Perturbed path integrals in imaginary time: Efficiently modeling nuclear quantum effects in molecules and materials

Igor Poltavsky, Robert A. DiStasio, and Alexandre Tkatchenko

Citation: The Journal of Chemical Physics **148**, 102325 (2018);
View online: https://doi.org/10.1063/1.5006596
View Table of Contents: http://aip.scitation.org/toc/jcp/148/10
Published by the American Institute of Physics

Perturbed path integrals in imaginary time: Efficiently modeling nuclear quantum effects in molecules and materials

Igor Poltavsky,[1] Robert A. DiStasio, Jr.,[2] and Alexandre Tkatchenko[1]
[1]*Physics and Materials Science Research Unit, University of Luxembourg, Luxembourg L-1511, Luxembourg*
[2]*Department of Chemistry and Chemical Biology, Cornell University, Ithaca, New York 14853, USA*

(Received 26 September 2017; accepted 13 November 2017; published online 6 December 2017)

Nuclear quantum effects (NQE), which include both zero-point motion and tunneling, exhibit quite an impressive range of influence over the equilibrium and dynamical properties of molecules and materials. In this work, we extend our recently proposed *perturbed path-integral* (PPI) approach for modeling NQE in molecular systems [I. Poltavsky and A. Tkatchenko, Chem. Sci. **7**, 1368 (2016)], which successfully combines the advantages of thermodynamic perturbation theory with path-integral molecular dynamics (PIMD), in a number of important directions. First, we demonstrate the accuracy, performance, and general applicability of the PPI approach to both molecules and extended (condensed-phase) materials. Second, we derive a series of estimators within the PPI approach to enable calculations of structural properties such as radial distribution functions (RDFs) that exhibit rapid convergence with respect to the number of beads in the PIMD simulation. Finally, we introduce an *effective nuclear temperature* formalism within the framework of the PPI approach and demonstrate that such effective temperatures can be an extremely useful tool in quantitatively estimating the "quantumness" associated with different degrees of freedom in the system as well as providing a reliable quantitative assessment of the convergence of PIMD simulations. Since the PPI approach only requires the use of standard second-order imaginary-time PIMD simulations, these developments enable one to include a treatment of NQE in equilibrium thermodynamic properties (such as energies, heat capacities, and RDFs) with the accuracy of higher-order methods but at a fraction of the computational cost, thereby enabling first-principles modeling that simultaneously accounts for the quantum mechanical nature of both electrons and nuclei in large-scale molecules and materials. *Published by AIP Publishing.* https://doi.org/10.1063/1.5006596

I. INTRODUCTION

The quantum mechanical nature of small particles such as (light) nuclei results in substantive delocalization of their classical positions, which in turn necessarily influences the equilibrium and dynamical properties of molecules and materials. In fact, such nuclear quantum effects (NQE), which include both zero-point motion and tunneling, exhibit quite an impressive range of influence that encompasses simple molecular properties such as bond dissociation energies[1] as well as complex condensed-phase properties such as the relative thermodynamic stabilities of competing molecular crystal polymorphs.[2,3] In liquid water, where the underlying tetrahedral network is primarily governed by the positions of the hydrogen atoms, it has been shown that accurate and reliable descriptions of the hydrogen bonding,[4–6] proton momentum distribution,[7] hydrated excess proton,[8,9] transport mechanism of aqueous hydroxide ions,[10] and numerous other phenomena[11] require a theoretical treatment of NQE beyond the harmonic approximation. In this regard, such non-classical nuclear behavior has also been observed in a number of systems throughout biology, chemistry, physics, and materials science, including DNA base pairs,[12] aromatic molecules,[13] graphene,[14,15] and equilibrium fractionation of stable aqueous Li isotopes,[16] to name a few.

A proper account for the NQE in all of the aforementioned systems also requires an accurate quantum mechanical description of the accompanying electronic degrees of freedom. In this regard, the *ab initio* quantum chemical and density functional methods provide a systematic route toward attaining this goal and have therefore become an essential tool in many fields of research.[17–19] In fact, state-of-the-art density-functional theory (DFT) approximations[20–22] that simultaneously reduce self-interaction error while treating the many-body nature of non-covalent interactions[23–25] are now able to furnish predictions for the binding energies of small molecules[26] and supramolecular systems[27] with so-called "chemical accuracy" (i.e., with errors less than 1 kcal mol^{-1}), thereby extending the regime of applicability of such high-accuracy approaches to systems that contain 1000s of atoms. Other methods that are quite useful for treating significantly larger systems include the classical force fields, in which the quantum mechanical description of the electronic degrees of freedom is mapped onto an analytical (and therefore computationally efficient) interatomic/interfragment potential. When compared to the largest system sizes accessible with the most efficient DFT implementations, accurately designed force fields such as the q-TIP4P/F water model[28] allow us to consider systems containing orders of magnitude more atoms. Furthermore, machine-learned potentials[29] can become

indispensable tools for cases where suitably reliable force fields are not available, for instance, the recently designed gradient-domain machine learning approach,[30] which uses only 1000 conformational geometries for training, is able to reproduce global potential energy surfaces of small- and intermediate-sized molecules with an accuracy of 0.3 kcal mol^{-1} for energies and 1 kcal mol^{-1} Å^{-1} for atomic forces. Successes such as these are making it more evident that the combination of machine-learned potentials with highly accurate quantum chemistry[31] or quantum Monte Carlo[32] approaches will pave the way toward a new era of electronic structure calculations.

Despite the fact that there are many systems of practical interest (in addition to those mentioned earlier), where NQE play an important role,[33–52] routine applications of these successful approaches for (explicitly or implicitly) describing the quantum mechanical nature of the electrons are almost always accompanied by the following series of approximations regarding the nuclei: (i) the electronic and nuclear degrees of freedom are separated by invoking the Born-Oppenheimer (BO) approximation and (ii) the nuclei are considered as purely classical point particles. By employing such approximations, one must of course consider the fact that the temperature, the mass of the nuclei, and the interactions between the nuclei and their environment are all non-trivial factors that determine the relative importance of NQE in a given system or process of interest. For instance, strong interatomic interactions in molecules and crystals may lead to non-classical behavior in the light nuclei at room (and even higher) temperature.[53,54] Extreme thermodynamic conditions, such as found with liquid hydrogen[55] or dense lithium[56] under high pressures, constitute additional factors that can substantially increase the relative importance and influence of NQE.

An accurate description of a number of physical properties in these cases is still possible within the BO approximation, assuming that the electrons and nuclei are both treated on an equivalent quantum mechanical footing. In this regard, the standard theoretical technique for taking such NQE into account is the Feynman-Kac imaginary-time path-integral molecular dynamics (PIMD) approach,[57,58] which can in principle be utilized in conjunction with any of the aforementioned methods for quantum mechanically describing the corresponding electronic degrees of freedom. In short, this approach maps a quantum system onto P classical copies (or "beads") that interact with their nearest neighbors via a harmonic potential and constitute a so-called ring polymer.[59] Equilibrium properties are then computed by sampling the phase space of this ring polymer by following the corresponding Newtonian trajectories outfitted with a canonical ensemble thermostat to ensure ergodicity. Hence, the main computational challenges associated with these simulations are the considerable costs that arises when obtaining the thermodynamic averages for the ring-polymer system, which is now comprised of P times the number of degrees of freedom found in the original system.

A number of methods have been reported in the literature that attempt to reduce the computational cost associated with PIMD simulations by decreasing the number of beads required for convergence of a given equilibrium property of interest.[33,41,53,60–68] Although all of these approaches are exact in the limit where $P \to \infty$, the results of these methods can differ quite substantially for finite values of P.[69] In order to choose the most appropriate method, it is important that one correctly and quantitatively identifies the degrees of freedom in the system that are most affected by NQE. In this work, we use the perturbed path-integral (PPI) approach[53] to address this challenge by introducing an *effective nuclear temperature* formalism[70] and demonstrate that such effective temperatures can be an extremely useful tool in quantitatively estimating the "quantumness" associated with different degrees of freedom in the system under study. Using the results of MD simulations with classical nuclei only, this approach furnishes a series of effective temperatures that can also be employed to estimate the convergence of PIMD simulations with respect to P without having to perform more expensive PIMD simulations with larger Trotter numbers to approximate residual errors. We also show here that the PPI approach allows one to estimate any equilibrium observable by developing the appropriate correction to the conventional second-order radial distribution function (RDF) estimator in addition to the previously proposed corrections to the energy and heat capacity.[53] This improved RDF estimator only requires knowledge of the atomic positions and forces and converges quite rapidly with P (i.e., by approximately a factor of five) when compared to the conventional RDF estimator, which is a significant advance in the modeling of NQE in large condensed-phase systems.

The structure of the remainder of this manuscript is as follows. In Sec. II, we briefly review the PPI approach (see Ref. 53 for more details). We then present the effective nuclear temperature formalism within the PPI framework in Sec. III, a development which enables (i) identification and quantitative estimation of the "quantumness" associated with different nuclear degrees of freedom in the system under study, (ii) estimation of the residual errors in PIMD simulations, and (iii) verification of the convergence of PIMD simulation results with respect to P. In Sec. IV, we then derive the PPI correction to the RDF estimator and benchmark it using the results of PIMD simulations for bulk water at ambient conditions. The manuscript is then ended with some conclusions regarding the future outlook of the field in Sec. V.

II. PERTURBED PATH INTEGRALS IN IMAGINARY TIME: HIGH-ORDER EFFICIENCY WITH THE SECOND-ORDER PROPAGATOR

Imaginary-time PIMD simulations are based on the representation of the partition function for a quantum system as a multi-dimensional integral over P interacting classical copies (or beads) of the system,

$$Z_{\rm PI} = A \int \prod_{is} d\vec{q}_i^{\,(s)} e^{-\beta U_{\rm eff}}, \tag{1}$$

where $\prod_{i,s} \equiv \prod_{i=1}^{N} \prod_{s=1}^{P}$ runs over the N particles comprising each of the P beads, A is a normalization constant, $\beta = 1/k_B T$ is the inverse temperature, $\vec{q}_i^{\,(s)}$ is the coordinate vector for particle i in bead s, and $U_{\rm eff}$ is the effective potential. Within the commonly used second-order approach, the form of $U_{\rm eff}$ is

derived by employing the Trotter theorem for the exponential of the Hamiltonian ($\hat{H}$),[59] i.e.,

$$e^{-\beta\hat{H}} = e^{-\beta(\hat{K}+\hat{U})} = \lim_{P\to\infty}\left[e^{-\frac{\beta\hat{U}}{2P}}e^{-\frac{\beta\hat{K}}{P}}e^{-\frac{\beta\hat{U}}{2P}}\right]^P, \quad (2)$$

in which $\widehat{K}$ and $\widehat{U}$ are the kinetic and potential energy operators, respectively. As a result, one obtains that

$$U_{\text{eff}} = \sum_i \frac{m_i\omega_P^2}{2}\sum_s\left(\vec{q}_i^{\,(s+1)} - \vec{q}_i^{\,(s)}\right)^2 + \frac{1}{P}\sum_s U_s, \quad (3)$$

where m_i is the mass of particle i, $\omega_P^2 = P/\beta^2$ is the chain frequency, and U_s is the intrabead potential energy corresponding to the particles in bead s, respectively.

When compared to higher-order expansions of $e^{-\beta\hat{H}}$,[60,62] the main advantage of the second-order factorization scheme in Eq. (2) is that U_{eff} simply depends on U_s (and not higher-order derivatives of U_s), allowing one to employ the PIMD approach to compute the multi-dimensional integral required for evaluation of the quantum partition function in Eq. (1). The main disadvantage of this expansion is the relatively slow convergence of associated thermodynamic properties with respect to P [typically $\mathcal{O}(P^{-2})$], which therefore requires increased computational effort for a sufficiently accurate and well-converged treatment of NQE.

A particularly efficient way to improve the accuracy of conventional second-order PIMD simulations (instead of simply increasing P) is to combine quantum mechanical perturbation theory with the Feynman-Kac imaginary-time path-integral technique. As explained in detail in Ref. 53, the resulting perturbed path-integral (PPI) method is based on the refinement of the second-order Z_{PI} by considering a ring-polymer consisting of quasi-classical copies of the initial quantum system. Mathematically speaking, this ansatz results in modification of the second-order ring-polymer partition function to yield $Z_{\text{PPI}} = Z_{\text{PI}}Z_{\text{q}}$ via

$$Z_{\text{q}} = \exp\left[-\beta\Delta F\right] = \exp\left[-\frac{\hbar^2\beta^3}{24P^3}\sum_i\frac{1}{m_i}\sum_s\left\langle\left(\vec{f}_i^{\,(s)}\right)^2\right\rangle\right]. \quad (4)$$

In this expression, ΔF is the PPI free-energy correction, $\vec{f}_i^{\,(s)}$ is the force acting on particle i in bead s, and $\langle\cdots\rangle$ indicates a thermodynamic average.

The advantages of the PPI approach over other existing schemes include:

1. no requirement for empirical parameters or elaborate fitting procedures;
2. no issues arising from statistical convergence problems upon increasing the system size;
3. applicability to the calculation of any thermodynamic observable;
4. no additional computational cost is needed since the only required ingredients—the atomic forces—are computed throughout the PIMD trajectory;
5. can be accounted for on-the-fly or *a posteriori* using any existing code that can perform PIMD simulations.

Letting ϕ be an estimator defined as

$$\phi \equiv \frac{\hbar^2\beta^2}{24P^3}\sum_i\frac{1}{m_i}\sum_s\left(\vec{f}_i^{\,(s)}\right)^2 \quad (5)$$

such that $\Delta F = \langle\phi\rangle$ in Eq. (4) and $\mathrm{D}(A,B)$ be a generic correlation function defined as

$$\mathrm{D}(A,B) = \langle AB\rangle - \langle A\rangle\langle B\rangle, \quad (6)$$

one can write the PPI corrections for the total (ΔE), potential (ΔU), and kinetic (ΔK) energy components as well as the (constant volume) heat capacity (ΔC) estimators[53] in the following

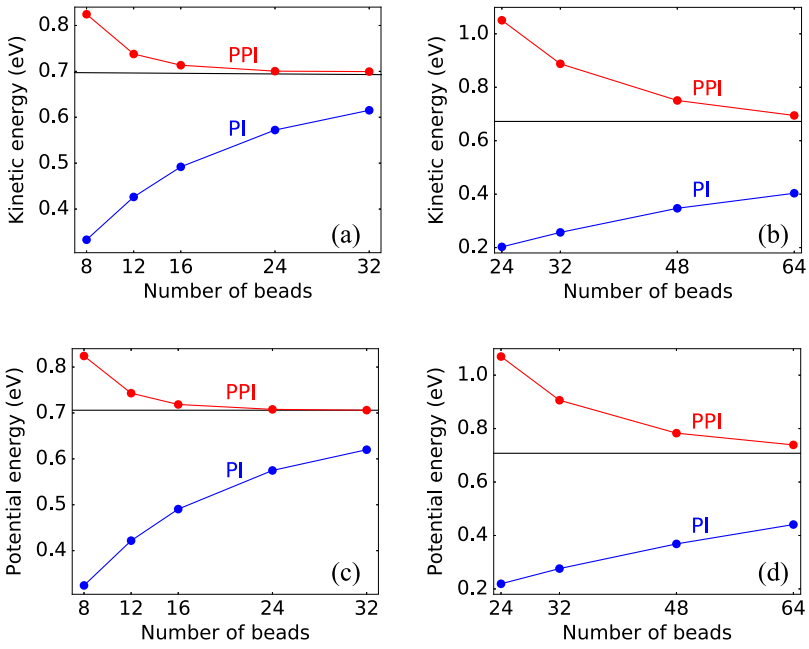

FIG. 1. Average kinetic [(a) and (b)] and potential [(c) and (d)] energies of the CH_5^+ molecule as a function of P for PIMD simulations carried out at 100 K [(a) and (c)] and 20 K [(b) and (d)] using the POSflex force field.[71] The solid black lines represent the converged results.[69]

transparent fashion:

$$\Delta E = 3\Delta F - \beta\, \mathrm{D}\,(\epsilon,\, \phi)\,, \tag{7}$$

$$\Delta U = 2\Delta F - \beta\, \mathrm{D}\,(u,\, \phi)\,, \tag{8}$$

$$\Delta K = \Delta F - \beta\, \mathrm{D}\,(\kappa,\, \phi)\,, \tag{9}$$

and

$$\Delta C = 2\,(5 + 3\beta E)\,\beta\Delta F - 2\,(3 + \beta E)\,\beta\Delta E + 2\beta\Delta K - \beta^3\, \mathrm{D}\left(\phi, \epsilon^2\right). \tag{10}$$

In these expressions, ϵ is the conventional *primitive* total energy estimator,

$$\epsilon = u + \kappa, \tag{11}$$

written in terms of κ, the conventional *primitive* kinetic energy estimator,

$$\kappa = \frac{3PN}{2\beta} - \sum_i \frac{m_i \omega_P^2}{2} \sum_s \left(\vec{q}_i^{\,(s+1)} - \vec{q}_i^{\,(s)}\right)^2, \tag{12}$$

and u, the conventional potential energy estimator,

$$u = \frac{1}{P} \sum_s U_s. \tag{13}$$

Figure 1 demonstrates the convergence of the kinetic and potential energy components as a function of P for the CH_5^+ molecule (using the POSflex force field[71]) at both 100 K and 20 K, obtained from PIMD simulations employing second-order estimators with and without PPI corrections. From these data, one can easily see that the PPI method provides a simple and efficient way to improve the accuracy of conventional second-order PI estimators.

III. EFFECTIVE TEMPERATURES FOR NUCLEI OR COLLECTIVE VIBRATIONAL MODES

Despite the fact that NQE are present in all molecules and materials, their role in a particular system or process of interest is not always clear. Thus, an important challenge in the theoretical modeling of NQE is that of quantitatively ascertaining the degree of "quantumness" associated with different degrees of freedom. In fact, such estimates can be used as indicators for approximating the convergence of PIMD simulations and therefore aid in choosing the best approach for modeling NQE in a given system. In this section, we demonstrate that the PPI method provides a simple and elegant tool to address this challenge. We note here that the results described herein can only be obtained by considering the PPI method as perturbation theory applied to conventional second-order PIMD. Such "effective temperatures" cannot be introduced by assuming that PPI is a second-order cumulant decomposition of the fourth-order Takahashi-Imada PI approach, which is an alternative derivation of Eq. (4) (for more details, see Ref. 53).

To explore the concept of "quantumness" associated with the different degrees of freedom in a system, we will consider the influence that interparticle interactions exert on NQE by analyzing the expansion of the free energy F of a quantum system in terms of powers of the reduced Planck's constant ($\hbar$),[70]

$$F = F_{\mathrm{c}} + \frac{\hbar^2\beta^2}{24} \sum_i \frac{\left\langle \vec{f}_i^{\,2} \right\rangle}{m_i} + \mathcal{O}(\hbar^3), \tag{14}$$

where F_{c} is the classical free energy. The first non-classical term, which is $\mathcal{O}(\hbar^2)$ on the RHS of Eq. (14), is proportional to the average of the square forces acting on the particle. Thus, strong attractions (or repulsions) acting on light nuclei can lead to pronounced NQE even at relatively high temperatures. For instance, a H atom bonded to an O atom is usually noticeably delocalized due to the combination of the relatively small mass of H and the large forces acting on it. As a result, quantum delocalization of H atoms in protonated water is an important factor in the formation of long water chain structures at ambient conditions.[72] Another example of the crucial role played by NQE occurs in proton transport through pristine graphene. In this case, strong repulsion at small proton-graphene distances leads to pronounced proton delocalization and the main transport mechanism is best described as quantum under-barrier tunneling.[73]

As was demonstrated in Ref. 53, Eq. (14) can also be utilized for imaginary-time PIMD approaches. In this case, the classical free energy term, F_c, transforms into the free energy of a ring-polymer system. For the second quasi-classical term in Eq. (14), one has to perform the summation over all of the force components in every bead, and the temperature becomes P times the equilibrium temperature, T, i.e., β is replaced by $\beta_P \equiv \beta/P$. As a result, this term transforms into the free-energy correction entering Eq. (4).

Following the standard perturbation-theory procedure,[70] one can obtain the following correction to the probability distribution function for a particular component j of the momentum associated with a general degree of freedom in the system (e.g., an individual atom, a phonon mode, etc.) in the following form:

$$dw_{p_j} \sim \exp\left[-\frac{\beta_P p_j^2}{2m_j}\left(1 - \frac{\hbar^2\beta_P^3}{12m_j} \sum_s \left\langle \left(f_j^{(s)}\right)^2 \right\rangle\right)\right] dp_j. \tag{15}$$

Here, the same corrections correspond to identical degrees of freedom in all beads due to their equivalence within the second-order PI approach.

The distribution in Eq. (15) has a Maxwellian form with an effective temperature given by

$$T_j^{\mathrm{eff}} = T + \frac{\hbar^2\beta_P^2}{12m_j P k_B} \sum_s \left\langle \left(f_j^{(s)}\right)^2 \right\rangle, \tag{16}$$

where k_B is the Boltzmann constant. The difference $\Delta_j \equiv T_j^{\mathrm{eff}} - T$ computed using MD simulations with classical nuclei therefore suggests a measure of the quantumness associated with the j-th degree of freedom. As such, when $\Delta_j \ll T$, the corresponding degree of freedom has a classical nature. For $\Delta_j < T$, classical MD trajectories can be used to compute the equilibrium properties of the system by introducing quantum corrections, see Secs. II and IV. Finally, when $\Delta_j \geq T$, one has to perform PIMD simulations for an accurate description of any process where this particular degree of freedom plays an important role. Note that different effective

temperatures may correspond to different degrees of freedom associated with the same nuclei and are indicative of *anisotropic* nuclear quantum fluctuations in molecular and crystal systems.

As an example which clearly demonstrates how this effective temperature formalism works, let us now consider a one-dimensional (1D) classical harmonic oscillator (HO) with frequency ω. Computing the average square force using the classical partition function yields

$$T_{\mathrm{HO}}^{\mathrm{eff}} = T + \frac{\hbar^2\omega^2}{12k_BT}. \tag{17}$$

By substituting $T_{\mathrm{HO}}^{\mathrm{eff}}$ back into the classical partition function and computing the average total energy, one obtains

$$E = \frac{1}{\beta} + \frac{\beta\hbar^2\omega^2}{12}, \tag{18}$$

a result that coincides with the first two terms in the expansion of the exact total energy of the 1D quantum harmonic oscillator (QHO) in powers of $\hbar$.

Now, we will apply this effective temperature formalism to a ring-polymer system. In this case, the first three terms in the expansion of the total energy of the 1D QHO (in powers of $\hbar$), obtained as a result of imaginary-time PI calculations with $P = 2$, were found as

$$E_{\mathrm{PI}} = \frac{1}{\beta} + \frac{\beta\hbar^2\omega^2}{16} - \frac{\beta^2\hbar^3\omega^3}{256}. \tag{19}$$

Note here that the numerical prefactors in the second and third terms of Eq. (19) are rather different from the exact values of 1/12 and 1/720. By employing Eq. (16) with the two-bead ring-polymer to compute $\left\langle \left(f_i^s\right)^2 \right\rangle$ and then recalculating the total energy using the same partition function with this new effective temperature, one finds that

$$E_{\mathrm{PPI}} = \frac{1}{\beta} + \frac{\beta\hbar^2\omega^2}{12} - \frac{\beta^2\hbar^3\omega^3}{768}, \tag{20}$$

which is in much better agreement with the exact results. In the PPI approach, only the numerical prefactor in the third term is slightly smaller than the exact value of 1/720.

This simple example above demonstrates that the effective temperatures introduced by Eq. (16) are also applicable for a ring polymer. Thus, one can use Δ_j as a measure of the convergence of PIMD results without the need to run increasingly more expensive PIMD simulations with larger values of P. As such, the employed Trotter number is deemed sufficient for an accurate quantitative description of the NQE when the auxiliary ring-polymer system displays pronounced classical behavior. In this effective temperature formalism, this is equivalent to the requirement that all of Δ_j [obtained using Eq. (16)] are small compared to T. Since this approach does not require any *a priori* knowledge regarding the characteristic frequencies of the system, the PPI effective temperature method can be very useful for large systems. As mentioned earlier, this effective temperature formalism can be applied to a given Cartesian degree of freedom as well as any collective vibrational degrees of freedom (e.g., the normal modes of the system).

TABLE I. Effective nuclear temperatures (in K) extracted from AI-MD simulations at the PBE+TS-vdW level of theory performed at 300 K.

System	H	D	O	C
Water	2866	...	602	...
Benzene	2385	...	...	898
H/D-doped PGS	1732	1049	...	907

A. Effective temperatures for classical MD

Table I presents the effective temperatures for different atomic species computed from classical *ab initio* MD (AI-MD) simulations at 300 K for a water molecule, benzene molecule, and pristine graphene sheet (PGS) doped with protons and deuterons (see Fig. 2). The systems considered here are of great practical interest and demonstrate pronounced NQE even at ambient conditions.[11,13,14] We note in passing that the system shown in Fig. 2 does not correspond to a stable state, although such a system plays an important role in the study of hydrogen isotope transport.[73]

The calculations here (and below) have been performed using the open-source i-PI code[74] for generating MD and PIMD trajectories and the all-electron FHI-aims package[75] to compute DFT energies and forces. The non-empirical exchange-correlation functional of Perdew, Burke, and Ernzerhof[76] (with the "tight" basis set setting) and the Tkatchenko-Scheffler method to account for van der Waals interactions[77] have been employed in all cases. All simulation time steps were set to 0.2 fs. For simulations of the PGS doped with protons (or deuterons), the out-of-plane coordinate of the centroid created by the dopant (H/D) ion and two C atoms in the surrounding hexagon were fixed at the same height, whereas all other degrees of freedom were allowed to fluctuate freely.

Considering the data in Table I, one clearly sees that the obtained ratios, Δ_j/T, are in the range of 1.0–8.6, which unambiguously indicates that all of the nuclei in these systems require a full quantum mechanical treatment. Furthermore, a closer look at the different components of the effective temperature corresponding to the in-plane and out-of-plane components of the H, D, and C atoms in the doped PGS system reveals considerable anisotropy in the NQE, as shown in Table II. Here, we observe that the NQE for H and D are strongly pronounced in the out-of-plane direction, while the in-plane components of the effective temperatures are

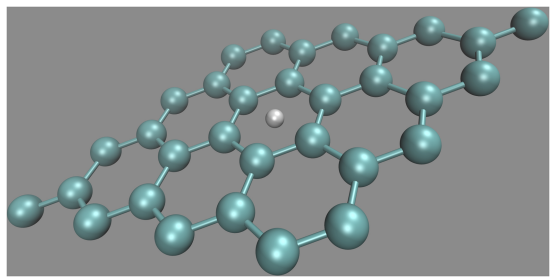

FIG. 2. Pristine graphene sheet (PGS) doped with an in-plane proton or deuteron (depicted by the gray sphere).

TABLE II. Effective nuclear temperatures (in K) corresponding to the in-plane and out-of-plane nuclear components in H/D-doped PGS extracted from AI-MD simulations at the PBE+TS-vdW level of theory performed at 300 K.

System	H	D	C
H/D-doped PGS (in-plane)	1306	794	2246
H/D-doped PGS (out-of-plane)	2584	1560	493

TABLE III. Effective nuclear temperatures (in K) extracted from the AI-PIMD simulations at the PBE+TS-vdW level of theory performed at 300 K.

System	H	D	O	C
Water	431	...	316	...
Benzene	381	...	...	316
H/D-doped PGS	380	326	...	318

approximately twice smaller. Indeed, the graphene-ion repulsion perpendicular to the graphene sheet is the major contribution to the force acting on the ions, leading to strong delocalization of these nuclei in this direction. In contrast, the in-plane forces are considerably stronger than the out-of-plane forces for the C atoms, resulting in an in-plane effective temperature that is almost five times higher. Hence, the in-plane fluctuations of the C atoms should have a pronounced quantum nature and will considerably deviate from classical predictions, while the out-of-plane modes will be significantly less affected by NQE.

This is clearly demonstrated in Fig. 3, where a comparison of the classical (AI-MD) and quantum (AI-PIMD) in-plane and out-of-plane fluctuations is shown for the PGS system at 300 K. From this figure, one can see that the in-plane quantum delocalization of the C atoms in graphene is at least as important as thermal vibrations. In contrast, the out-of-plane thermal fluctuations obviously dominate over nuclear quantum fluctuations.

B. Effective temperatures for PIMD

The effective temperatures extracted from AI-PIMD simulations performed at 300 K for a water molecule, a benzene molecule, and the H/D-doped PGS system are presented in Table III. Here, we employed $P = 10$ beads for the AI-PIMD simulations of the PGS system and $P = 12$ beads for the water and benzene molecules. All other simulation details are the same as that used for the AI-MD simulations described directly above.

From these data, one can see that the effective temperatures for D, O, and C exceed the equilibrium temperature of 300 K by less than 10%. In contrast, the effective temperatures for the H nuclei are still higher than 300 K by 27%–44%. This means that the employed Trotter number is not sufficient to obtain accurate predictions for processes involving H in these systems. Thus, either an *a posteriori* correction method should be employed to compute accurate thermodynamic properties for these systems[53,63] or an AI-PIMD simulation should be repeated with a larger value of P.

Figure 4 shows a comparison of the probability distribution functions for the nearest C–C and C–H distances in the benzene molecule computed from AI-PIMD trajectories with $P = 12$ using the conventional second-order estimator (blue line) and the corresponding converged results (red line). In full agreement with the conclusions based on the effective temperatures in Table III, the C–C distribution curve obtained from AI-PIMD simulations with $P = 12$ is in very good agreement

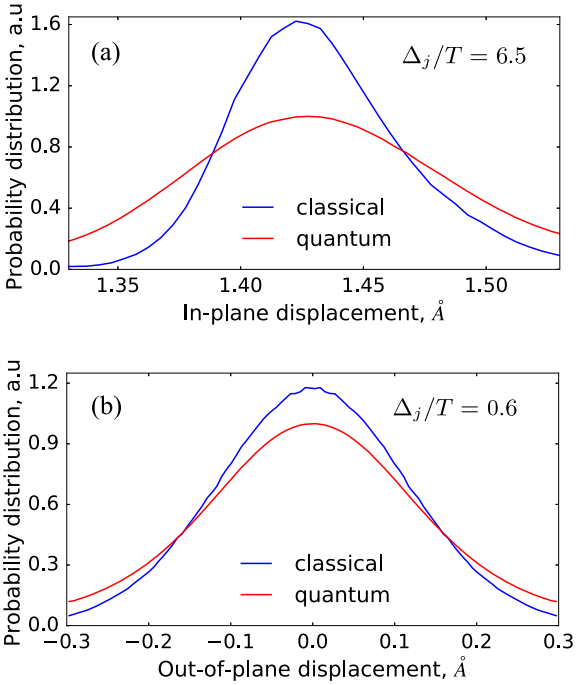

FIG. 3. Probability distribution functions [in arbitrary units (a.u.)] for finding two nearest C atoms at the given (a) in-plane distance and (b) out-of-plane distance in the PGS system at 300 K.

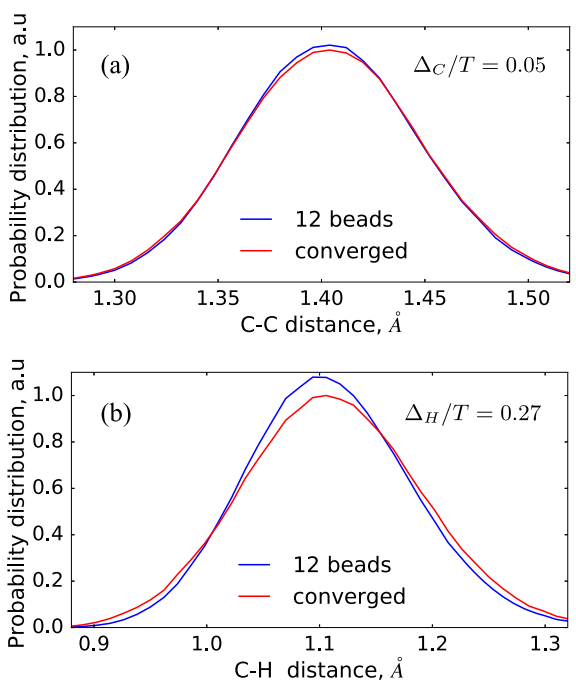

FIG. 4. Probability distribution functions [in arbitrary units (a.u.)] for finding two nearest (a) C atoms and (b) C and H atoms at a given distance in the benzene molecule at 300 K.

with the converged result. In the same breath, the C–H fluctuations noticeably deviate from the accurate results due to an underestimation of the NQE for H atoms in the AI-PIMD simulations with $P = 12$.

Knowledge of these effective nuclear temperatures already provides enough information to predict some accurate (and well-converged) thermodynamic properties using non-converged simulation results. For instance, the relative difference between the effective and true temperatures for C in the AI-PIMD simulations of the PGS system at 300 K is 6%. In this system, the NQE contribution to the kinetic energy obtained using the conventional virial centroid estimator is 52.8 meV/atom, while the converged answer is 56.9 meV/atom. Thus, the kinetic energy is underestimated by approximately 7%, which is in very good agreement with the Δ_j/T ratio. The same analysis for the benzene molecule leads to a relative underestimation of the NQE contribution to the kinetic energy by 12% in the AI-PIMD simulations with $P = 12$ at 300 K. This value is in between the values of Δ_j/T for the C and H nuclei in the benzene molecule, which were found to be 5% and 25%, respectively. This contribution is also quite close to the expected mean error of 15%. Hence, one can use these effective temperatures to check the convergence of PIMD simulations with respect to P as well as quantitatively estimate the converged results. A more accurate calculation of different thermodynamic observables is possible by employing PPI corrections for the properties of interest, which brings us to the topic of Sec. IV.

IV. PERTURBED PATH-INTEGRAL CORRECTION FOR THE RADIAL DISTRIBUTION FUNCTION

Equilibrium structural properties, such as the radial distribution function (RDF), are of great practical interest in the study of condensed-phase systems. The central quantity required for obtaining a RDF is the number of pairs of particles (of a given type) separated by a distance r in a given system of interest. In practice, one counts all pairs $N(r)$ with an interparticle distance in the range $(r - \Delta r/2,\ r + \Delta r/2)$, namely,

$$N(r) = \left\langle \hat{N}(r) \right\rangle = \left\langle {\sum_{ij}}' \frac{1}{P} \sum_s \hat{n}\left[r_{ij}^{(s)} - r \right] \right\rangle, \tag{21}$$

where $\langle \cdots \rangle$ represents an average over PIMD trajectories, the primed summation indicates omission of the $i = j$ term, $r_{ij}^{(s)}$ is the distance between particles i and j in bead s, and the operator

$$\hat{n}\left[r_{ij}^{(s)} - r \right] = \lim_{\Delta r \to 0} \frac{1}{\Delta r} \left(\sigma\left[r_{ij}^{(s)} - (r + \Delta r/2) \right] - \sigma\left[r_{ij}^{(s)} - (r - \Delta r/2) \right] \right) \tag{22}$$

is the difference between two Heaviside step (σ) functions. When $\Delta r \to 0$, the RHS of Eq. (22) transforms into a delta (δ) function, which is the standard mathematically rigorous definition of $\hat{n}\left[r_{ij}^{(s)} - r \right]$ in the continuous limit.

Using Eq. (12) from Ref. 53, the PPI correction for $N(r)$ can be written as

$$\Delta N(r) = -\frac{\hbar^2 \beta^3}{24 P^3} \mathrm{D}\left(N(r), \sum_i \frac{1}{m_i} \sum_s \left(\vec{f}_i^{(s)} \right)^2 \right) - \frac{\hbar^2 \beta^2}{12 P^3} \sum_i \frac{1}{m_i} \sum_s \left\langle \vec{f}_i^{(s)} \cdot \left(\frac{\partial N(r)}{\partial \vec{r}_i^{(s)}} \right) \right\rangle, \tag{23}$$

where the sum over i includes *all* particles in the system. Since the correlation function D on the RHS of Eq. (23) is trivial to compute, we will focus our attention on the second term. Here, the scalar product vanishes if the i-th particle does not belong to the set which enters $\hat{N}(r)$, otherwise

$$\vec{f}_i^{(s)} \cdot \left(\frac{\partial N(r)}{\partial \vec{r}_i^{(s)}} \right) = \frac{\vec{f}_i^{(s)}}{P} {\sum_j}' \frac{\vec{u}_{ij}^{(s)}}{\Delta r} \left(\delta\left[r_{ij}^{(s)} - (r + \Delta r/2) \right] - \delta\left[r_{ij}^{(s)} - (r - \Delta r/2) \right] \right), \tag{24}$$

where $\vec{u}_{ij}^{(s)} \equiv \vec{r}_{ij}^{(s)} / r_{ij}^{(s)}$ is a unit vector along the direction connecting particles j and i. Hereafter we omit the limit $\Delta r \to 0$ for clarity. In contrast to Eq. (23) above, the summation here is taken over the particles which contribute to $N(r)$ only.

Substituting Eq. (24) into the second term on the RHS of Eq. (23), representing the δ functions as the difference between two σ functions divided by $\Delta r'$ [as was done in Eq. (22)], and setting $\Delta r' = 2\Delta r$, one obtains the final expression for $\Delta N(r)$,

$$\Delta N(r) = -\frac{\hbar^2 \beta^3}{24 P^3} \mathrm{D}\left(N(r), \sum_i \frac{1}{m_i} \sum_s \left(\vec{f}_i^{(s)} \right)^2 \right) + \frac{\hbar^2 \beta^2}{24 P^4} {\sum_{ij}}' \sum_s \left\langle \left(\frac{\vec{f}_i^{(s)}}{m_i} - \frac{\vec{f}_j^{(s)}}{m_j} \right) \frac{\vec{u}_{ij}^{(s)}}{\Delta r} \times \left(\hat{n}\left[r_{ij}^{(s)} - (r - \Delta r) \right] - \hat{n}\left[r_{ij}^{(s)} - (r + \Delta r) \right] \right) \right\rangle. \tag{25}$$

The $\hat{n}$ operators in the second term on the RHS of Eq. (25) indicate that the value found for the ith bin of the $N(r)$ distribution should be added to the previous $(i - 1)$ bin and subtracted from the next $(i + 1)$ bin.

Dividing $N(r)$ by the volume of a spherical shell with radius dr yields the desired RDF

$$g(r) = 4\pi r^2 N(r) dr, \tag{26}$$

which is subject to the following normalization condition:

$$\frac{4\pi}{V} \int_V dr\, r^2 N(r) = N, \tag{27}$$

where N is the total number of particles in the system and V is the total volume. Since Eq. (21) already satisfies this normalization condition, we have

$$\frac{4\pi}{V} \int_V dr\, r^2 \Delta N(r) = 0. \tag{28}$$

Moreover, the corresponding integral contributions to $\Delta N(r)$ from both terms in Eq. (25) are also independently equal to

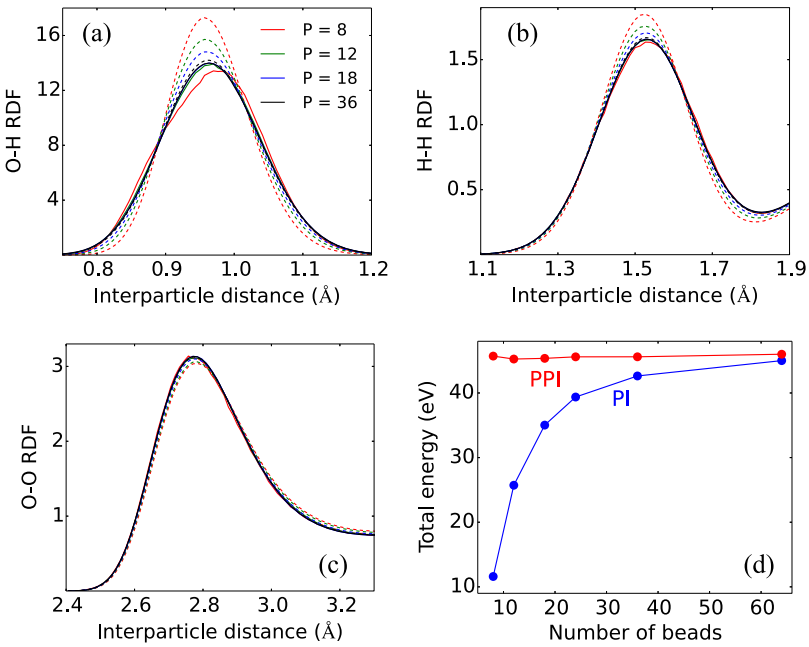

FIG. 5. O–H (a), H–H (b), and O–O (c) radial distribution functions (RDFs) as well as the total energy (d) as a function of P for PIMD simulations of liquid water at 298 K with (solid lines) and without (dashed lines) the corresponding PPI correction. In this work, liquid water was modeled using the q-TIP4P/F force field as implemented in the i-PI code and a simulation cell containing 216 water molecules subject to periodic boundary conditions.

zero. This is due to the fact that Eq. (27) implies that the correlation function in Eq. (25) vanishes since one of its arguments is a constant. To fulfill the condition in Eq. (28), the second integral on the RHS of Eq. (25) must also vanish. This simply reflects the fact that any RDF correction should conserve N and only modify the shape of the curve.

The efficiency of the PPI RDF correction in Eq. (25) is demonstrated in Fig. 5 by considering the equilibrium structure of bulk liquid water via PIMD simulations. To model liquid water, all calculations used the q-TIP4P/F water model[28] in the i-PI code[74] to perform *NVT* simulations on a box containing 216 water molecules subject to standard periodic boundary conditions. All simulations were carried out at 298 K with a time step of 0.1 fs for a total of 40 ns (with an additional 10 ns for equilibration). The results of applying conventional and *improved* PPI-based RDF estimators for the same PIMD trajectories are shown in Figs. 5(a)–5(c). In Fig. 5(d), we also show the convergence of the conventional virial centroid total energy estimator as a function of P with and without the PPI correction from Eqs. (9) and (10).

From these data, one can see that $P = 12$ is already sufficient to obtain highly converged RDF curves in addition to the total energy when PPI corrections are added to the conventional estimators. Without such PPI-corrected estimators, accurate RDFs require at least $P = 64$ beads to properly reproduce the first O–H peak. The results for the conventional RDF estimators applied to the PIMD trajectories with $P = 64$ exactly coincide with the PPI estimators starting from $P = 12$ onward and are not shown in the figure. *Quite importantly*, the exclusion of H (O) forces in Eq. (25) would lead to incorrect RDFs since the positions of the O (H) atoms do correlate with the forces acting on the H (O) atoms.

Figure 5(d) demonstrates that the convergence of the total energy is even more challenging. In this case, simulations with up to 64 beads still do not lead to fully converged results within the conventional second-order PIMD approach. Deployment of PPI corrections significantly improves the situation. In fact, PPI-corrected simulations with $P = 12$ are sufficient to accurately reproduce the first O–H peak, which is the largest challenge due to the pronounced quantum behavior of H atoms in water. All of the other computed properties can be reproduced with reasonable accuracy (~3%–5%) and for most practical purposes using only 8 beads, which is often acceptable even for computationally demanding AI-PIMD simulations of large realistic systems.

V. CONCLUSIONS AND FUTURE OUTLOOK

In this work, we have demonstrated the potential of thermodynamic perturbation theory to extend the applicability of (PI)MD simulations. We proposed a definition of effective temperatures (for atoms or other collective degrees of freedom) that allows one to quantitatively assess the "degree of quantumness" associated with nuclear fluctuations in molecules and materials. This approach only requires information about the interparticle forces, which are readily available in MD simulations. The proposed perturbed path-integral (PPI) formalism[53] allows for generalization of the effective temperatures to imaginary-time path integrals, which significantly accelerates the PIMD simulations and paves the way toward rigorously assessing their convergence. We demonstrate that PPI estimators can be developed for any thermodynamic observable of interest, with a particular focus on the radial distribution function estimator. We found that the strong interatomic interactions in molecules and crystals lead to pronounced nuclear quantum fluctuations at ambient conditions not only for hydrogen but also for heavier nuclei such as oxygen and carbon atoms. Due to the anisotropic nature of intramolecular and intracrystal interactions, NQE can be strongly anisotropic in such systems and cannot be correctly captured by merely increasing the overall temperature of the system. All of the employed estimators are implemented in the i-PI code[74] and can be used out of the box.

ACKNOWLEDGMENTS

I.P. and A.T. acknowledge the financial support from the Luxembourg National Research within the FNR-CORE program (No. FNR-11360857). R.A.D. acknowledges partial support from Cornell University through start-up funding and the Cornell Center for Materials Research (CCMR) with funding from the NSF MRSEC program (No. DMR-1120296).

[1] S. J. Blanksby and G. B. Ellison, Acc. Chem. Res. **36**, 255 (2003).
[2] A. M. Reilly and A. Tkatchenko, Phys. Rev. Lett. **113**, 055701 (2014).
[3] M. Rossi, P. Gasparotto, and M. Ceriotti, Phys. Rev. Lett. **117**, 115702 (2016).
[4] B. Chen, I. Ivanov, M. L. Klein, and M. Parrinello, Phys. Rev. Lett. **91**, 215503 (2003).
[5] M. E. Tuckerman, D. Marx, M. L. Klein, and M. Parrinello, Science **275**, 817 (1997).
[6] X.-Z. Li, B. Walker, and A. Michaelides, Proc. Natl. Acad. Sci. U. S. A. **108**, 6369 (2011).
[7] J. A. Morrone and R. Car, Phys. Rev. Lett. **101**, 017801 (2008).
[8] D. Marx, M. E. Tuckerman, J. Hutter, and M. Parrinello, Nature **397**, 601 (1999).
[9] D. Marx, M. E. Tuckerman, and M. Parrinello, J. Phys.: Condens. Matter **12**, A153 (2000).
[10] M. E. Tuckerman, D. Marx, and M. Parrinello, Nature **417**, 925 (2002).
[11] M. Ceriotti, W. Fang, P. G. Kusalik, R. H. McKenzie, A. Michaelides, M. A. Morales, and T. E. Markland, Chem. Rev. **116**, 7529 (2016).
[12] W. Fang, J. Chen, M. Rossi, Y. Feng, X.-Z. Li, and A. Michaelides, J. Phys. Chem. Lett. **7**, 2125 (2016).
[13] J. Schulte, R. Ramírez, and M. C. Böhm, Mol. Phys. **99**, 1155 (2001).
[14] C. P. Herrero and R. Ramírez, J. Chem. Phys. **145**, 224701 (2016).
[15] C. P. Herrero and R. Ramírez, J. Chem. Phys. **148**, 102302 (2018).
[16] R. Dupuis, M. Benoit, M. E. Tuckerman, and M. Méheut, Acc. Chem. Res. **50**, 1597 (2017).
[17] L. Kronik and A. Tkatchenko, Acc. Chem. Res. **47**, 3208 (2014).
[18] J. Landers, G. Y. Gor, and A. V. Neimark, Colloids Surf., A **437**, 3 (2013).
[19] I. B. Obot, D. D. Macdonald, and Z. M. Gasem, Corros. Sci. **99**, 1 (2015).
[20] K. Burke, J. Chem. Phys. **136**, 150901 (2012).
[21] J. Klimeš and A. Michaelides, J. Chem. Phys. **137**, 120901 (2012).
[22] J. Hermann, R. A. DiStasio, Jr., and A. Tkatchenko, Chem. Rev. **117**, 4714 (2017).
[23] A. Tkatchenko, R. A. DiStasio, Jr., R. Car, and M. Scheffler, Phys. Rev. Lett. **108**, 236402 (2012).
[24] A. Ambrosetti, A. M. Reilly, R. A. DiStasio, Jr., and A. Tkatchenko, J. Chem. Phys. **140**, 18A508 (2014).
[25] R. A. DiStasio, Jr., O. A. von Lilienfeld, and A. Tkatchenko, Proc. Natl. Acad. Sci. U. S. A. **109**, 14791 (2012).
[26] R. A. DiStasio, Jr., V. V. Gobre, and A. Tkatchenko, J. Phys.: Condens. Matter **26**, 213202 (2014).
[27] A. Ambrosetti, D. Alfè, R. A. DiStasio, Jr., and A. Tkatchenko, J. Phys. Chem. Lett. **5**, 849 (2014).
[28] S. Habershon, T. E. Markland, and D. E. Manolopoulos, J. Chem. Phys. **131**, 024501 (2009).
[29] J. Behler, J. Chem. Phys. **145**, 170901 (2016).
[30] S. Chmiela, A. Tkatchenko, H. E. Sauceda, I. Poltavsky, K. T. Schütt, and K.-R. Müller, Sci. Adv. **3**, e1603015 (2017).
[31] J. Řezáč and P. Hobza, Chem. Rev. **116**, 5038 (2016).
[32] A. Zen, S. Sorella, M. J. Gillan, A. Michaelides, and D. Alfè, Phys. Rev. B **93**, 241118 (2016).
[33] S. Jang, S. Jang, and G. A. Voth, J. Chem. Phys. **115**, 7832 (2001).
[34] K. R. Glaesemann and L. E. Fried, J. Chem. Phys. **117**, 3020 (2002).
[35] T. M. Yamamoto, J. Chem. Phys. **123**, 104101 (2005).
[36] M. E. Tuckerman, D. Marx, M. L. Klein, and M. Parrinello, J. Chem. Phys. **104**, 5579 (1996).
[37] R. O. Weht, J. Kohanoff, D. A. Estrin, and C. Chakravarty, J. Chem. Phys. **108**, 8848 (1998).
[38] M. Pavese, D. R. Berard, and G. A. Voth, Chem. Phys. Lett. **300**, 93 (1999).
[39] M. E. Tuckerman and D. Marx, Phys. Rev. Lett. **86**, 4946 (2001).
[40] S. Miura, M. E. Tuckerman, and M. L. Klein, J. Chem. Phys. **109**, 5290 (1998).
[41] A. Pérez, M. E. Tuckerman, H. P. Hjalmarson, and O. A. von Lilienfeld, J. Am. Chem. Soc. **132**, 11510 (2010).
[42] A. Pérez and M. E. Tuckerman, J. Chem. Phys. **135**, 064104 (2011).
[43] O. Marsalek and T. E. Markland, J. Phys. Chem. Lett. **8**, 1545 (2017).
[44] L. Wang, S. D. Fried, S. G. Boxer, and T. E. Markland, Proc. Natl. Acad. Sci. U. S. A. **111**, 18454 (2014).
[45] L. Kong, X. Wu, and R. Car, Phys. Rev. B **86**, 134203 (2012).
[46] G. A. Voth, D. Chandler, and W. H. Miller, J. Chem. Phys. **91**, 7749 (1989).
[47] U. W. Schmitt and G. A. Voth, J. Chem. Phys. **111**, 9361 (1999).
[48] A. R. Menzeleev, N. Ananth, and T. F. Miller III, J. Chem. Phys. **135**, 074106 (2011).
[49] B. Pamuk, J. M. Soler, R. Ramírez, C. P. Herrero, P. W. Stephens, P. B. Allen, and M.-V. Fernández-Serra, Phys. Rev. Lett. **108**, 193003 (2012).
[50] F. Paesani, S. Iuchi, and G. A. Voth, J. Chem. Phys. **127**, 074506 (2007).
[51] J. Liu, W. H. Miller, F. Paesani, W. Zhang, and D. A. Case, J. Chem. Phys. **131**, 164509 (2009).
[52] D. M. Wilkins, D. E. Manolopoulos, and L. X. Dang, J. Chem. Phys. **142**, 064509 (2015).
[53] I. Poltavsky and A. Tkatchenko, Chem. Sci. **7**, 1368 (2016).
[54] C. Vega, M. M. Conde, C. McBride, J. L. F. Abascal, E. G. Noya, R. Ramirez, and L. M. Sesé, J. Chem. Phys. **132**, 046101 (2010).
[55] M. A. Morales, J. M. McMahon, C. Pierleoni, and D. M. Ceperley, Phys. Rev. Lett. **110**, 065702 (2013).
[56] Y. Feng, J. Chen, D. Alfé, X.-Z. Li, and E. Wang, J. Chem. Phys. **142**, 064506 (2015).
[57] B. J. Berner and D. Thirumalai, Annu. Rev. Phys. Chem. **37**, 401 (1986).
[58] K. E. Schmidt and D. M. Ceperley, in *The Monte Carlo Method in Condensed Matter Physics*, Volume 71 of Topics in Applied Physics, edited by K. Binder (Springer Berlin Heidelberg, 1995), pp. 205–248.
[59] M. E. Tuckerman, in *Quantum Simulations of Complex Many-Body Systems: From Theory to Algorithms*, NIC Series, Vol. 10, edited by J. Grotendorst, D. Marx, and M. Alejandro (John von Neumann Institute for Computing, Jülich, 2002), pp. 269–298.
[60] M. Takahashi and M. Imada, J. Phys. Soc. Jpn. **53**, 3765 (1984).
[61] S. A. Chin, Phys. Lett. A **226**, 344 (1997).
[62] M. Suzuki, J. Phys. Soc. Jpn. **61**, 3015 (1992).
[63] O. Marsalek, P.-Y. Chen, R. Dupuis, M. Benoit, M. Mćheut, Z. Bačić, and M. E. Tuckerman, J. Chem. Theory Comput. **10**, 1440 (2014).
[64] M. Ceriotti, G. Bussi, and M. Parrinello, Phys. Rev. Lett. **102**, 020601 (2009).
[65] M. Ceriotti, G. Bussi, and M. Parrinello, Phys. Rev. Lett. **103**, 030603 (2009).
[66] O. Marsalek and T. E. Markland, J. Chem. Phys. **144**, 054112 (2016).
[67] V. Kapil, J. Behler, and M. Ceriotti, J. Chem. Phys. **145**, 234103 (2016).
[68] A. V. Sinitskiy and G. A. Voth, J. Chem. Phys. **143**, 094104 (2015).
[69] F. Uhl, D. Marx, and M. Ceriotti, J. Chem. Phys. **145**, 054101 (2016).
[70] L. D. Landau and E. M. Lifshitz, *Statistical Physics* (Butterworth–Heinemann, Oxford, 1980).
[71] F. Uhl, L. Walewski, H. Forbert, and D. Marx, J. Chem. Phys. **141**, 104110 (2014).
[72] H. S. Mei, M. E. Tuckerman, D. E. Sagnella, and M. L. Klein, J. Phys. Chem. B **102**, 10446 (1998).
[73] I. Poltavsky, L. Zheng, M. Mortazavi, and A. Tkatchenko, e-print arXiv:1605.06341v2 (2016).
[74] M. Ceriotti, J. More, and D. E. Manolopoulos, Comput. Phys. Commun. **185**, 1019 (2014).
[75] V. Blum, R. Gehrke, F. Hanke, P. Havu, V. Havu, X. Ren, K. Reuter, and M. Scheffler, Comput. Phys. Commun. **180**, 2175 (2009).
[76] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).
[77] A. Tkatchenko and M. Scheffler, Phys. Rev. Lett. **102**, 073005 (2009).