JCTC Journal of Chemical Theory and Computation

pubs.acs.org/JCTC Article

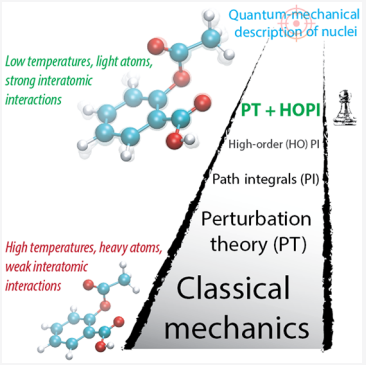
Accurate Description of Nuclear Quantum Effects with High-Order Perturbed Path Integrals (HOPPI)

Igor Poltavsky, Venkat Kapil, Michele Ceriotti, Kwang S. Kim, and Alexandre Tkatchenko*

Cite This: J. Chem. Theory Comput. 2020, 16, 1128−1135 Read Online

ACCESS | Metrics & More | Article Recommendations

ABSTRACT: Imaginary time path-integral (PI) simulations that account for nuclear quantum effects (NQE) beyond the harmonic approximation are increasingly employed together with modern electronic-structure calculations. Existing PI methods are applicable to molecules, liquids, and solids; however, the computational cost of such simulations increases dramatically with decreasing temperature. To address this challenge, here, we propose to combine high-order PI factorization with perturbation theory (PT). Already for conventional second-order PI simulations, the PT ansatz increases the accuracy 2-fold compared to fourth-order schemes with the same settings. In turn, applying PT to high-order path integrals (HOPI) further improves the efficiency of simulations for molecular and condensed matter systems especially at low temperatures. We present results for bulk liquid water, the aspirin molecule, and the CH_5^+ molecule. Perturbed HOPI simulations remain both efficient and accurate down to 20 K and provide a convenient method to estimate the convergence of quantum-mechanical observables.



1. INTRODUCTION

The nonclassical behavior of nuclei has been shown to play an important role for many systems of practical interest.[1−15] Nuclear quantum effects (NQE) affect the stability of biomolecules[16] and molecular crystals,[17,18] ion tunneling through low-dimensional materials,[9] the phase diagram of high pressure melts,[19] and fluctuations of carbon atoms in graphene,[20] among other examples. In general, the importance of NQE is determined by two particular factors: (i) the mass of atoms and (ii) their interaction with each other and the environment. Both of these factors often lead to significant NQE in molecules and nanosystems. Nevertheless, even for those systems, the majority of simulations up to now have been done by treating the nuclei as classical particles. The main reason is the large computational cost of methods which treat nuclei quantum mechanically. Relatively cheap approaches based on the harmonic approximation have often led to dissatisfying results due to pronounced anharmonicity of NQE arising from the complexity of real interatomic interactions. Examples are the quantum fluctuations in molecular crystals such as paracetamol[17] or the negative thermal expansion of silicon,[21] to name a few. The state-of-the-art technique which treats both electrons and nuclei on a quantum-mechanical level of accuracy beyond the harmonic approximation is the Feynman−Kac imaginary time path integral method.[22,23] Within this method, after employing the Born−Oppenheimer approximation, the nuclear subsystem is mapped onto a so-called ring polymer (RP)[24]—an extended system containing P classical copies of the quantum system (beads, replicas) with a partition function which is equivalent to that of the original quantum system.

In practice, the applications of PI are limited by the large size of RP required for an accurate description of NQE. The number of degrees of freedom in RP simulations is P times larger than the corresponding total number of degrees of freedom in the classical system, and thus, the computational costs of calculating the energy and forces are increased by P times. The value of P that is required for convergence depends on temperature and the characteristic frequencies in the system but also on the factorization scheme employed for the decoupling of kinetic and potential energy operators in the expression for the partition function of the nuclei. In particular, high-order path integrals (HOPI)[25−28] schemes provide faster convergence and therefore are of great interest especially for low temperature simulations. High-order factorization schemes, such as Suzuki−Chin[25,26] (SC) or Takahashi−Imada[27] (TI), typically reduce the number of beads required for converged simulations by 2-fold, without introducing any approximations. By itself, this is already a considerable step forward from the so-called *primitive* factorization which is the essence of most of the recently developed alternative approaches.[10,29−31] The main drawback of high-order methods is that the derivative of the RP Hamiltonian involves terms that

Received: September 3, 2019
Published: January 8, 2020

https://dx.doi.org/10.1021/acs.jctc.9b00881
J. Chem. Theory Comput. 2020, 16, 1128−1135

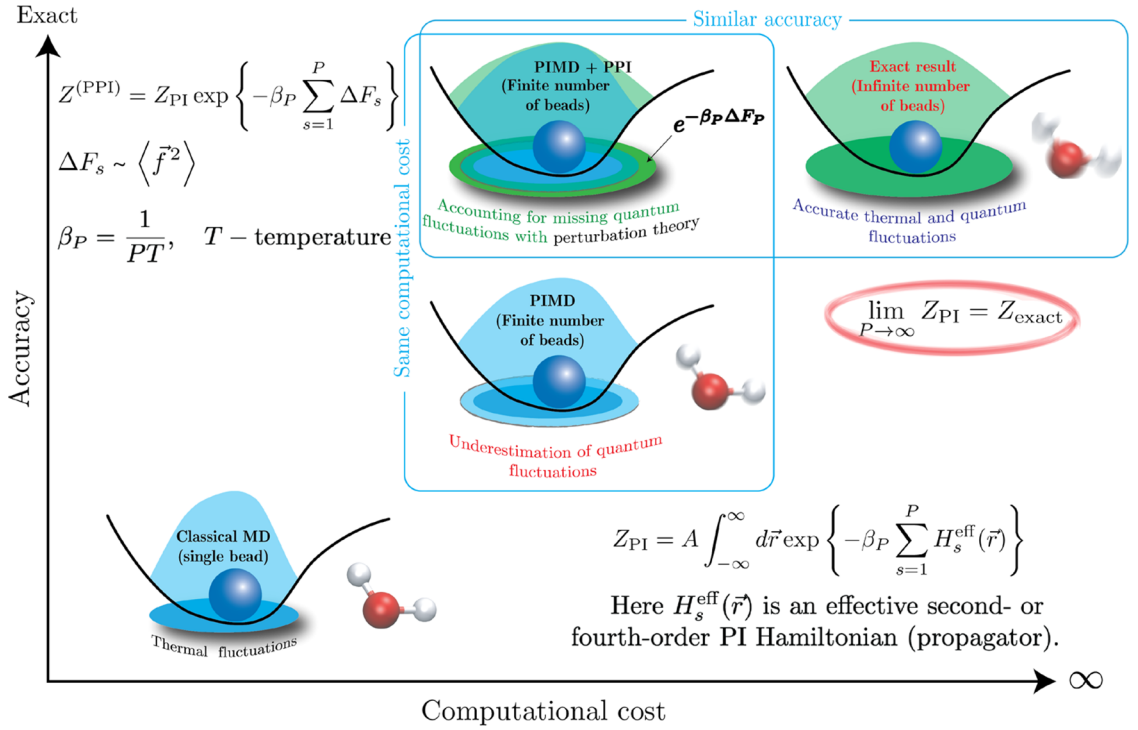


Figure 1. Employment of perturbation theory allows refining the ring polymer partition function, Z_{PI}, using the information encoded in atomic forces $\vec{f}$. As a result, converged thermodynamic observables can be computed using the perturbed partition function, $Z^{(\mathrm{PPI})}$, from PIMD trajectories obtained within conventional simulations with a finite number of beads.

are difficult to compute, which has prevented the extension of most of the accelerated path integral schemes to HOPI. Recently, however, it has been shown that these terms can be computed by finite differences with a small computational overhead.[32] Extending accelerated schemes to benefit from fourth-order sampling is, therefore, the most promising route to further reduce the cost of PI simulations, also given that going beyond the fourth-order propagators requires writing the partition function as a sum of nonequivalent terms, which are challenging to converge simultaneously within *ab initio* calculations.

In this work, we demonstrate an alternative way to improve the efficiency of HOPI by considering RP as a quasiclassical system. This ansatz has been previously applied to the conventional second-order PI approach, which has been shown to significantly reduce the required number of beads.[33,34] The limiting factor there was the inability of second-order PI to capture NQE accurately enough at a reasonable number of beads when lowering the temperature. Here, we employ fourth-order PI as a starting point for the perturbation theory,[35] which considerably extends the lower boundary of the accessible temperatures. Importantly, this approach does not add any additional computational cost beyond calculating the improved estimators.

2. PERTURBED HIGH-ORDER PATH INTEGRALS (HOPPI)

As a starting point in this work, we choose the fourth-order SC PI factorization, proposed in ref 32. Note that a similar method can be implemented with TI PI. The SC RP partition function can be written as

$$Z_{\mathrm{sc}} = A \int \prod_{i,s} d\vec{r}_i^{(s)} \exp\left\{-\beta V_{\mathrm{eff}}^{(\mathrm{sc})}\right\} \tag{1}$$

Here, A is the normalization constant, $\vec{r}_i^{(s)}$ is the position of the particle i within bead s, and $\beta = 1/k_{\mathrm{B}}T$ is the inverse temperature. Finally, $V_{\mathrm{eff}}^{(\mathrm{sc})}$ is the effective SC potential

$$V_{\mathrm{eff}}^{(\mathrm{sc})} = \sum_s \left[\frac{P}{2\beta^2\hbar^2} \sum_i m_i \left(\vec{r}_i^{(s+1)} - \vec{r}_i^{(s)}\right)^2 + \frac{1}{P} w_s V_s + \frac{w_s d_s \hbar^2 \beta^2}{P^3} \sum_i \frac{1}{m_i} \left(\vec{f}_i^{(s)}\right)^2 \right] \tag{2}$$

where P is the number of beads, $\vec{r}_{P+i} = \vec{r}_i$, m_i is the mass of the particle i, V_s is the potential energy of bead s, $\vec{f}_i^{(s)}$ is the physical force acting on the particle i within bead s, and w_i and d_i are decomposition coefficients.

By choosing the decomposition coefficients similar to those in ref 32

$$w_i = 2/3, \quad d_i = 0, \quad i \text{ is even}$$
$$w_i = 4/3 \; d_i = 1/12, \quad i \text{ is odd} \tag{3}$$

we only have to calculate for an MD step the spacial derivative from $(\vec{f}_i^{(s)})^2$ for even beads. Moreover, for "smooth" realistic potentials, this term can be recomputed every n steps (n = 4–8) without a significant loss in accuracy. Altogether, these make SC PIMD simulations an efficient and powerful tool for accounting for NQE, which is a vital step for studying molecular and condensed matter systems. Unfortunately, even within this optimal choice of parameters, the computational cost of PIMD simulations grows quickly when decreasing the temperature. Already, at 100 K, one has to use more than 30

https://dx.doi.org/10.1021/acs.jctc.9b00881
J. Chem. Theory Comput. 2020, 16, 1128–1135

beads to compute the energy of molecules such as water or CH_5^+. More complex properties, such as heat capacity, would require even larger computational effort.

2.1. Improving the Partition Function. To resolve this challenge, we apply perturbation theory to refine the SC RP partition function. The idea of the method is schematically shown in Figure 1 and can be utilized with any factorization scheme. The quantum fluctuations, that are typically underestimated within PIMD simulations with a finite number of beads, are accounted for by considering each bead as a quasiclassical system. In the limit $P \to \infty$, the effective temperature of RP goes to infinity, and the additional multiplier, $\exp(-\beta_P \Delta F)$, goes to unity. Hence, the exact solution of the PI in the limit of an infinite number of beads is not violated. By contrast, at a finite number of beads, the additional term leads to a more accurate expression for the partition function of the quantum system. This results in more precise values for the thermodynamic observables computed with this partition function. In practice, no one does PI simulations with an arbitrarily large number of beads to capture NQE. The limitation for the accuracy often comes from the quality of employed force fields or electronic structure methods. Figure 1, in this case, means that this limit can be reached much faster with the more accurate estimators obtained with the help of perturbation theory with a negligible increase in computational cost compared to that associated with generating PI trajectories.

Following the procedure proposed in ref 33 for the primitive factorization, one can write a "refined" SC partition function as

$$Z_{sc}^{(HOPPI)} = Z_{sc} Z_{corr} \tag{4}$$

where the correction term (called the PPI correction hereafter) has the form

$$Z_{corr} = \exp\left\{ -\frac{\hbar^2}{24}\left(\frac{\beta}{P}\right)^3 \sum_{s,i,\mu} \frac{\alpha_{i\mu}}{m_i} \left\langle f_{i\mu}^{(s)} \xi_{i\mu}^{(s)} \right\rangle_{sc} \right\} \tag{5}$$

Here, $\alpha_{i\mu}$ are damping parameters that can be determined *a priori*, $f_{i\mu}^{(s)} = -w_s \partial V_s / \partial r_{i\mu}^{(s)}$ and $\xi_{i\mu}^{(s)} = -2 w_s d_s \hbar^2 \beta^2 f_{i\mu}^{(s)} (\partial f_{i\mu(s)} / \partial r_{i\mu}^{(s)}) / (m_i P^3)$ are the μ components of the renormalized interparticle interaction force and the effective SC force arising from the last term in the right-hand side (RHS) of eq 2 acting on a particle i within bead s, respectively.

From eq 5, one can see that $\ln Z_{corr}$ is inversely proportional to P^5 (we have explicitly P^6 in the denominator, while the sum growth as P^1 since all its terms are finite). This guarantees the correct limit $Z_{sc}^{(HOPPI)} \to Z_{sc}$ when $P \to \infty$. We also emphasize here that the proposed method is an *a posteriori* procedure which does not modify the PI trajectories computed within the standard SC method. If these trajectories miss some important configurations, the HOPPI approach cannot fix this. Nevertheless, the values of PPI corrections can be estimated locally for different parts of the trajectory. This may reveal those configurations of the system where the underlying SC scheme with the given number of beads fails since the PPI corrections become large compared to the default estimators. A suitable instrument for such studies are the *effective temperatures* introduced in ref 34.

2.2. Avoiding Double Counting for NQE. The choice of the product of forces $f_{i\mu}^{(s)} \xi_{i\mu}^{(s)}$ and the introduction of $\alpha_{i\mu}$ in eq 5 are needed to avoid double counting for NQE already incorporated in eq 1. Hence, we omit the terms appearing as a result of perturbation theory in the expression for Z_{corr} which are proportional to $(\vec{f}_{i\mu}^{(s)})^2$ and $(\vec{\xi}_{i\mu}^{(s)})^2$. The former term is already included in Z_{sc}. The latter is proportional to $\hbar^6$ and is part of a higher-order correction, whose full functional form is unknown. Using only $f_{i\mu}^{(s)} \xi_{i\mu}^{(s)}$ in eq 5 still does not guarantee the avoidance of the double counting for NQE in eq 4. Both the effective SC force in the RP potential, eq 2, and the correction, eq 5, lead to an infinite number of terms in the decomposition of Z_{sc} in powers of $\hbar$. Hence, the damping parameters $\alpha_{i\mu}$, which are, in general, different for different degrees of freedom, must be introduced to avoid the double counting for NQE in $Z_{sc}^{(HOPPI)}$. Note that no damping parameters are needed when the method is applied to the second-order factorization scheme[33,34] since no fictitious potential energy terms appear in the effective Hamiltonian for this case.

The values of $\alpha_{i\mu}$ are determined by making $Z_{sc}^{(HOPPI)}$ exact for a quantum harmonic oscillator (QHO) at any given number of beads. That is, we obtain α for a specific degree of freedom j by solving the equation

$$Z_{exact}^{QHO}(\chi_j) - Z_{sc}^{(HOPPI)}(\chi_j, \alpha_j, P) = 0 \tag{6}$$

where Z_{exact}^{QHO} is the exact partition function of a one-dimensional QHO and

$$\chi_j = \exp\left\{ -\frac{\hbar^2}{24 m_j}\left(\frac{\beta}{P}\right)^3 \sum_{s \in even} \left\langle f_j^{(s)} \xi_j^{(s)} \right\rangle_{sc} \right\} \tag{7}$$

The new variables χ_j characterize the "quantumness" of different degrees of freedom in the system and are easily accessible within the conventional SC simulation procedure. The advantages of employing χ_j are the following: (i) χ_j are well defined for arbitrary, even anharmonic, degrees of freedom. (ii) They can be computed once for a given value of P and then rescaled to the case of simulations with a different number of beads L by using the scaling rule

$$\alpha_j(\chi_j, L) \simeq \alpha_j(\chi_j^{P/L}, P) \tag{8}$$

In practice, we use the analytical expression for the SC HOPPI partition function of a one-dimensional harmonic oscillator obtained in the two beads limit. In this case, we first extract the variable x from

$$\frac{2}{L}\ln(\chi_j) = -\frac{x^4(12 + x^2)}{72(432 + 36x^2 + x^4)} \tag{9}$$

with the value of χ_j computed by using eq 7 along the whole PIMD trajectory. Then, α_j (Figure 2) can be found as

$$\alpha_j = \frac{36(432 + 36x^2 + x^4)}{x^4(12 + x^2)} \times \left(\ln(1728) - 2\ln\left(x\sqrt{432 + 36x^2 + x^4}\, \mathrm{csch}\left(\frac{x}{2}\right)\right)\right) \tag{10}$$

Hence, the fitting of the damping parameters does not introduce any noticeable computational overhead to the proposed approach.

For the harmonic oscillator $x = \hbar\omega\beta$, the damping parameters obtained in this way are functions of temperature, and the corresponding derivatives must be accounted for when deriving estimators for different thermodynamic properties such as energy or heat capacity. Importantly, although we have

https://dx.doi.org/10.1021/acs.jctc.9b00881
J. Chem. Theory Comput. 2020, 16, 1128−1135

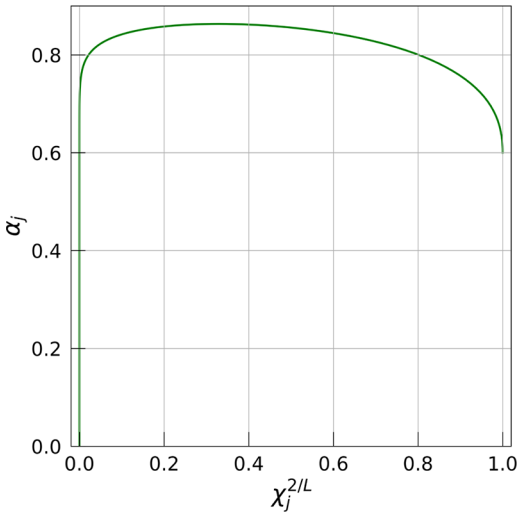


Figure 2. Damping parameter as a function of the "quantumness" of a given degree of freedom.

introduced the additional parameters α_j, whose number is equal to the number of degrees of freedom, the additional computational cost of this procedure is negligible since we have to compute the averages entering eq 7 anyway for PPI estimators.

2.3. Total Energy Estimator. To derive the HOPPI estimator for the total energy, E, we follow the standard procedure defining it as

$$E = -\frac{\partial \ln Z}{\partial \beta} \tag{11}$$

Other estimators, for instance, for kinetic or potential energy, can be derived in a similar way.

Substituting the expression for the improved partition function, eq 4, into eq 11 one obtains

$$E_{sc}^{(HOPPI)} = E_{sc} + \frac{\partial}{\partial \beta}\left\{\frac{\hbar^2}{24}\left(\frac{\beta}{P}\right)^3 \sum_{s,i,\mu} \frac{\alpha_{i\mu}}{m_i}\left\langle f_{i\mu}^{(s)} \xi_{i\mu}^{(s)}\right\rangle_{sc}\right\} \tag{12}$$

Here, the first term in the RHS is the total energy computed by employing any conventional SC total energy estimator, whereas the second term is the desired correction. Both $\alpha_{i\mu}$ and $\xi_{i\mu}^{(s)}$ in eq 12 are functions of temperature. The SC effective force, $\xi_{i\mu}^{(s)}$, is proportional to β^2, while the temperature dependence of $\alpha_{i\mu}$ is more challenging. To compute $\alpha_{i\mu}$ and $\partial\alpha_{i\mu}/\partial\beta$, we use the scaling rule of eq 8 and the analytical expression for the partition function of a QHO obtained within SC factorization scheme with two beads.

The final expression for the correction to the total energy can be written as

$$E_{corr} = 5\Delta F - D(\phi, \varepsilon) + \frac{\hbar^2}{24}\left(\frac{\beta}{P}\right)^3 \sum_{s,i,\mu} \frac{1}{m_i}\frac{\partial \alpha_{i\mu}}{\partial \beta}\left\langle f_{i\mu}^{(s)} \xi_{i\mu}^{(s)}\right\rangle_{sc} \tag{13}$$

Here, $\Delta F = -\ln(Z_{corr})/\beta = \langle\phi\rangle_{sc}$ is the free energy correction, ϵ is the primitive SC total energy estimator, and $D(A, B) = \langle AB\rangle_{sc} - \langle A\rangle_{sc}\langle B\rangle_{sc}$.

The efficiency of PI methods also depends upon the accuracy of the estimator employed for computing E_{sc}. While all estimators are bound to give the same (exact) result in the $P \to \infty$ limit, the accuracy at finite P varies depending on the estimator and the system. For the cases we consider here, we find that the most reliable option is the *primitive* thermodynamic estimator. In combination with the developed correction, it leads to the most accurate results at all values of P. The drawback is the poor statistical convergence of this estimator with increasing number of beads. The best alternative for the systems considered in this publication is the *virial* thermodynamic estimator (see ref 32). This estimator yields results similar to those obtained with the primitive one. By contrast, the so-called *operator* estimator[32] overestimates the energy at small numbers of beads compared to the primitive or virial estimators. We want to emphasize here that this finding is not general and may depend upon the system under study, temperature, and other conditions.

2.4. Estimators Depending upon Coordinates. In principle, one can obtain a general expression for an arbitrary observable, which is a function of coordinates similar to the derivation of eq 12 in ref 33. Unfortunately, the practical applications of such correction are minimal. The main problem is the emergence of third-order derivatives of the potential energy versus coordinates as well as complex combinations of first- and second-order derivatives. These derivatives are computationally costly, making PPI corrections for the properties such as the radius of gyration or radial distribution function within the SC method inefficient and impractical. Whenever someone is interested in these properties for a system at reasonably high temperatures (ambient conditions or close), more accurate results can be obtained by employing the combination of second-order PI and perturbation theory rather than the SC method. Below, we provide the expressions for the radius of gyration within the conventional second-order PI, corresponding PPI correction, and for the SC method. These expressions are used to compare the efficiency of different methods in the following section.

Estimators for the radius of gyration and centroid within the second-order PI

$$\rho_i^2 = \frac{1}{P}\sum_s (\Delta\vec{r}_i^{(s)})^2, \quad \Delta\vec{r}_i^{(s)} = \vec{r}_i^{(c)} - \vec{r}_i^{(s)} \tag{14}$$

$$\vec{r}_i^{(c)} = \frac{1}{P}\sum_s \vec{r}_i^{(s)} \tag{15}$$

An estimator for the PPI correction for the radius of gyration within the second-order PI

$$\Delta\rho_i^2 = \frac{\hbar^2\beta^2}{6P^3}\sum_s \frac{\vec{f}_i^{(s)}}{m_i}\Delta\vec{r}_i^{(s)} - \frac{\hbar^2\beta^3}{24P^3}D(\phi, \rho_i^2) \tag{16}$$

An estimator for the radius of gyration and centroid within the SC approach

$$\tilde{\rho}_i^2 = \frac{1}{P}\sum_s w_s(\Delta\vec{r}_i^{(s)})^2 - \frac{4\beta^2\hbar^2}{m_i P^3}\sum_s d_s w_s\left(\frac{w_s}{P} - 1\right)\vec{f}_i^{(s)}\Delta\vec{r}_i^{(s)} - \frac{4\beta^2\hbar^2}{m_i P^4}\sum_s d_s w_s \vec{\xi}_i^{(s)}\Delta\vec{r}_i^{(s)} \tag{17}$$

where the centroid entering $\Delta\vec{r}_i^{(s)}$ is

$$\tilde{\vec{r}}_i^{(c)} = \frac{1}{P}\sum_s w_s \vec{r}_i^{(s)} - \frac{2\hbar^2\beta^2}{m_i P^3}\sum_s d_s w_s \vec{f}_i^{(s)} \tag{18}$$

https://dx.doi.org/10.1021/acs.jctc.9b00881
J. Chem. Theory Comput. 2020, 16, 1128−1135

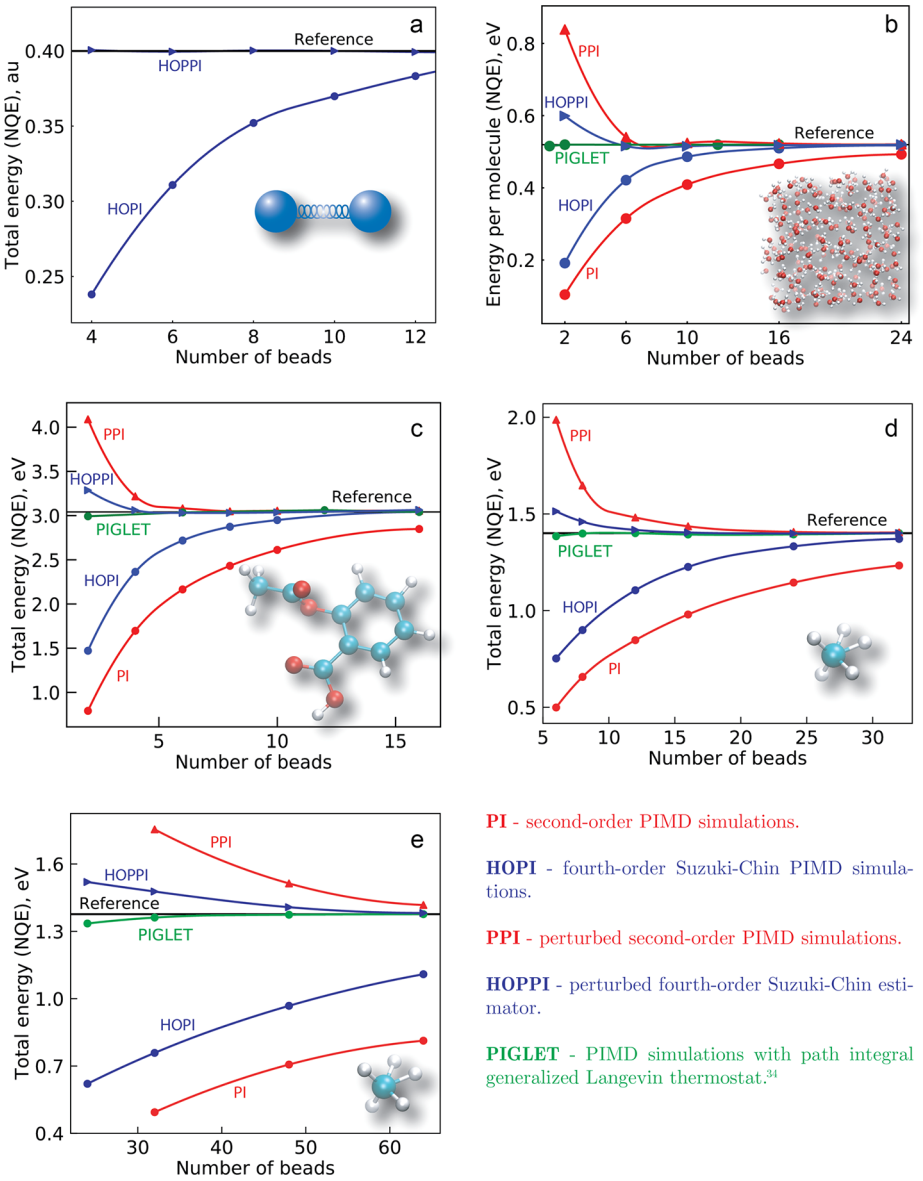


Figure 3. NQE contribution to the average total energy of (a) a three-dimensional isotropic quantum harmonic oscillator, (b) a qTIP4/PF model (simulations were done for 216 water molecules at 298 K in periodic boundary conditions), (c) an aspirin molecule described by sGDML potential at 300 K, and a CH_5^+ molecule described by a POSflex force field at (d) 100 K and (e) 20 K. PIGLET stands for the simulations with the path-integral generalized Langevin equation thermostat.[41] Symbols are the result of calculations, and solid lines serve as guidance for the eyes.

Note that the expressions for $\tilde{\vec{r}}_i^{(c)}$ and $\tilde{\rho}_i^2$ are obtained following the standard procedure

$$\langle \varepsilon \rangle = -\frac{1}{\beta}\frac{\partial \ln Z(V+\lambda\varepsilon)}{\partial \lambda}\bigg|_{\lambda=0} \tag{19}$$

These explain the force-dependent terms appearing in eqs 17 and 18.

3. RESULTS AND DISCUSSION

As practical examples, we consider below several benchmark systems in the temperature range starting from 300 K and down to 20 K. Figure 3 shows the total energy obtained with SC and second-order PIMD simulations with and without the PPI correction. All the simulations hereafter have been performed using the i-PI code.[36] A time step of 0.1 fs has been chosen for all systems. The CH_5^+ molecule is described by the POSflex force field,[37] as implemented in CP2K. For the aspirin molecule, we have used a recently developed sGDML force field trained on the CCSD/cc-pVDZ data with mean square error: 0.76 kcal/mol/Å for forces and 0.16 kcal/mol for the energy.[38,39] The liquid water is described by the qTIP4/PF model[40] with 216 water molecules using periodic boundary conditions. The notation PI stands for the second-order PIMD simulations with the Langevin equation thermostat, whereas HOPPI refers to the fourth-order Suzuki−Chin decomposition. Finally, the abbreviation PPI means the perturbed path-integrals correction developed in ref 33 for the second-order PI, and HOPPI stands for the improved SC RP partition function, eq 4.

By construction, the proposed approach reproduces the exact total energy of the isotropic three-dimensional QHO at an arbitrary number of beads within statistical errors (Figure 3(a)). For molecular systems, the deviation of the obtained results from the reference data at small number of beads

https://dx.doi.org/10.1021/acs.jctc.9b00881
J. Chem. Theory Comput. 2020, 16, 1128−1135

(Figure 3(b)−(e)) are explained by the quasiclassical nature of the employed correction, eq 5, and the anharmonicity of interatomic interactions. The reference hereafter is the result of fully converged (large P limit) second-order PIMD simulations.

We want to clarify that the comparisons of the efficiency of different estimators hereafter are made regarding the required number of beads rather than actual computational costs. A cost-wise comparison is beyond the scope of this article as the expense of performing HOPI and second-order PI simulations depends on the efficiency of the sampling methods which is orthogonal to the efficiency of estimators used to compute expectation values. Indeed, methods such as multiple time stepping, ring polymer contraction, and their combination can be used to reduce the cost of both HOPI and second-order PI simulations. In this work, however, we use the same simulation settings for every considered method and discuss only the numbers of beads. The computational overhead of the HOPPI scheme by itself is negligible, and the statistical convergences of the estimators, mainly defined by the underlying sampling method,[33] are also similar.

Figure 3(b) shows the NQE contribution to the total energy (per molecule) of liquid water obtained for different numbers of beads. The simulations were done at 298 K and at the density corresponding to atmospheric pressure. One can see that the PPI method, together with the standard PILE and SC PIMD simulations, yields total energies with less than a 2.5% error starting from 10 and 6 beads, respectively. The same level of accuracy with the conventional estimators would require approximately 20 beads within the SC method, while second-order PIMD simulations even with 24 beads still underestimate the NQE contribution to the total energy by 14%.

Another system of practical interest, which exhibits pronounced nonclassical behavior of nuclei even at 300 K, is the aspirin molecule (Figure 3(c)). Here, we set the convergence criterion for the NQE part of the total energy as 1% error with respect to the reference data. To fulfill this criterion within the conventional SC PIMD simulations, one would have to use at least 16 beads. The second-order approach underestimates the total energy by more than 2% even with 28 beads. Employing the PPI method proposed in ref 33, the required accuracy is reached by using the PILE trajectories with 8 beads. Application of perturbation theory to the more accurate SC partition function improves the efficiency of PIMD simulations even further. By using only 4 beads we can reproduce the NQE contribution to the total energy with ∼0.6% error.

As the next example, we consider the CH_5^+ molecule described by the POSflex force field[37] (Figure 3(d) and (e)). To enhance the role of NQE, the simulations have been done at lower temperatures of 100 and 20 K. At 100 K, similar to the previous two cases, the second-order method fails to reproduce the NQE accurately at a reasonably small number of beads leading to a 12% error even when 32 beads are used for the simulations. Using of the SC decomposition decreases the error with the same number of beads to 2.5%. The combination of PILE and PPI achieves the same accuracy using twice fewer beads compared to the SC approach, which considerably increases the efficiency. Finally, the HOPPI total energy estimator makes it possible to perform simulations with only 10 beads and stay within the same 2.5% error bar.

In addition, Figure 3(d) shows the results obtained by employing the so-called path-integral generalized Langevin equation thermostat (PIGLET).[41] GLE methods share some similarities with high-order perturbed path integrals, in that they are designed to yield exact results in the harmonic limit by means of a history-dependent Langevin dynamics that drives the system to a stationary state that involves frequency-dependent fluctuations. In PPI, information on the mean curvature of the potential is encoded in the average value of the squared forces, and hence, the convergence can be hampered for fluxional molecules such as CH_5^+, in which the stiffness of the potential experienced by protons changes over time. In GLE methods, information on the curvature is determined by the history-dependent friction term, and the correction to the phase-space distribution can adapt to slowly varying changes in the behavior of particles. The downside to the adaptive nature of PIGLET and related methods is that the probability distribution that is sampled by the dynamics cannot be written as a simple function of the coordinates of the ring polymer, which means that nonstandard estimators such as heat capacity[42] or those for isotope−fractionation ratios[43,44] do not necessarily benefit from enhanced convergence. Furthermore, advanced sampling techniques such as replica exchange or metadynamics cannot be combined with the colored-noise dynamics. Conversely, PPI involve only conventional thermostatting and sample a simply defined ring polymer ensemble, which means that they can be combined with most accelerated or biased sampling schemes, and that it is possible, in principle, to evaluate any path integral estimator. In either of these scenarios, using the standard second- or high-order PIMD schemes with the PPI correction will be beneficial.

Figure 4 shows the radius of gyration computed for a hydrogen atom within the conventional second-order, second-order plus PPI correction, and SC approaches. One can see that the combination of second-order PI and perturbation theory provides the most accurate results considerably outperforming the SC method. As a reference, we use the radius of gyration computed within the second-order PIMD simulations with 84 beads. For completeness, we present here

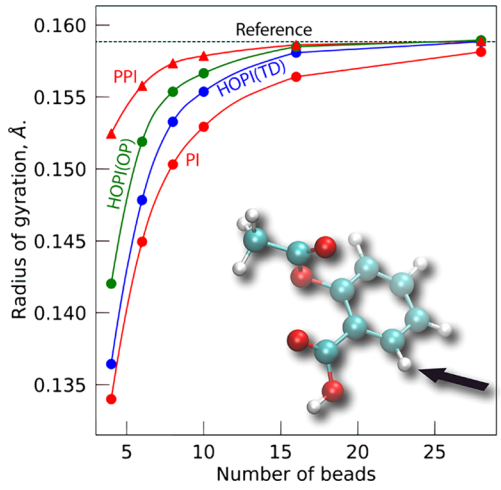


Figure 4. Radius of gyration for the hydrogen atom (indicated by an arrow) in the molecule of aspirin computed at 300 K using four different methods as a function of the number of beads. HOPI(TD) stands for the Suzuki−Chin thermodynamic estimator eq 17, while HOPI(OP) is the so-called operator estimator (see the text for details). Symbols are the result of calculations, and solid lines serve as guidance for the eyes.

https://dx.doi.org/10.1021/acs.jctc.9b00881
J. Chem. Theory Comput. 2020, 16, 1128−1135

the results obtained with the SC thermodynamic estimator (eq 17) and so-called SC operator estimator obtained by applying the conventional definition for the radius of gyration to all beads and centroid to only even beads (see ref 32). One can see a noticeable difference between the results obtained with TD and OP estimators at small numbers of beads.

An important practical advantage of the PPI approach is the possibility to estimate the convergence of PIMD simulations with respect to the number of beads without a necessity to run time-consuming benchmark simulations. If the PPI correction amounts to only a reasonably small part of the conventional SC estimator, the result of the perturbed SC estimator can be considered as converged for a given observable. Otherwise, the simulations have to be repeated with an increased number of beads. Importantly, such an analysis can be done for any property of interest from a short PIMD trajectory. Finally, we emphasize here that while one can mistakenly conclude from analyzing Figure 3 that PPI gives an upper bound estimation for the total energy, this is not true. The proposed scheme is the perturbation approach, and its results usually oscillate around the exact results when changing the number of beads.

4. CONCLUSIONS

Summarizing, we have shown that the combination of perturbation theory and PIMD simulations of second and fourth orders considerably increase the accuracy of the total energy estimators. The proposed approach is free of any additional computational cost, is equally applicable for any thermodynamic observable, and can be utilized both *a posteriori* and on-the-fly. Employing the perturbed PI with second-order PIMD simulations makes it possible to use twice fewer beads compared to the fourth-order Suzuki-Chin PI, while achieving the same accuracy. When applied to the Suzuki−Chin factorization, perturbed path integrals enable efficient simulations of molecular systems with considerable NQE at low temperatures. The main limitation of the method comes from the necessity to know high-order derivatives for computing position dependent properties. The appearance of reliable machine learning force fields providing accurate and computationally cheap high-order derivatives can remove this limitation, making HOPPI a method of choice for low-temperatures simulations. With all these developments, the bottleneck of PIMD simulations for realistic molecular or condensed systems becomes the sampling procedure rather than the accuracy of the estimators. Hence, the development of efficient sampling techniques and accurate force fields should become central problems of future studies related to imaginary time path integral simulations.

AUTHOR INFORMATION

Corresponding Author

Alexandre Tkatchenko − *Physics and Materials Science Research Unit, University of Luxembourg, Luxembourg City L-1511, Luxembourg;* orcid.org/0000-0002-1012-4854; Email: alexandre.tkatchenko@uni.lu

Other Authors

Igor Poltavsky − *Physics and Materials Science Research Unit, University of Luxembourg, Luxembourg City L-1511, Luxembourg;* orcid.org/0000-0002-3188-7017

Venkat Kapil − *Laboratory of Computational Science and Modelling, Institute of Materials, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland;* orcid.org/0000-0003-0324-2198

Michele Ceriotti − *Laboratory of Computational Science and Modelling, Institute of Materials, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland;* orcid.org/0000-0003-2571-2832

Kwang S. Kim − *Department of Chemistry, School of Natural Science, Ulsan National Institute of Science and Technology, Ulsan 44919, Korea;* orcid.org/0000-0002-6929-5359

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jctc.9b00881

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

I.P. and A.T. acknowledge financial support from the Luxembourg National Research within the FNR-CORE program (FNR-11360857). M.C. and V.K. acknowledge financial support by the Swiss National Science Foundation (Project ID 200021-159896).

REFERENCES

(1) Liang, C.; Tocci, G.; Wilkins, D. M.; Grisafi, A.; Roke, S.; Ceriotti, M. Solvent fluctuations and nuclear quantum effects modulate the molecular hyperpolarizability of water. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2017**, *96*, No. 041407.

(2) Ceriotti, M.; Fang, W.; Kusalik, P. G.; McKenzie, R. H.; Michaelides, A.; Morales, M. A.; Markland, T. E. Nuclear Quantum Effects in Water and Aqueous Systems: Experiment, Theory, and Current Challenges. *Chem. Rev.* **2016**, *116*, 7529−7550.

(3) Morrone, J. A.; Car, R. Nuclear Quantum Effects in Water. *Phys. Rev. Lett.* **2008**, *101*, No. 017801.

(4) Marsalek, O.; Chen, P.-Y.; Dupuis, R.; Benoit, M.; Méheut, M.; Bačić, Z.; Tuckerman, M. E. Efficient Calculation of Free Energy Differences Associated with Isotopic Substitution Using Path-Integral Molecular Dynamics. *J. Chem. Theory Comput.* **2014**, *10*, 1440−1453.

(5) Marx, D.; Tuckerman, M. E.; Hutter, J.; Parrinello, M. The nature of the hydrated excess proton in water. *Nature* **1999**, *397*, 601−604.

(6) Marx, D.; Tuckerman, M. E.; Parrinello, M. Solvated excess protons in water: quantum effects on the hydration structure. *J. Phys.: Condens. Matter* **2000**, *12*, A153.

(7) Mei, H. S.; Tuckerman, M. E.; Sagnella, D. E.; Klein, M. L. Quantum Nuclear ab Initio Molecular Dynamics Study of Water Wires. *J. Phys. Chem. B* **1998**, *102*, 10446−10458.

(8) Vega, C.; Conde, M. M.; McBride, C.; Abascal, J. L.; Noya, E. G.; Ramirez, R.; Sesé, L. M. Heat capacity of water: A signature of nuclear quantum effects. *J. Chem. Phys.* **2010**, *132*, No. 046101.

(9) Poltavsky, I.; Zheng, L.; Mortazavi, M.; Tkatchenko, A. Quantum tunneling of thermal protons through pristine graphene. *J. Chem. Phys.* **2018**, *148*, 204707.

(10) Ceriotti, M.; Bussi, G.; Parrinello, M. Nuclear Quantum Effects in Solids Using a Colored-Noise Thermostat. *Phys. Rev. Lett.* **2009**, *103*, No. 030603.

(11) Li, X.-Z.; Walker, B.; Michaelides, A. Quantum nature of the hydrogen bond. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 6369−6373.

(12) Chen, B.; Ivanov, I.; Klein, M. L.; Parrinello, M. Hydrogen Bonding in Water. *Phys. Rev. Lett.* **2003**, *91*, 215503.

(13) Ceriotti, M.; Cuny, J.; Parrinello, M.; Manolopoulos, D. E. Nuclear quantum effects and hydrogen bond fluctuations in water. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 15591−15596.

(14) Mella, M.; Kuo, J.-L.; Clary, D. C.; Klein, M. L. Nuclear quantum effects on the structure and energetics of (H2O)6H+. *Phys. Chem. Chem. Phys.* **2005**, *7*, 2324−2332.

(15) Kapil, V.; Cuzzocrea, A.; Ceriotti, M. Anisotropy of the Proton Momentum Distribution in Water. *J. Phys. Chem. B* **2018**, *122*, 6048−6054.
(16) Rossi, M.; Fang, W.; Michaelides, A. Stability of Complex Biomolecular Structures: van der Waals, Hydrogen Bond Cooperativity, and Nuclear Quantum Effects. *J. Phys. Chem. Lett.* **2015**, *6*, 4233−4238.
(17) Rossi, M.; Gasparotto, P.; Ceriotti, M. Anharmonic and Quantum Fluctuations in Molecular Crystals: A First-Principles Study of the Stability of Paracetamol. *Phys. Rev. Lett.* **2016**, *117*, 115702.
(18) Kapil, V.; Engel, E.; Rossi, M.; Ceriotti, M. Assessment of Approximate Methods for Anharmonic Free Energies. *J. Chem. Theory Comput.* **2019**, *15*, 5845−5857.
(19) Morales, M. A.; McMahon, J. M.; Pierleoni, C.; Ceperley, D. M. Nuclear Quantum Effects and Nonlocal Exchange-Correlation Functionals Applied to Liquid Hydrogen at High Pressure. *Phys. Rev. Lett.* **2013**, *110*, No. 065702.
(20) Herrero, C. P.; Ramírez, R. Quantum effects in graphene monolayers: Path-integral simulations. *J. Chem. Phys.* **2016**, *145*, 224701.
(21) Kim, D. S.; Hellman, O.; Herriman, J.; Smith, H. L.; Lin, J. Y. Y.; Shulumba, N.; Niedziela, J. L.; Li, C. W.; Abernathy, D. L.; Fultz, B. Nuclear quantum effect with pure anharmonicity and the anomalous thermal expansion of silicon. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115*, 1992−1997.
(22) Berne, B. J.; Thirumalai, D. On the Simulation of Quantum Systems: Path Integral Methods. *Annu. Rev. Phys. Chem.* **1986**, *37*, 401−424.
(23) Schmidt, K. E.; Ceperley, D. M. In *The Monte Carlo Method in Condensed Matter Physics*; Binder, K., Ed.; Topics in Applied Physics; Springer: Berlin Heidelberg, 1995; Vol. *71*; pp 205−248.
(24) Tuckerman, M. E. In *Quantum Simulations of Complex Many-Body Systems: From Theory to Algorithms*; Grotendorst, J., Marx, D., Alejandro, M., Eds.; NIC Series; John von Neumann Institute for Computing: Jülich, 2002; Vol. *10*; pp 269−298.
(25) Suzuki, M. General Nonsymmetric Higher-Order Decomposition of Exponential Operators and Symplectic Integrators. *J. Phys. Soc. Jpn.* **1992**, *61*, 3015−3019.
(26) Chin, S. A. Symplectic integrators from composite operator factorizations. *Phys. Lett. A* **1997**, *226*, 344−348.
(27) Takahashi, M.; Imada, M. Monte Carlo Calculation of Quantum Systems. II. Higher Order Correction. *J. Phys. Soc. Jpn.* **1984**, *53*, 3765−3769.
(28) Jang, S.; Jang, S.; Voth, G. A. Applications of higher order composite factorization schemes in imaginary time path integral simulations. *J. Chem. Phys.* **2001**, *115*, 7832−7842.
(29) Markland, T. E.; Manolopoulos, D. E. A refined ring polymer contraction scheme for systems with electrostatic interactions. *Chem. Phys. Lett.* **2008**, *464*, 256−261.
(30) Marsalek, O.; Markland, T. E. Ab initio molecular dynamics with nuclear quantum effects at classical cost: Ring polymer contraction for density functional theory. *J. Chem. Phys.* **2016**, *144*, No. 054112.
(31) Kapil, V.; VandeVondele, J.; Ceriotti, M. Accurate molecular dynamics and nuclear quantum effects at low cost by multiple steps in real and imaginary time: Using density functional theory to accelerate wavefunction methods. *J. Chem. Phys.* **2016**, *144*, No. 054111.
(32) Kapil, V.; Behler, J.; Ceriotti, M. High order path integrals made easy. *J. Chem. Phys.* **2016**, *145*, 234103.
(33) Poltavsky, I.; Tkatchenko, A. Modeling quantum nuclei with perturbed path integral molecular dynamics. *Chem. Sci.* **2016**, *7*, 1368−1372.
(34) Poltavsky, I.; DiStasio, R. A.; Tkatchenko, A. Perturbed path integrals in imaginary time: Efficiently modeling nuclear quantum effects in molecules and materials. *J. Chem. Phys.* **2018**, *148*, 102325.
(35) Landau, L. D.; Lifshitz, E. M. *Statistical Physics*; Butterworth−Heinemann: Oxford, 1980.
(36) Kapil, V.; Rossi, M.; Marsalek, O.; Petraglia, R.; Litman, Y.; Spura, T.; Cheng, B.; Cuzzocrea, A.; Meißner, R. H.; Wilkins, D. M.; Helfrecht, B. A.; Juda, P.ła.; Bienvenue, S. P.; Fang, W.; Kessler, J.; Poltavsky, I.; Vandenbrande, S.; Wieme, J.; Corminboeuf, C.; Kuhne, T. D.; Manolopoulos, D. E.; Markland, T. E.; Richardson, J. O.; Tkatchenko, A.; Tribello, G. A.; Van Speybroeck, V.; Ceriotti, M. i-PI 2.0: A universal force engine for advanced molecular simulations. *Comput. Phys. Commun.* **2019**, *236*, 214.
(37) Uhl, F.; Walewski, L.; Forbert, H.; Marx, D. Adding flexibility to the "particles-on-asphere" model for large-amplitude motion: POSflex force field for protonated methane. *J. Chem. Phys.* **2014**, *141*, 104110.
(38) Chmiela, S.; Sauceda, H. E.; Müller, K.-R.; Tkatchenko, A. Towards exact molecular dynamics simulations with machine-learned force fields. *Nat. Commun.* **2018**, *9*, 3887.
(39) Chmiela, S.; Tkatchenko, A.; Sauceda, H. E.; Poltavsky, I.; Schütt, K. T.; Müller, K.-R. Machine learning of accurate energy-conserving molecular force fields. *Sci. Adv.* **2017**, *3*, No. e1603015.
(40) Habershon, S.; Markland, T. E.; Manolopoulos, D. E. Competing quantum effects in the dynamics of a flexible water model. *J. Chem. Phys.* **2009**, *131*, No. 024501.
(41) Uhl, F.; Marx, D.; Ceriotti, M. Accelerated path integral methods for atomistic simulations at ultra-low temperatures. *J. Chem. Phys.* **2016**, *145*, No. 054101.
(42) Kapil, V.; Wieme, J.; Vandenbrande, S.; Lamaire, A.; Van Speybroeck, V.; Ceriotti, M. Modeling the structural and thermal properties of loaded metal-organic frameworks. An interplay of quantum and anharmonic fluctuations. *J. Chem. Theory Comput.* **2019**, *15* (5), 3237−3249.
(43) Ceriotti, M.; Markland, T. E. Efficient methods and practical guidelines for simulating isotope effects. *J. Chem. Phys.* **2013**, *138*, No. 014112.
(44) Cheng, B.; Ceriotti, M. Direct path integral estimators for isotope fractionation ratios. *J. Chem. Phys.* **2014**, *141*, 244112.

https://dx.doi.org/10.1021/acs.jctc.9b00881
J. Chem. Theory Comput. 2020, 16, 1128−1135