac{\partial \boldsymbol{\sigma}}{\partial \mathbf{x}(0)}\boldsymbol{\mu}\right) = \mathbf{0} \\ \tilde{\tilde{\mathbf{p}}}\left(\frac{\Delta t}{2}\right) \leftarrow \tilde{\mathbf{p}}\left(\frac{\Delta t}{2}\right) + \frac{\partial \boldsymbol{\sigma}}{\partial \mathbf{x}(0)}\boldsymbol{\mu} \end{cases} \\
&\tilde{\mathbf{x}}\left(\frac{\Delta t}{2}\right) \leftarrow \mathbf{x}(0) + \mathbf{M}^{-1}\tilde{\tilde{\mathbf{p}}}\left(\frac{\Delta t}{2}\right)\frac{\Delta t}{2} \\
&\textit{Thermostat for a full time step } \Delta t \textit{ (in which } \tilde{\tilde{\mathbf{p}}}\left(\frac{\Delta t}{2}\right) \textit{ is updated)} \\
&\tilde{\mathbf{x}}(\Delta t) \leftarrow \tilde{\mathbf{x}}\left(\frac{\Delta t}{2}\right) + \mathbf{M}^{-1}\tilde{\tilde{\mathbf{p}}}\left(\frac{\Delta t}{2}\right)\frac{\Delta t}{2} \\
&\mathrm{C}_1 : \begin{cases} \text{Solve } \boldsymbol{\lambda} : \boldsymbol{\sigma}\left(\tilde{\mathbf{x}}(\Delta t) + \mathbf{M}^{-1}\frac{\partial \boldsymbol{\sigma}}{\partial \mathbf{x}(0)}\boldsymbol{\lambda}\right) = \mathbf{0} \\ \mathbf{x}(\Delta t) \leftarrow \tilde{\mathbf{x}}(\Delta t) + \mathbf{M}^{-1}\frac{\partial \boldsymbol{\sigma}}{\partial \mathbf{x}(0)}\boldsymbol{\lambda} \\ \tilde{\tilde{\mathbf{p}}}\left(\frac{\Delta t}{2}\right) \leftarrow \tilde{\tilde{\mathbf{p}}}\left(\frac{\Delta t}{2}\right) + \frac{1}{\Delta t}\frac{\partial \boldsymbol{\sigma}}{\partial \mathbf{x}(0)}\boldsymbol{\lambda} \end{cases} \\
&\mathrm{C}_2 : \begin{cases} \text{Solve } \boldsymbol{\mu} : \left(\frac{\partial \boldsymbol{\sigma}}{\partial \mathbf{x}(\Delta t)}\right)^T \mathbf{M}^{-1}\left(\tilde{\tilde{\mathbf{p}}}\left(\frac{\Delta t}{2}\right) + \frac{\partial \boldsymbol{\sigma}}{\partial \mathbf{x}(\Delta t)}\boldsymbol{\mu}\right) = \mathbf{0} \\ \mathbf{p}\left(\frac{\Delta t}{2}\right) \leftarrow \tilde{\tilde{\mathbf{p}}}\left(\frac{\Delta t}{2}\right) + \frac{\partial \boldsymbol{\sigma}}{\partial \mathbf{x}(\Delta t)}\boldsymbol{\mu} \end{cases}
\end{aligned}
\tag{68}
$$

We denote it "C2-C1-x-T-x-C2-p".

We use AMBER2018 [55] to simulate a liquid water system subject to intramolecular bond length and angle constraints. The system contains a box of 216 water molecules (with periodic boundary conditions). The TIP3P model are employed as the force field. Fig. 5a implies that the two "middle" schemes (Eqs. 67 and 68) perform better in configurational sampling than the conventional "side" scheme as well as the BBK algorithm [25] for systems subject to constraints. Fig. 5b shows that "LF-Middle" leads to a more accurate marginal momentum distribution than "VV-Middle".

4.2 Isokinetic constraints in the "middle" scheme

Nonholonomic constraints (that involve the constraint of momenta) are also widely used in MD simulation in the canonical ensemble. In Ref. [12] we have recently extended the application of the "middle" thermostat scheme to the isokinetic constraint as well as to the MTS technique, which leads to a more efficient version for the Stochastic-Iso-NH-RESPA [SIN(R)] algorithm [56].

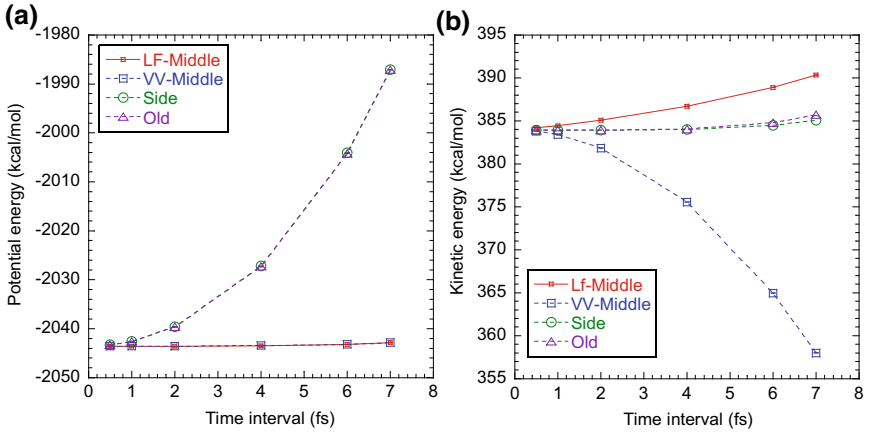

Fig. 5 MD results of liquid water with intramolecular O–H bond length and H–O–H angle constraints at $T = 298.15$ K using different schemes, **a** average potential energy per atom $\langle U(\mathbf{x})\rangle/(N_{\text{atom}}k_B)$ (unit: Kelvin) **b** average kinetic energy per atom $\langle \mathbf{p}^T\mathbf{M}^{-1}\mathbf{p}\rangle/(2N_{\text{atom}}k_B)$ (unit: Kelvin). "Old" stands for the BBK algorithm [25] used in AMBER (Reproduced with permission from Ref. [11])

Given $\Delta t = n\delta t$, with Δt as the outer time interval and δt the inner time interval, respectively. The propagating order of original SIN(R) [56] is expressed as

$$\begin{aligned} e^{\mathcal{L}\Delta t} \approx\ & e^{\mathcal{L}_N\delta t/2}e^{\mathcal{L}_{\mathbf{p}}^{(f)}\delta t/2+\mathcal{L}_{\mathbf{p}}^{(s)}\Delta t/2}e^{\mathcal{L}_{\mathbf{x}}\delta t/2}e^{\mathcal{L}_O\delta t}e^{\mathcal{L}_{\mathbf{x}}\delta t/2}e^{\mathcal{L}_{\mathbf{p}}^{(f)}\delta t/2}e^{\mathcal{L}_N\delta t/2} \\ & \times\left(e^{\mathcal{L}_N\delta t/2}e^{\mathcal{L}_{\mathbf{p}}^{(f)}\delta t/2}e^{\mathcal{L}_{\mathbf{x}}\delta t/2}e^{\mathcal{L}_O\delta t}e^{\mathcal{L}_{\mathbf{x}}\delta t/2}e^{\mathcal{L}_{\mathbf{p}}^{(f)}\delta t/2}e^{\mathcal{L}_N\delta t/2}\right)^{n-2} \\ & \times e^{\mathcal{L}_N\delta t/2}e^{\mathcal{L}_{\mathbf{p}}^{(f)}\delta t/2}e^{\mathcal{L}_{\mathbf{x}}\delta t/2}e^{\mathcal{L}_O\delta t}e^{\mathcal{L}_{\mathbf{x}}\delta t/2}e^{\mathcal{L}_{\mathbf{p}}^{(f)}\delta t/2+\mathcal{L}_{\mathbf{p}}^{(s)}\Delta t/2}e^{\mathcal{L}_N\delta t/2}, \end{aligned} \tag{69}$$

where the Kolmogorov operators in Eq. (69) are explicitly defined in Ref. [12]. Refs. [9, 11] have already suggested that applying thermostat in the middle of the propagation step always leads to much better performance in sampling the configuration space. When the middle thermostat scheme is used to design new SIN(R) algorithms, the propagating order of the "VV-Middle" version is

$$\begin{aligned} e^{\mathcal{L}\Delta t} \approx\ & e^{\mathcal{L}_{\mathbf{p}}^{(f)}\delta t/2+\mathcal{L}_{\mathbf{p}}^{(s)}\Delta t/2}e^{\mathcal{L}_{\mathbf{x}}\delta t/2}e^{\mathcal{L}_N\delta t/2}e^{\mathcal{L}_O\delta t}e^{\mathcal{L}_N\delta t/2}e^{\mathcal{L}_{\mathbf{x}}\delta t/2}e^{\mathcal{L}_{\mathbf{p}}^{(f)}\delta t/2} \\ & \times\left(e^{\mathcal{L}_{\mathbf{p}}^{(f)}\delta t/2}e^{\mathcal{L}_{\mathbf{x}}\delta t/2}e^{\mathcal{L}_N\delta t/2}e^{\mathcal{L}_O\delta t}e^{\mathcal{L}_N\delta t/2}e^{\mathcal{L}_{\mathbf{x}}\delta t/2}e^{\mathcal{L}_{\mathbf{p}}^{(f)}\delta t/2}\right)^{n-2} \\ & \times e^{\mathcal{L}_{\mathbf{p}}^{(f)}\delta t/2}e^{\mathcal{L}_{\mathbf{x}}\delta t/2}e^{\mathcal{L}_N\delta t/2}e^{\mathcal{L}_O\delta t}e^{\mathcal{L}_N\delta t/2}e^{\mathcal{L}_{\mathbf{x}}\delta t/2}e^{\mathcal{L}_{\mathbf{p}}^{(f)}\delta t/2+\mathcal{L}_{\mathbf{p}}^{(s)}\Delta t/2}, \end{aligned} \tag{70}$$

[denoted "VV-Middle-SIN(R)"] and that of the "LF-Middle" version becomes

$$e^{\mathcal{L}\Delta t} \approx e^{\mathcal{L}_{\mathbf{x}}\delta t/2}e^{\mathcal{L}_N\delta t/2}e^{\mathcal{L}_O\delta t}e^{\mathcal{L}_N\delta t/2}e^{\mathcal{L}_{\mathbf{x}}\delta t/2}e^{\mathcal{L}_{\mathbf{p}}^{(f)}\delta t}$$

$$\times \left(e^{\mathcal{L}_{\mathbf{x}}\delta t/2} e^{\mathcal{L}_N \delta t/2} e^{\mathcal{L}_O \delta t} e^{\mathcal{L}_N \delta t/2} e^{\mathcal{L}_{\mathbf{x}}\delta t/2} e^{\mathcal{L}_{\mathbf{p}}^{(f)}\delta t} \right)^{n-2}$$
$$\times e^{\mathcal{L}_{\mathbf{x}}\delta t/2} e^{\mathcal{L}_N \delta t/2} e^{\mathcal{L}_O \delta t} e^{\mathcal{L}_N \delta t/2} e^{\mathcal{L}_{\mathbf{x}}\delta t/2} e^{\mathcal{L}_{\mathbf{p}}^{(f)}\delta t + \mathcal{L}_{\mathbf{p}}^{(s)}\Delta t} \tag{71}$$

[denoted "LF-Middle-SIN(R)"]. Eqs. (70) and (71) yield the same efficiency and accuracy for sampling the configuration space. These versions in the "middle" scheme are all easily extended with more than two time steps for SIN(R).

We have simulated a system of liquid water [216 H2O molecules in a box with periodic boundary conditions at the state point $T = 298.15$ K(temperature) and 0.997 g·cm^{-3} (density)] to compare the original version of SIN(R) [Eq. (69)] and its new version with the "middle" scheme ("VV-Middle-SIN(R)") [Eq. (70)]. POLI2VS-a polarizable and flexible force field [57] for liquid water is used. In MTS, one may treat the force contributed by the intramolecular interactions as the fast part and consider the force derived from the noncovalent interactions as the slow part. Fig. 6 demonstrates the results with the inner time interval δt varied from 0.1 to 2fs while the outer time interval Δt are fixed at 2 or 4fs. As the inner time step δt increases, "VV-Middle-SIN(R)" performs more efficiently than conventional SIN(R).

All versions in the "middle" scheme for MD in Sect. 4.1 (with holonomic constraints) and Sect. 4.2 (with nonholonomic constraints as well as resonance-free multiple time-step techniques) can easily be extended to PIMD. We have already applied some of them in PIMD simulations for liquid water [10].

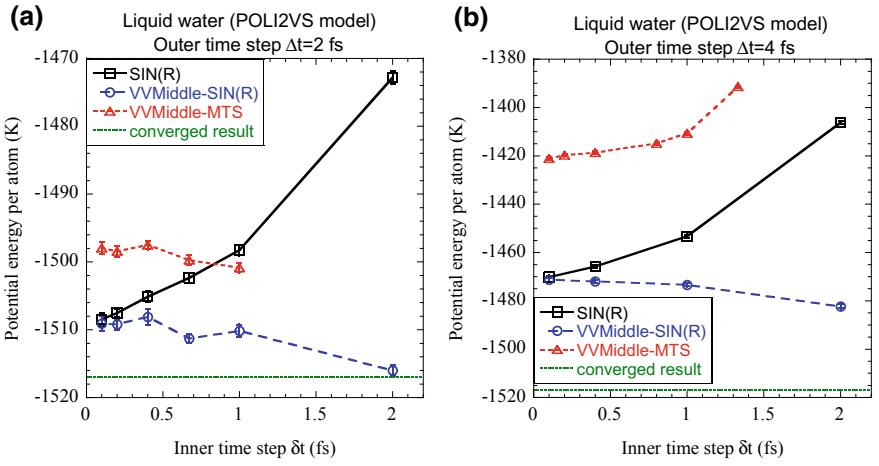

Fig. 6 Average potential energy per atom $\langle U(\mathbf{x})\rangle/(N_{\text{atom}}k_B)$ (unit: Kelvin) of liquid water at $T = 298.15$ K as a function of the inner time step δt with a fixed outer time step. **a** $\Delta t = 2$ fs **b** $\Delta t = 4$ fs (Reproduced with permission from Ref. [12])

5 Conclusions

In many remarks, the "middle" thermostat scheme [7, 9, 11, 12, 15, 16] (either "VV-Middle" or "LF-Middle") provides a promising approach to design efficient MD/PIMD algorithms for sampling the configuration space of the canonical ensemble with or without constraints. Combination of the "middle" scheme with resonance-free MTS techniques leads to more efficient and robust algorithms for sampling the configuration space for multi-time-scale systems [12].

It is straightforward to employ the "middle" thermostat scheme with various advanced enhanced sampling techniques [55, 58–68] to accelerate configurational sampling for molecular systems where rare events become important [12]. Since the "middle" thermostat scheme is useful for any types of potential energy surfaces or force fields (for realistic molecular systems), it will be helpful to implement the "middle" scheme for ab initio MD or ab initio PIMD[21, 69] to reduce the computation cost. It will be interesting to develop more efficient MD and PIMD algorithms in the unified theoretical framework for these purposes.

Acknowledgements This work was supported by the Ministry of Science and Technology of China (MOST) Grants No. 2016YFC0202803 and No. 2017YFA0204901, by the National Natural Science Foundation of China (NSFC) Grants No. 21373018 and No. 21573007, by the Recruitment Program of Global Experts, by Specialized Research Fund for the Doctoral Program of Higher Education No. 20130001110009, and by Special Program for Applied Research on Super Computation of the NSFC-Guangdong Joint Fund (the second phase) under Grant No. U1501501.

References

1. Allen MP, Tildesley DJ (1989) Computer simulation of liquids. Clarendon Press
2. Frenkel D, Smit B (2002) Understanding molecular simulation, 2nd edn. Academic Press, San Diego
3. Chandler D, Wolynes PG (1981) Exploiting the isomorphism between quantum theory and the classical statistical mechanics of polyatomic fluids. J. Chem. Phys. 74(7):4078–4095
4. Parrinello M, Rahman A (1984) Study of an f center in molten kcl. J. Chem. Phys. 80(2):860–867
5. Berne BJ, Thirumalai D (1986) On the simulation of quantum systems: path integral methods. Annu Rev Phys Chem 37:401–424
6. Tuckerman ME (2010) Statistical mechanics: theory and molecular simulation. Oxford University Press, New York
7. Liu J, Li D, Liu X (2016) A simple and accurate algorithm for path integral molecular dynamics with the langevin thermostat. J Chem Phys 145(2):024103
8. Markland TE, Ceriotti M (2018) Nuclear quantum effects enter the mainstream. Nat Rev Chem 2(3):14
9. Zhang Z, Liu X, Chen Z, Zheng H, Yan K, Liu J (2017) A unified thermostat scheme for efficient configurational sampling for classical/quantum canonical ensembles via molecular dynamics. J Chem Phys 147(3):034109
10. Liu X, Liu J (2018) Critical role of quantum dynamical effects in the raman spectroscopy of liquid water. Mol Phys 116(7–8):755–779
11. Zhang Z, Yan K, Liu X, Liu J (2018) A leap-frog algorithm-based efficient unified thermostat scheme for molecular dynamics. Chin Sci Bull 63(0023–074X):3467

12. Zhang Z, Liu X, Yan K, Tuckerman ME, Liu J (2019) Unified efficient thermostat scheme for the canonical ensemble with holonomic or isokinetic constraints via molecular dynamics. J Phys Chem A 123(28):6056–6079
13. Liu J, Li D, Liu X (2016) Further study of path integral liouville dynamics
14. Liu J, Zhang Z (2016) Path integral liouville dynamics: Applications to infrared spectra of oh, water, ammonia, and methane. J Chem Phys 144(3):034307
15. Li D, Han X, Chai Y, Wang C, Zhang Z, Chen Z, Liu J, Shao J (2017) Stationary state distribution and efficiency analysis of the langevin equation via real or virtual dynamics. J Chem Phys 147(18):184104
16. Li D-z, Chen Z-f, Zhang Z-j, Liu J (2017) Understanding molecular dynamics with stochastic processes via real or virtual dynamics. Chin J Chem Phys 30(6):735–760
17. Case DA, Ben-Shalom IY, Brozell SR, Cerutti DS, Cheatham TE III, Cruzeiro VWD, Darden TA, Duke RE, Ghoreishi D, Gilson MK, Gohlke H, Goetz AW, Greene D, Harris R, Homeyer N, Izadi S, Kovalenko A, Kurtzman T, Lee TS, LeGrand S, Li P, Lin C, Liu J, Luchko T, Luo R, Mermelstein DJ, Merz KM, Miao Y, Monard G, Nguyen C, Nguyen H, Omelyan I, Onufriev A, Pan F, Qi R, Roe DR, Roitberg A, Sagui C, Schott-Verdugo S, Shen J, Simmerling CL, Smith J, Salomon-Ferrer R, Swails J, Walker RC, Wang J, Wei H, Wolf RM, Wu X, Xiao L, York DM (2018) Kollman PA Amber 2018. University of California, San Francisco
18. Andersen HC (1980) Molecular dynamics simulations at constant pressure and/or temperature. Journal of Chemical Physics 72(4):2384–2393
19. Nosé S (1984) A molecular dynamics method for simulations in the canonical ensemble. Mol Phys 52(2):255–268
20. Hoover WG (1985) Canonical dynamics: equilibrium phase-space distributions. Phys Rev A 31(3):1695–1697
21. Martyna GJ, Klein ML, Tuckerman M (1992) Nosé–hoover chains: The canonical ensemble via continuous dynamics. J Chem Phys 97(4):2635–2643
22. Martyna GJ, Tuckerman ME, Tobias DJ, Klein ML (1996) Explicit reversible integrators for extended systems dynamics. Mol Phys 87(5):1117–1157
23. Tuckerman M, Berne BJ, Martyna GJ (1992) Reversible multiple time scale molecular dynamics. J Chem Phys 97(3):1990–2001
24. Andrea TA, Swope WC, Andersen HC (1983) The role of long ranged forces in determining the structure and properties of liquid water. J Chem Phys 79(9):4576–4584
25. Brünger A, Brooks Iii CL, Karplus M (1984) Stochastic boundary conditions for molecular dynamics simulations of st2 water. Chem Phys Lett 105(5):495–500
26. Goga N, Rzepiela AJ, de Vries AH, Marrink SJ, Berendsen HJC (2012) Efficient algorithms for langevin and dpd dynamics. J Chem Theory Comput 8(10):3637–3649
27. Bussi G, Parrinello M (2007) Accurate sampling using langevin dynamics. Phys Rev E 75(5):056707
28. Grønbech-Jensen N, Farago O (2013) A simple and effective verlet-type algorithm for simulating langevin dynamics. Mol Phys 111(8):983–991
29. Leimkuhler B, Matthews C (2012) Rational construction of stochastic numerical methods for molecular sampling. Appl Math Res Express 2013(1):34–56
30. Leimkuhler B, Matthews C (2013) Robust and efficient configurational molecular sampling via langevin dynamics. J Chem Phys 138(17):174102
31. Leimkuhler B, Matthews C (2016) Efficient molecular dynamics using geodesic integration and solvent–solute splitting. Proc R Soc A: Math, Phys Eng Sci 472(2189)
32. Leimkuhler B, Matthews C (2015) Molecular dynamics with deterministic and stochastic numerical methods. Springer
33. Hall R, Berne BJ (1984) Nonergodicity in path integral molecular dynamics. J. Chem. Phys. 81(8):3641–3643
34. Gillan MJ (1987) Quantum simulation of hydrogen in metals. Phys Rev Lett 58(6):563–566
35. Singer K, Smith W (1988) Path integral simulations of condensed phase lennard-jones systems. Mol Phys 64(6):1215–1231

36. Müser MH (2002) On new efficient algorithms for pimc and pimd. Comput Phys Commun 147(1–2):83–86
37. Drozdov AN, Talkner P (1998) Path integrals for fokker-planck dynamics with singular diffusion: accurate factorization for the time evolution operator. J Chem Phys 109(6):2080–2091
38. Tuckerman ME, Marx D, Klein ML, Parrinello M (1996) Efficient and general algorithms for path integral car-parrinello molecular dynamics. J Chem Phys 104(14):5579–5588
39. Ceriotti M, Parrinello M, Markland TE, Manolopoulos DE (2010) Efficient stochastic thermostatting of path integral molecular dynamics. J Chem Phys 133(12):124104
40. Liu X, Liu J (2018) Path integral molecular dynamics for exact quantum statistics of multi-electronic-state systems. J Chem Phys 148(10):102319
41. Wang H, Liu X, Liu J (2018) Accurate calculation of equilibrium reduced density matrix for the system-bath model: A multilayer multiconfiguration time-dependent hartree approach and its comparison to a multi-electronic-state path integral molecular dynamics approach. Chin J Chem Phys 31(4):446–456
42. Suzuki M (1985) Decomposition formulas of exponential operators and lie exponentials with some applications to quantum mechanics and statistical physics. J Math Phys 26(4):601–612
43. Yoshida H (1990) Construction of higher order symplectic integrators. Phys Lett A 150(5):262–268
44. Suzuki M (1991) General theory of fractal path integrals with applications to many-body theories and statistical physics. J Math Phys 32(2):400–407
45. Ceperley DM (1995) Path integrals in the theory of condensed helium. Rev Mod Phys 67(2):279–355
46. Feynman RP (1953) Atomic theory of the lambda-transition in helium. Phys Rev 91(6):1291–1301
47. Tuckerman ME, Berne BJ, Martyna GJ, Klein ML (1993) Efficient molecular-dynamics and hybrid monte-carlo algorithms for path-integrals. J Chem Phys 99(4):2796–2808
48. Herman MF, Bruskin EJ, Berne BJ (1982) On path integral monte-carlo simulations. J Chem Phys 76(10):5150–5155
49. Pollock EL, Ceperley DM (1984) Simulation of quantum many-body systems by path-integral methods. Phys. Rev. B 30(5):2555–2568
50. Liu J, Li D, Liu X (2016) Supplementary material for the paper 'a simple and accurate algorithm for path integral molecular dynamics'. J Chem Phys 145:024103. ftp://ftp.aip.org/epaps/journ_chem_phys/E-JCPSA6-145-007626
51. Feynman RP, Hibbs AR (1965) Quantum mechanics and path integrals. McGraw-Hill, New York
52. Cao J, Berne BJ (1993) A Born-Oppenheimer approximation for path-integrals with an application to electron solvation in polarizable fluids. J Chem Phys 99(4):2902–2916
53. Ryckaert J-P, Ciccotti G, Berendsen HJC (1977) Numerical integration of the cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes. J Comput Phys 23(3):327–341
54. Andersen HC (1983) Rattle: A "velocity" version of the shake algorithm for molecular dynamics calculations. J Comput Phys 52(1):24–34
55. Hamelberg D, Mongan J, McCammon JA (2004) Accelerated molecular dynamics: A promising and efficient simulation method for biomolecules. J Chem Phys 120(24):11919–11929
56. Leimkuhler B, Margul DT, Tuckerman ME (2013) Stochastic, resonance-free multiple time-step algorithm for molecular dynamics with very large time steps. Mol Phys 111(22–23):3579–3594
57. Hasegawa T, Tanimura Y (2011) A polarizable water model for intramolecular and intermolecular vibrational spectroscopies. J Phys Chem B 115(18):5545–5553
58. Torrie GM, Valleau JP (1977) Nonphysical sampling distributions in monte carlo free-energy estimation: Umbrella sampling. J Comput Phys 23(2):187–199
59. Laio A, Parrinello M (2002) Escaping free-energy minima. Proc Natl Acad Sci 99(20):12562
60. Gao YQ (2008) An integrate-over-temperature approach for enhanced sampling. J Chem Phys 128(6):064105

61. Valsson O, Parrinello M (2014) Variational approach to enhanced sampling and free energy calculations. Phys Rev Lett 113(9):090601
62. Valsson O, Tiwary P, Parrinello M (2016) Enhancing important fluctuations: Rare events and metadynamics from a conceptual viewpoint. Annu Rev Phys Chem 67(1):159–184
63. Sugita Y, Okamoto Y (1999) Replica-exchange molecular dynamics method for protein folding. Chem Phys Lett 314(1):141–151
64. Peters B (2017) Reaction rate theory and rare events simulations. Elsevier, Amsterdam, Netherlands
65. Carter EA, Ciccotti G, Hynes JT, Kapral R (1989) Constrained reaction coordinate dynamics for the simulation of rare events. Chem Phys Lett 156(5):472–477
66. Sprik M, Ciccotti G (1998) Free energy from constrained molecular dynamics. J Chem Phys 109(18):7737–7744
67. Sergi A, Ciccotti G, Falconi M, Desideri A, Ferrario M (2002) Effective binding force calculation in a dimeric protein by molecular dynamics simulation. J Chem Phys 116(14):6329–6338
68. Bello-Rivas JM, Elber R (2015) Exact milestoning. J Chem Phys 142(9):094102
69. Anandakrishnan R, Drozdetski A, Walker Ross C, Onufriev Alexey V (2015) Speed of conformational change: Comparing explicit and implicit solvent molecular dynamics simulations. Biophys J 108(5):1153–1164