Check for updates

Using a monomer potential energy surface to perform approximate path integral molecular dynamics simulation of *ab initio* water at near-zero added cost

Daniel C. Elton, [ID] †*[a] Michelle Fritz[b] and Marivi Fernández-Serra [ID] *[c]

It is now established that nuclear quantum motion plays an important role in determining water's hydrogen bonding, structure, and dynamics. Such effects are important to include in density functional theory (DFT) based molecular dynamics simulation of water. The standard way of treating nuclear quantum effects, path integral molecular dynamics (PIMD), multiplies the number of energy/force calculations by the number of beads required. In this work we introduce a method whereby PIMD can be incorporated into a DFT simulation with little extra cost and little loss in accuracy. The method is based on the many body expansion of the energy and has the benefit of including a monomer level correction to the DFT energy. Our method calculates intramolecular forces using the highly accurate monomer potential energy surface developed by Partridge–Schwenke, which is cheap to evaluate. Intermolecular forces and energies are calculated with DFT only once per timestep using the centroid positions. We show how our method may be used in conjunction with a multiple time step algorithm for an additional speedup and how it relates to ring polymer contraction and other schemes that have been introduced recently to speed up PIMD simulations. We show that our method, which we call "monomer PIMD", correctly captures changes in the structure of water found in a full PIMD simulation but at much lower computational cost.

There is great interest in being able to accurately simulate liquid water at the quantum mechanical level.[1–4] The most widely used methodology for this is density functional theory. However, many density functionals fail to accurately reproduce all of the key properties of water such as its density, compressibility, and diffusion constant. Moreover, different density functionals fail in different ways. For instance, PBE creates a overstructured liquid, while many van der Waals (vdW) functionals create an understructured liquid.[5,6] There are nonetheless new meta-GGA functionals such as SCAN[7] or empirically optimized hybrid functionals such as B97M-rV[8] which are producing promising results for liquid water.

Most *ab initio* techniques are based on the Born–Oppenheimer approximation and the assumption that nuclear dynamics can be treated classically. However, over the past two decades a wide range of studies have demonstrated that this is not a good assumption for water because the OH stretching mode of water is very quantum mechanical (zero point temperature $T_z = \hbar\omega/2k_b = 2600$ K), and hydrogen nuclei are delocalized, leading to a large number of non-negligible nuclear quantum effects (NQEs) – for a recent review, see Ceriotti, *et al.*[9]

In the primary isotope effect, the OH distance is observed to be longer than the OD distance. In the secondary isotope effect, also called the Ubbelöhde effect, the H-bond donor–acceptor (oxygen–oxygen) distance R changes upon isotopic substitution. The magnitude and direction of the change depends on the strength of the hydrogen bond, due to competing quantum effects.[10–15] In particular, the zero-point motion of hydrogen in the out-of-plane direction (a type of librational motion) acts to increase R while the zero point motion of the stretching mode acts to decrease R.[11]

In materials with strong H-bonds, NQEs decrease the donor–acceptor distance (positive Ubbelöhde effect), while in materials with weaker H-bonds the opposite effect occurs (negative Ubbelöhde effect). The crossover from positive to negative Ubbelöhde effect has been estimated to be around $R = 2.6$–2.7 Å.[11,13] Both water and ice have H-bonds that lie near this crossover point,[4] and therefore the magnitude and direction of the secondary isotope effect in simulations of water and ice is particularly sensitive to the details of the

[a] Department of Physics and Astronomy, Stony Brook University, Stony Brook, New York 11794-3800, USA. E-mail: delton@umd.edu

[b] Universidad Autonoma de Madrid, 28049 Madrid, Spain

[c] Department of Physics and Astronomy, Stony Brook University, Stony Brook, New York 11794-3800, USA. E-mail: maria.fernandez-serra@stonybrook.edu

† Present address: Department of Mechanical Engineering, University of Maryland, College Park, Maryland 20740, USA.

This journal is © the Owner Societies 2019

Phys. Chem. Chem. Phys., 2019, **21**, 409–417 | 409

water geometry. The secondary isotope effect in ice is known to be positive (NQEs decrease R), leading to the anomalous isotope effects discovered by Pamuk et al.[16,17] This anomalous isotope effect occurs in several phases of ice and persists even in room temperature water.[18]

The ''gold standard'' technique for treating NQEs is path integral molecular dynamics (PIMD).[19] The sensitivity of competing quantum effects to the water geometry and degree of anharmonicity in the OH potential leads to a broad spectrum of sometimes conflicting results obtained from PIMD simulations of water with different forcefield models and DFT functionals.[11,13,20] As an example, the change in the dipole moment of H_2O when NQEs are included may be either positive or negative depending on the functional or forcefield being employed.[12,21] Because of the high cost of incorporating NQEs with PIMD, the testing of DFT functionals is often done with D_2O, where NQEs are much smaller due to the higher mass of deuterium and can therefore be ignored. This may be reasonable for testing density functionals, but the structure and dynamics of D_2O are different than H_2O due to NQEs. In the past, some people have introduced ''effective NQEs'' by raising the temperature of their DFT simulation. This can be justified theoretically for weakly interacting systems such as gases or van der Waals bonding materials,[22] but the same justification does not apply to hydrogen bonded materials. Increasing the temperature can be useful for compensating for the overstructuring of GGA functionals, but should not considered as an effective treatment of NQEs. A better option for approximately simulating NQEs is to use colored noise thermostats tuned to quantum zero point temperatures of different modes in liquid water.[23]

We note that classical forcefield models are not a rigorous way of studying NQEs because they are parametrized to experimental data, leading to a double counting of NQEs when used with PIMD simulation. Additionally, harmonic models do not allow for a change in the average OH distance from NQEs, and thus cannot capture primary or secondary isotope effects. Even worse, we have found that PIMD simulation with the harmonic model SPC-f[24] (and to a lesser extent q-SPC/Fw) shows an unphysical decrease in r_{OH},[21] which must be due to the ''curvature problem'' intrinsic to PIMD simulation. In the curvature problem, beads curve around a spherical shell of near constant r_{OH}, causing the centroid to lie in the interior, leading to a shorter r_{OH}.[21,25,26] While classical forcefields have been reparametrized specifically for use with PIMD,[27,28] and also parametrized from Born Oppenheimer ab initio simulations,[29,30] the most rigorous and computationally attainable way of studying NQEs in liquid water is by means of DFT-based PIMD simulations.

1 Path integral molecular dynamics methods

PIMD maps the partition function for the quantum mechanical system onto the partition function of a classical system with the following Hamiltonian:

$$H = \sum_{i=1}^{N} \sum_{j=1}^{N_b} \left(\frac{(p_i^j)^2}{2m_i'} + \frac{m_i \omega_n^2}{2} \left(q_i^j - q_i^{j+1} \right)^2 \right) + \sum_{j=1}^{N_b} V\left(q_1^j, \cdots, q_N^j \right) \tag{1}$$

here q_i^k are (x,y,z) vectors containing the bead coordinates and $i = 1 \ldots N$ is the atomic index and $k = 1 \ldots N_b$ is the bead index. We have put a prime on m_i' to indicate that these masses (called fictitious masses) may be different than the physical masses m_i. A full derivation and description of the PIMD method can be found elsewhere.[21] Craig and Manolopoulos argue that simply setting $m_i' = m_i$ for all i does the best job of reproducing the actual quantum dynamics, and call this methodology ''Ring Polymer Molecular Dynamics'' (RPMD). However, when RPMD is used the spectra are contaminated by spurious peaks caused by the normal mode frequencies which span the entire spectrum from 0 to $2\omega_n$ where $\omega_n = k_B T N_b / \hbar$.[31,32]

In this section we discuss different options for rescaling the fictitious masses as $m_j' = \sigma_j m_j$, where σ_j is a ''mass rescaling factor''. The rescaling is typically done in normal mode coordinates, so j indexes the normal modes of the ring polymer. The mass rescaling factor rescales the bead normal mode frequencies as $\Omega_k' = \Omega_k / \sqrt{\sigma_k}$. Ignoring thermostating choices, the major different PIMD implementations that have been introduced are distinguished solely by their choice of mass rescaling factor.[21,25]

In their original paper on PIMD,[33] Parrinello et al. choose to bring all of the non-centroid frequencies to the value of ω_n. A better approach is to scale the frequencies of the normal modes to above the highest frequency of interest in the system, thus avoiding the problem of normal mode contamination.[32,34] In effect, centroid molecular dynamics (CMD) rescales the normal modes to a very high frequency.[35] The disadvantage of this is that it requires using a very short timestep, even when an exact propagator is used to evolve the normal mode coordinates. The PIMD simulation methodology we use is called ''partially adiabatic centroid molecular dynamics'', denoted PA-CMD, because we choose an intermediate rescaling.[21,34] In most of our work we scale all normal modes to $10\,000\ cm^{-1}$, well above the overtones found at $5260\ cm^{-1}$ and $6800\ cm^{-1}$.

The other ingredient to PIMD is to attach thermostats to each degree of freedom to overcome the ergodicity problems first pointed out by Hall and Berne (1984).[36] We use Nosé–Hoover chain thermostats, with a chain length of 2. Alternatively, our code allows for Langevin thermostats to be used. The thermostating is done in normal-mode space, with the thermostats optimally tuned to each normal mode as they are in the PILE thermostat scheme of Ceriotti et al.[37] Importantly, the centroid mode is not thermostated, since doing so washes out the dynamics (as shown in Fig. 1).

1.1 The many body expansion

Our method is based on the many body expansion, which gives an exact decomposition of the potential energy into 1-body,
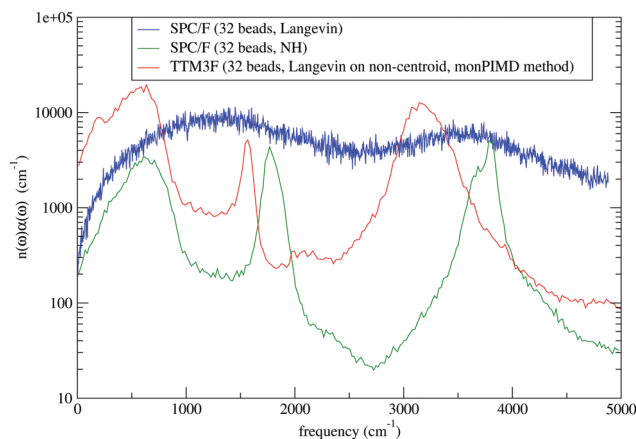
410 | Phys. Chem. Chem. Phys., 2019, **21**, 409–417

This journal is © the Owner Societies 2019

Fig. 1 Comparison of IR spectra using Langevin (PILE) and Nosé–Hoover thermostating. The IR spectra from PIMD simulation are shown for SPC/F with Langevin thermostating on all the modes, which washes out the dynamics. We do not thermostat the centroid mode with PILE, which preserves the dynamics, as shown for TTM3F and the monomer PIMD method.

2-body, 3-body, and higher order terms:

$$
\begin{aligned}
V(\{R_I\}) \;=\; & \sum_{I=1}^{N_{\text{mol}}} V_1(R_I) + \sum_{I<J}^{N_{\text{mol}}} V_2(R_I, R_J) \\
& + \sum_{I<J<K}^{N_{\text{mol}}} V_3(R_I, R_J, R_K) + \cdots
\end{aligned}
\tag{2}
$$

here R_I refers to the set of nuclear coordinates of molecule I, and N_{mol} is the number of molecules. In our method, we first subtract off the DFT monomer energies using a monomer potential energy surface (described below) fitted to the DFT functional being used. By subtracting off this term, this allows us to calculate the intramolecular energy using PIMD with the Partridge–Schwenke monomer potential energy surface (PES),[38] which is a highly accurate surface derived from CCSD calculations. This can thought of as a monomer correction to the DFT potential:

$$
\begin{aligned}
V'(\{R_I\}) = V_{\text{DFT}}(\{R_I\}) - \sum_{I=1}^{N_{\text{mol}}} V_{\text{1DFT}}(R_I) \\
+ \sum_{I=1}^{N_{\text{mol}}} V_{\text{1PS}}^{(\text{PIMD})}(R_I)
\end{aligned}
\tag{3}
$$

Therefore, the intramolecular energies and forces are calculated with PIMD, while the intermolecular forces and energies are calculated using standard techniques. The intermolecular forces on the beads are all set equal to the intermolecular forces computed from the bead centroids. Thus, in each timestep we only have to do one DFT calculation, using the centroid coordinates.

In addition to allowing for more efficient calculation of NQEs, our method has the added advantage of including a monomer correction to the DFT energy.[39] It has previously been shown that a large contribution to DFT error is in the monomer term.[40] A comparison of radial distribution functions (RDFs)

for conventional PBE and monomer-corrected PBE with 64 molecules is shown in Fig. 4. It is worth noting that in place of a monomer correction, the PES fit to the functional being used may be used instead, as may be desired for doing a rigorous comparison of different functionals with our method.

1.2 Monomer potential energy surface

The functional form of the potential energy surface developed by Patridge and Schwenke is:[38]

$$
V(r_1, r_2, \theta) = V^a(r_1) + V^a(r_2) + V^b(r_{\text{HH}}) + V^c(r_1, r_2, \theta)
\tag{4}
$$

where

$$
V^a(r) = D[e^{-2a(r-r_0)} - 2e^{-a(r-r_0)}]
$$

$$
V^b(r) = A\, e^{-br}
$$

$$
\begin{aligned}
V^c(r_1, r_2, \theta) = c_{000} + e^{-\beta[(r_1-r_e)^2+(r_2-r_e)^2]} \\
\times \sum_{ijk} c_{ijk}[(r_1-r_e)/r_e]^i[(r_2-r_e)/r_e]^j \\
\times [\cos(\theta) - \cos(\theta_e)]^k
\end{aligned}
\tag{5}
$$

here r_e and θ_e are fixed in advanced to match water's geometry and A, D, a, b, r_0, and c_{ijk} are all free parameters. As in the work of Partridge and Schwenke we truncate the polynomial expansion of V^c at $i+j \leq 8$ and $k \leq 14 - (i+j)$ for a total of 245c_{ijk}. We found that fitting this PES to DFT monomer data was the most technically challenging part of implementing our method. The fit was performed with a training set of DFT energies for 1176 monomer configurations. As was done by Partridge and Schwenke, we found that we had to fit to points calculated on a nonlinearly spaced grid, with more points where the PES changes rapidly (i.e. around r_{OH} = 0.95 Å). More specifically, we computed DFT energies at $r_{\text{OH1}}, r_{\text{OH2}} \in \{$ 0.65, 0.75, 0.85, 0.95, 0.975, 1.0, 1.05, 1.1, 1.2, 1.3, 1.5, 1.6, 1.7$\}$ Å and $\theta_{\text{HOH}} \in \{$85, 95, 100, 105, 110, 115$\}$. While Partridge and Schwenke computed their fit on a grid of points going out to only 1.4 Å, we found we had to add additional points out to 1.7 Å to obtain the correct asymptotic behaviour in the fit. Fitting to only 1.4 Å led to occasional water dissociation events in the simulation which would cause the simulation to fail. The results of our fitting are visualized in Fig. 2 and 3.
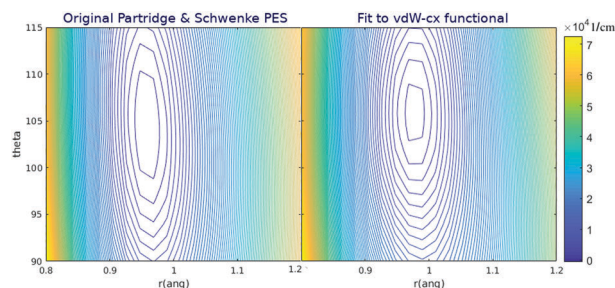

Fig. 2 The monomer potential energy surface of Partridge and Schwenke (left) and vdW-cx (right).
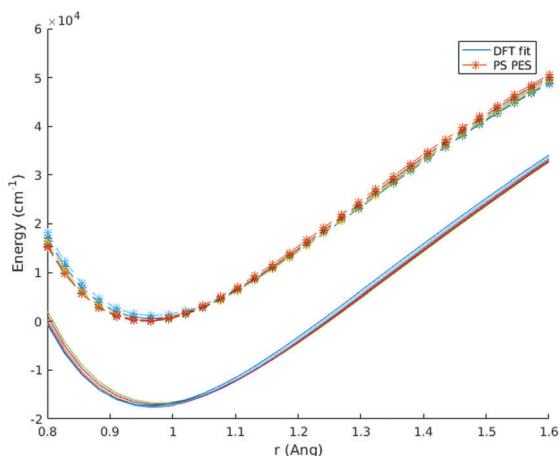
Fig. 3 Energy *vs.* r_{OH} for the case where $r_{OH1} = r_{OH2}$. Different HOH angles are shown in different colors. The Partridge and Schwenke energy surface is compared with a custom fit to PBE.
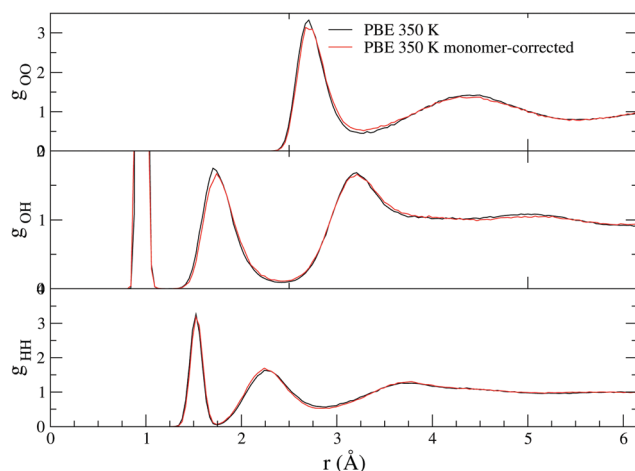

Fig. 4 Comparison of RDFs for conventional PBE and monomer corrected PBE. The simulations had 64 molecules and lengths of 35 and 27 ps, respectively.

1.3 Integration of the forces

The use of the monomer PES introduces a split between intramolecular and intermolecular forces. Tuckerman *et al.* show how to derive an integration scheme when there is a splitting between long range and short range forces.[41,42] The method is based on the classical propagator $e^{iL\Delta t}$, which exactly evolves the system from an initial phase space point $\Gamma(t) = \{\Sigma r, \Sigma p\}$ at time t to a final point at $t + \Delta t$ through $\Gamma(t + \delta t) = e^{iL\Delta t}\Gamma(t)$. The part of the Liouville operator that evolves momentum (L_p) is decomposed into short range (s) and long range (l) components:

$$
\begin{aligned}
L &= L_\gamma + L_r + L_p^s + L_p^l \\
&= L_\gamma + \sum_i^{n_a} \frac{p_i}{m_i} \cdot \frac{\partial}{\partial r_i} + \sum_i^{n_a} F_i^s \cdot \frac{\partial}{\partial p_i} + \sum_i^{n_a} F_i^l \cdot \frac{\partial}{\partial p_i}
\end{aligned}
\tag{6}
$$

here F_i^l and F_i^s are the long and short range forces on atom i, n_a is the number of atoms, L_r is the part of the operator which evolves

position and L_γ refers to the part of the operator which evolves the Nosé–Hoover thermostat. To obtain an integration method, the operator is split using the Trotter formula:

$$
e^{iL\Delta t} \approx e^{iL_\gamma \frac{\Delta t}{2}} e^{iL_p^l \frac{\Delta t}{2}} e^{iL_p^s \frac{\Delta t}{2}} e^{iL_r \Delta t} e^{iL_p^s \frac{\Delta t}{2}} e^{iL_p^l \frac{\Delta t}{2}} e^{iL_\gamma \frac{\Delta t}{2}}
\tag{7}
$$

This expression can be translated into an algorithm by reading the sequence of propagators from right to left – *i.e.* $e^{iL_p^l \frac{\Delta t}{2}}$ corresponds to an half timestep update of the momentum using the long range forces, *etc.*[41] The integration algorithm obtained is equivalent to a nested velocity-Verlet scheme. Multiple time steps (MTS) can be introduced by further splitting the inner part of eqn (7) so the short range forces are integrated M times for every time the long range forces are integrated:

$$
e^{iL\Delta t} \approx e^{iL_\gamma \frac{\Delta t}{2}} e^{iL_p^l \frac{\Delta t}{2}} \left[e^{iL_p^s \frac{\Delta t}{2M}} e^{iL_r \frac{\Delta t}{M}} e^{iL_p^s \frac{\Delta t}{2M}} \right]^M e^{iL_p^l \frac{\Delta t}{2}} e^{iL_\gamma \frac{\Delta t}{2}}
\tag{8}
$$

The inner timestep becomes $\frac{\Delta t}{M}$, where M is an integer. Lehr *et al.* have demonstrated how MTS can be implemented in Hartree–Fock calculations for water clusters by splitting the long and short range forces *via* a fragment-based approach, thus showing that MTS can be done in the context of an *ab initio* simulation.[43] When using MTS, one should be aware that resonances can occur between the fast timestep and the slower timestep. The first resonance occurs when the outer timestep becomes larger than $\Delta t_{\max} = \tau/\pi$, where τ is the period of the fastest mode in the problem. For water, this would be the OH stretch frequency ≈ 3600 cm^{-1} which leads to a value of $\Delta t_{\max} = 2.95$ fs. However, in PIMD simulation one must also consider the maximum frequency normal mode of the ring polymer when combined with the maximum OH stretching frequency, which is $\sqrt{\omega_{RP,\max}^2 + \omega_{OH,\max}^2}$. Our testing showed that the size of the outer timestep cannot go above ≈ 1.5 fs – any longer and the simulation quickly becomes unstable. However, Morrone, *et al.* have shown that the use of colored noise thermostats can stabilize resonances, offering the possibility of higher outer timesteps.[44]

2 Comparison to other methods

Our method can be understood as an extension to *ab initio* MD of the ring polymer contraction method introduced by Markland and Manoloupolos for classical MD.[45,46] In ring polymer contraction, long-range forces are analyzed using a contracted ring polymer with n' beads that are constructed by taking the n' lowest frequency ring polymer normal modes in Fourier space and transforming them into real space. Short range forces are analysed on all n beads. Our method corresponds to contraction of the long range forces to $n = 1$, *i.e.* the centroid mode (sometimes denoted as $n = 0$), and a separation between long range and short range forces that corresponds to intermolecular and intramolecular forces.

Recently, two separate groups have published a method called ''quantum ring polymer contraction'', which uses an auxiliary potential to perform *ab initio* PIMD with little added

cost.[28,47,48] The method they employed, while couched in different language, is similar to the method we present here. The principal difference is that they use self consistent charge density functional tight binding (SCC-DFTB) as the auxiliary potential in place of the monomer PES we use here. As discussed before, the use of a PES makes our simulation more accurate, while using SCC-DFTB has the opposite effect.

Recently a number of papers have been published that combine ring polymer contraction with a MTS integrator and the idea of mixing forces[49] from higher level and lower level *ab initio* methods.[47,50,51] In such methods, a lower level *ab initio* technique is used to handle the short timestep and full ring polymer, while a higher level (more expensive) technique is used with the longer timestep and contracted ring polymer. For example, in two recent studies, MP2 was combined with DFT in this manner to study small gas phase molecular systems.[50,51] A variation of this method called multilevel sampling has also been introduced and applied to FCC hydrogen, resulting in a 3–4× speedup in PIMD simulation.[52]

Another methodology introduced recently, ring polymer interpolation, achieves a 2.5–10× speedup, depending on the accuracy one desires.[53] Ring polymer interpolation could be combined with our method, resulting in a multiplicative speedup. Finally, another option for speeding up PIMD simulation is to incorporate adaptive resolution PIMD methods,[54] which allow for PIMD simulation of a small region to be combined with classical simulation of a larger region.[55]

3 Verification of the method

To verify that our method captures nuclear quantum effects with minimal losses in accuracy compared to a full PIMD simulation, we compare several observables – RDFs, dipole moments, density of states, the average bead radius of gyration, and OH distance histograms. The infrared spectrum is calculated using:[56]

$$n(\omega)\alpha(\omega) = \frac{\omega^2}{6k_B T \varepsilon_0 V c} \int_{-\infty}^{\infty} e^{-i\omega t} \langle P(0) \cdot P(t) \rangle dt \qquad (9)$$

here $\alpha(\omega)$ is the IR absorption coefficient per unit length, $n(\omega)$ is the index of refraction, and P is the dipole moment of the entire system. In PIMD simulation there are two ways to calculate the dipole moment – the first is to use the centroid positions:

$$\mu_i = \mu(\bar{r}_O, \bar{r}_{H1}, \bar{r}_{H2})$$

$$\bar{r}_i = \frac{1}{N_b} \sum_{j=1}^{N_b} r_i^j \qquad (10)$$

here r_X refers to the position of atom X, while r_i^j refers to the position of bead j in atom i. The second is to calculate the dipole moment separately for each bead "image" and then average them:

$$\mu_i = \frac{1}{N_b} \sum_{j=1}^{N_b} \mu(r_O^j, r_{H1}^j, r_{H2}^j) \qquad (11)$$

For a linear dipole function the results are the same, but for a non-linear dipole function, such as in TTM3F or DFT, the

results are not guaranteed to be the same. In practice no difference is observed between the two methods.[57] We implemented the second method (eqn (11)) because of its simplicity and because it is more in line with how estimators typically work in CMD. To calculate dipole moments for our DFT simulations, we calculated approximate dipoles using TTM3F (a polarizable model) using the centroid coordinates from DFT.[58] We found that polarization included in the TTM3F dipole model is necessary to correctly capture the intensity of the OH-stretching peak.

We calculate the "density of states" for hydrogen using the velocity–velocity autocorrelation function:

$$I(\omega) = \frac{1}{N_H} \int_{-\infty}^{\infty} e^{-i\omega t} \sum_{i=1}^{N_H} \langle v_i^H(0) \cdot v_i^H(t) \rangle dt \qquad (12)$$

The extent of delocalization of the hydrogen atoms is quantified through the radius of gyration, which is the root mean square displacement of ring polymer beads from the center of the ring:

$$r_{gyr,H} = \frac{1}{N_H N_b} \sum_{i=1}^{N_H} \sum_{j=1}^{N_b} \left\| r_i^j - r_i^c \right\| \qquad (13)$$

3.1 Tests with TTM3F

The first verification of our method was done with the polarizable TTM3F potential, which is parametrized from *ab initio* simulations and uses the PS potential energy surface natively, but modified to give the correct dissociation behaviour at large r_{OH}.[60] We simulated a system of 256 molecules for 200 ps with a 9 Å realspace Coulomb cutoff. Radial distribution functions (RDFs) are shown in Fig. 5. As has been noted elsewhere, TTM3F exhibits only small primary isotope effect and very little or no secondary isotope effect,[16] due to a lack of anharmonicity in the r_{OH} potential and competing quantum effects. Thus, the first O–O peak is only slightly lower and the nuclear quantum effects primarily manifest themselves in the broadening of the first O–H peak and decreased length of the second O–H peak,
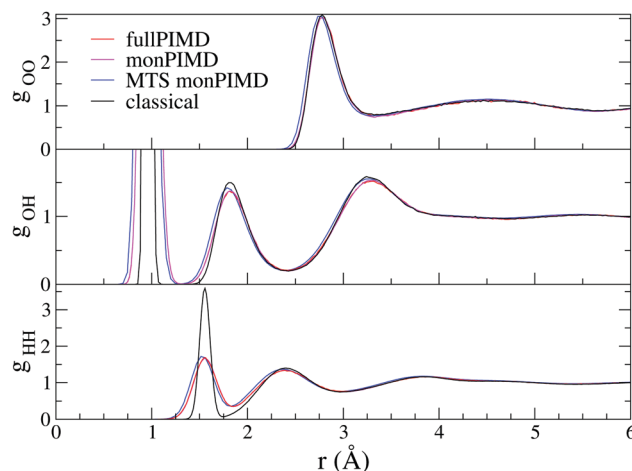
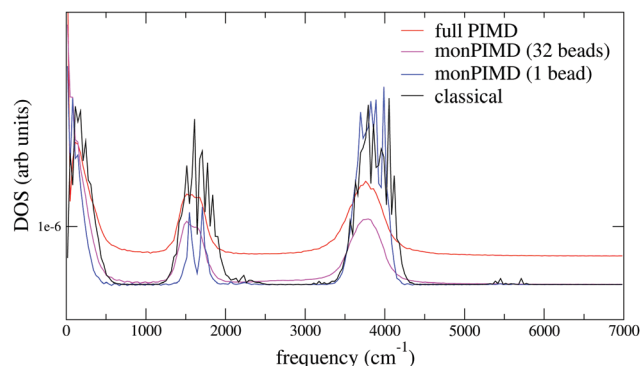Fig. 5 Validation with TTM3F: RDFs for the three methods at 300 K.

Fig. 6 Validation with TTM3F: hydrogen density of states (DOS) for one molecule (gas phase) at 300 K.

Table 1 Average OH distance, average HOH angle, average dipole moment, diffusion constant, average radius of gyration of the beads, max bead–bead OH distance and max centroid–centroid OH distance. Note: average OH distances for PIMD simulation are reported in the form centroid–centroid distance/bead–bead distance

| | TTM3F | | |
Property	Class.	FullPIMD	monPIMD
$\langle r_{OH} \rangle$	0.986	0.994/1.01	0.996/1.0
$\langle \theta_{HOH} \rangle$	105.43	105.4	105.66
$\langle \mu \rangle$	2.757	2.835	2.855
D (10^{-5} cm^2 s^{-1})	2.7	3.0	2.9
$\langle r_{gyr} \rangle$	0.0	0.1507	0.1515
Max bead r_{OH}	1.18	1.54	1.56
Max cent. r_{OH}	1.13	1.18	1.23

which indicates slightly shorter/stronger H-bonds. The monomer PIMD and full PIMD O–O RDFs are nearly the same, but the multiple time step monomer PIMD is noticeably shifted to smaller distances. The reason for this discrepancy is not clear, but very similar discrepancies are observed by Marsalek, *et al.* when applying their quantum ring polymer contraction method to RevPBE + D3.[47]

The density of states spectrum for a single molecule simulation with TTM3F is shown in Fig. 5 and 6 and the infrared spectrum for bulk water is shown in Fig. 7. Since some of the parameters of TTM3F, such as the dipole moment surface, are specifically tuned to reproduce the infrared spectrum at 300 K, the placement of the peaks in the classical simulation is quite good. When NQEs are incorporated, the OH-stretching band is redshifted and broadened. The HOH bending mode is also redshifted. Our monPIMD method reproduces the PIMD spectrum almost exactly, indicating very good capturing of NQEs. Further properties are given in Table 1. The diffusion constant of TTM3F is only slightly increased by NQEs due to competing quantum effects, as was previously discussed for TIP4P/2005f.[12] The nuclear delocalization, as measured by the radius of gyration was 1.54 Å for the full PIMD simulation and 1.56 Å for the monPIMD simulation. The max r_{OH} during the entire simulation, as measured by the centroid–centroid distance, was 1.18 Å for the full PIMD simulation and 1.23 Å for monPIMD simulation. A more
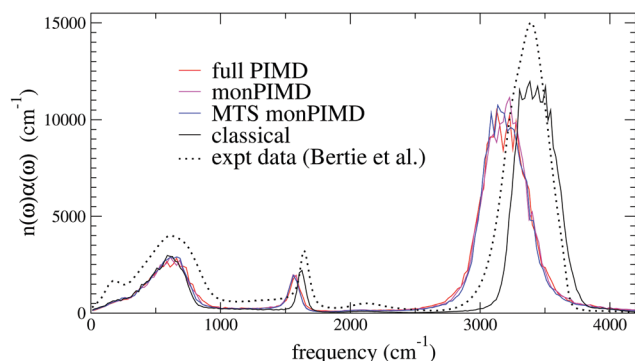
complete comparison of the bead–bead delocalization in full and monomer PIMD is obtained by looking at the histograms in Fig. 9. Together, the results in Table 1 and histograms in Fig. 9 indicate that the delocalization in the full and approximate methods are nearly the same.

3.2 Tests with DFT

We tested our method with PBE[61] and the Berland–Hyldgaard functional,[62] which is a version of the DRSLL vdW functional introduced by Dion *et al.* with modified exchange.[63] We choose this functional because of its consistent-exchange semilocal exchange choice (vdW-cx), which makes it very robust for simulating a variety of physical systems.[62] In addition, the functional performance on liquid water, has been analyzed in detail[39] and shown to be comparable to other vdW-based density functionals. We began by simulating isolated molecules with both full PIMD and monPIMD, and then progressed to simulating a pentamer cluster. The distributions of r_{OH} for the pentamer cluster simulations of vdW-cx with both full PIMD and monPIMD are shown in Fig. 9. The distributions of centroid–centroid and bead–bead r_{OH} distances are nearly the same, with slightly more delocalization observed in the full PIMD simulation as compared to monomer PIMD. Similar results were observed for PBE.

Fig. 12 shows the DOS for a single molecule simulated using the vdW-cx functional with conventional PIMD and our monomer PIMD method with 1 bead and 32 beads. The expected redshifting of the bending and stretching bands is observed, however additional peaks are observed at ≈ 2250 cm^{-1} and ≈ 5250 cm^{-1}. These frequencies correspond to the association band and the first overtone band, respectively. The association band also appears in the DOS in TTM3F simulation of bulk water (see Fig. 8) but has a tiny magnitude. The spurious enhancement of both peaks is observed both with 1 beads and 32 beads, indicating it stems from some aspect of the effective potential energy surface rather than bead normal mode contamination. Careful inspection of our fit potential energy surface did not reveal any irregularities. Attempts to refit the surface with more data points were not successful in reducing the intensity of either peak in the spectra. Interestingly, the intensity of the association band at ≈ 2100 cm^{-1}, which is due to a combination of libration and HOH bending, has been found

Fig. 7 Validation with TTM3F: infrared spectra for the three PIMD methods for 128 molecules compared to the classical spectra and experimental data at 300 K.[59]

414 | *Phys. Chem. Chem. Phys.*, 2019, **21**, 409–417
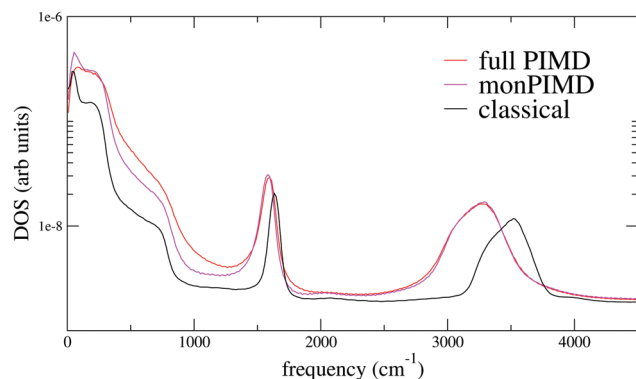
This journal is © the Owner Societies 2019

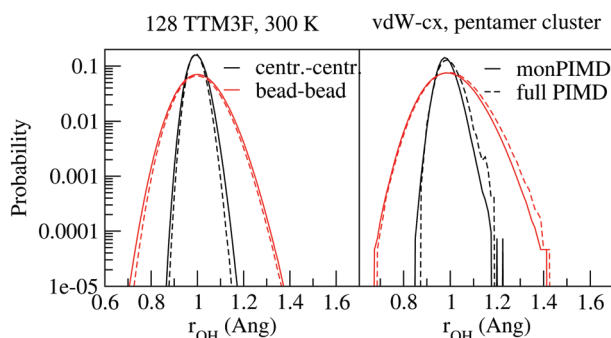Fig. 8 Validation with TTM3F: hydrogen density of states (DOS) for 128 molecules at 300 K.

Fig. 9 Histograms of the r_{OH} distance for a simulation of bulk water (128 molecules) with TTM3F and of a pentamer cluster with vdW-cx. Only slight differences are observed between full PIMD (solid lines) and the monomerPIMD method (dashed).

to be very sensitive to the coordinates used to construct the dipole moment surface and other factors such as H-bonding configuration.[64] Given the fact that PIMD is only rigorous for the calculation of equilibrium properties,[34] and that many methods
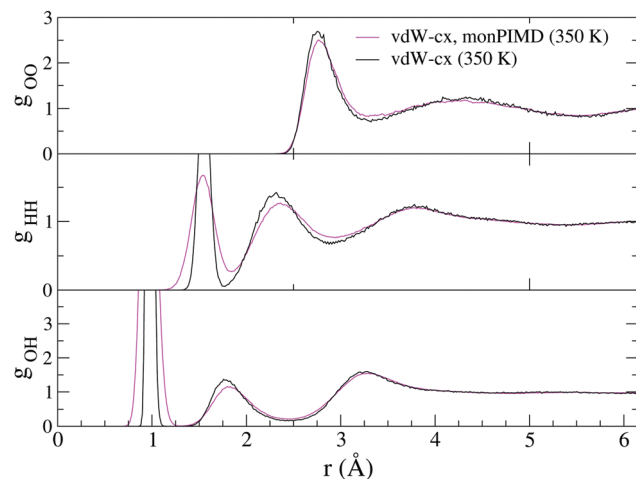

Fig. 10 Comparison of RDFs for vdW-cx simulated at 350 K with the monomer PIMD method (with the monomer correction) compared to a conventional vdW-cx simulation.

suffer from spurious peaks from normal mode contamination,[31,32] the presence of enhanced peaks in the spectrum is not as large of an issue as it may appear.

Next we performed a simulation of 64 molecules with the monomer PIMD method for both vdW-cx and PBE. A comparison of RDFs is shown in Fig. 10 for vdW-cx. We observe the

Table 2 Comparison of properties in classical *vs.* monPIMD simulation of 64H$_2$O molecules with the vdW-cx functional. Average OH distance, average HOH angle, average dipole moment, diffusion constant, average radius of gyration of the beads, max bead–bead OH distance and max centroid–centroid OH distance. Note: average OH distances for PIMD simulation are reported in the form centroid–centroid distance/bead–bead distance

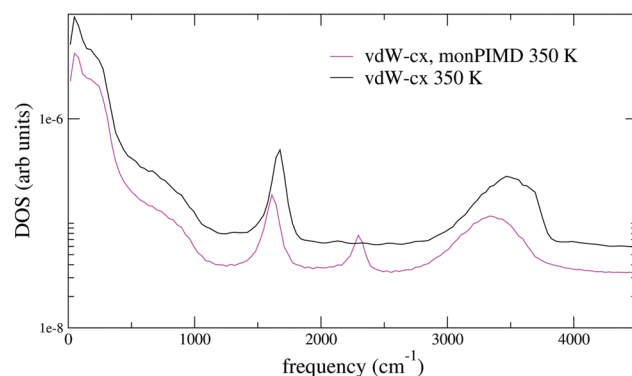| Property	Class.	monPIMD
$\langle r_{OH} \rangle$ (Å)	0.994	0.986/.997
$\langle \theta_{HOH} \rangle$	104.6	105.0/104.80
$\langle \mu \rangle$ (D)	3.68	3.66
D (10^{-5} cm^2 s^{-1})	2.3	3.4
$\langle r_{gyr} \rangle$ (Å)	0.0	0.146
Max bead r_{OH} (Å)	1.19	1.49
Max centroid r_{OH} (Å)	1.19	1.19

Fig. 11 Density of states (eqn (12)) for 64 molecules with conventional MD, compared with the monomer PIMD method (32 beads).
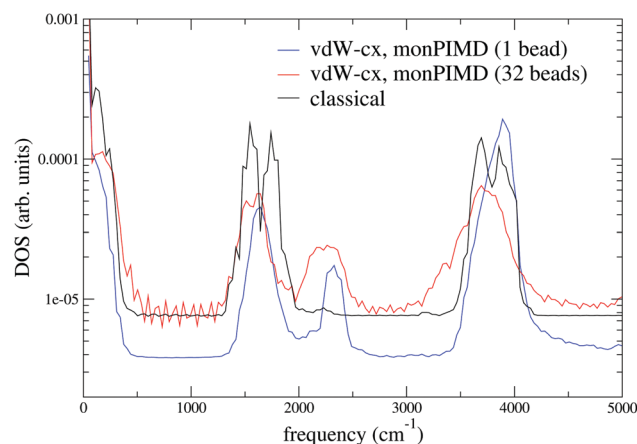

Fig. 12 Density of states (eqn (12)) for a single molecule with the vdW-cx functional simulated with traditional classical DFT and the monomer PIMD method with 1 bead and 32 beads at 350 K.

This journal is © the Owner Societies 2019

Phys. Chem. Chem. Phys., 2019, **21**, 409–417 | **415**

correct destructuring of the first O–O peak and first O–O valley as well as the expected destructuring of the O–H and H–H peaks. Information on the average water molecule geometry, dipole moment, and diffusion constant is shown in Table 2. Our simulation with monPIMD results in a slightly larger r_{OH} and HOH angle, and leads to a slightly smaller (-0.5%) dipole moment, and larger diffusion constant. The density of states for the 64 molecule vdW-cx simulation is shown in Fig. 11. Again we see that the same enhancement of the association band observed with the monomer.

4 Conclusion

We have introduced a new methodology for speeding up PIMD simulation with density functional theory and have shown that it allows for computationally tractable PIMD DFT calculations of the equilibrium properties of water. The Fortran-90 code we have written implementing this method is open source and available on GitHub.[65] In principle our method can be applied to any molecular system. The main hurdle to applying our method to other molecules is fitting an accurate potential energy surface, however recent work has shown how this can be done with neural networks.[66,67] Our method was fully validated for TTM3F simulation of water, where we showed that the method reproduces both the structure and dynamics of liquid water observed in full PIMD simulation. The advantage of our method is the $\approx 30\times$ speedup obtained, which makes *ab initio* PIMD simulations of water practical. The method nonetheless requires careful mapping and fitting of a monomer potential energy surface whenever the DFT functional or basis set is changed. However, this process is fast and easy to implement. While the structural properties of water found with our method are almost as good as the full PIMD simulation, we do observe enhancement of the association peak in the DOS for the *ab initio* functionals. We explored some possible causes for this effect and attempted to mitigate it, but more work is needed to fully understand it.

There are several variations of our method that could be explored. The first is to use a monomer DFT calculation to subtract off the monomer energies and forces (eqn (3)) and then perform monomer DFT calculations to obtain forces and energies for the monomer PIMD calculations. Doing this requires $(N_b + 1) \times N_{mol}$ additional DFT monomer calculations to be performed each timestep but has the benefit of avoiding the need for a PES and provides a more accurate representation of the DFT forces and energies. We estimate there should be at least a $2\times$ speedup over conventional PIMD with such a method, and possibly much higher. With such a method the monomer calculations can be trivially parallelized over many nodes on a cluster.

Funding

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

Notes and references

1 A. P. Gaiduk, F. Gygi and G. Galli, *J. Phys. Chem. Lett.*, 2015, **6**, 2902–2908.
2 E. Guardia, I. Skarmoutsos and M. Masia, *J. Phys. Chem. B*, 2015, **119**, 8926–8938.
3 F. Corsetti, E. Artacho, J. M. Soler, S. S. Alexandre and M.-V. Fernández-Serra, *J. Chem. Phys.*, 2013, **139**, 194502.
4 M. J. Gillan, D. Alfé and A. Michaelides, *J. Chem. Phys.*, 2016, **144**, 130901.
5 J. Wang, G. Román-Pérez, J. M. Soler, E. Artacho and M.-V. Fernández-Serra, *J. Chem. Phys.*, 2011, **134**, 024516.
6 A. Mogelhoj, A. K. Kelkkanen, K. T. Wikfeldt, J. Schiotz, J. J. Mortensen, L. G. M. Pettersson, B. I. Lundqvist, K. W. Jacobsen, A. Nilsson and J. K. Norskov, *J. Phys. Chem. B*, 2011, **115**, 14149–14160.
7 M. Chen, H.-Y. Ko, R. C. Remsing, M. F. Calegari Andrade, B. Santra, Z. Sun, A. Selloni, R. Car, M. L. Klein, J. P. Perdew and X. Wu, *Proc. Natl. Acad. Sci. U. S. A.*, 2017, **114**, 10846–10851.
8 L. Ruiz Pestana, O. Marsalek, T. E. Markland and T. Head-Gordon, *J. Phys. Lett.*, 2018, **9**, 5009–5016.
9 M. Ceriotti, W. Fang, P. G. Kusalik, R. H. McKenzie, A. Michaelides, M. A. Morales and T. E. Markland, *Chem. Rev.*, 2016, **116**, 7529–7550.
10 K. H. Kim, H. Pathak, A. Späh, F. Perakis, D. Mariedahl, J. A. Sellberg, T. Katayama, Y. Harada, H. Ogasawara, L. G. M. Pettersson and A. Nilsson, *Phys. Rev. Lett.*, 2017, **119**, 075502.
11 R. H. McKenzie, C. Bekker, B. Athokpam and S. G. Ramesh, *J. Chem. Phys.*, 2014, **140**, 174508.
12 S. Habershon, T. E. Markland and D. E. Manolopoulos, *J. Chem. Phys.*, 2009, **131**, 024501.
13 X.-Z. Li, B. Walker and A. Michaelides, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 6369–6373.
14 T. E. Markland and B. J. Berne, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 7988–7991.
15 G. Romanelli, M. Ceriotti, D. E. Manolopoulos, C. Pantalei, R. Senesi and C. Andreani, *J. Phys. Chem. Lett.*, 2013, **4**, 3251–3256.
16 B. Pamuk, J. M. Soler, R. Ramírez, C. P. Herrero, P. W. Stephens, P. B. Allen and M.-V. Fernández-Serra, *Phys. Rev. Lett.*, 2012, **108**, 193003.
17 B. Pamuk, P. B. Allen and M.-V. Fernández-Serra, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2015, **92**, 134105.
18 B. Pamuk, P. B. Allen and M.-V. Fernández-Serra, *J. Phys. Chem. B*, 2018, **122**, 5694–5706.

19 D. Chandler and P. G. Wolynes, *J. Chem. Phys.*, 1981, **74**, 4078–4095.

20 B. Chen, I. Ivanov, M. L. Klein and M. Parrinello, *Phys. Rev. Lett.*, 2003, **91**, 215503.

21 D. C. Elton, PhD thesis, Stony Brook University, 2016.

22 L. Landau and E. Lifshitz, *Statistical Physics*, Pergamon Press, 1969, p. 99.

23 S. Ganeshan, R. Ramírez and M. V. Fernández-Serra, *Phys. Rev. B*, 2013, **87**, 134207.

24 K. Toukan and A. Rahman, *Phys. Rev. B*, 1985, **31**, 2643–2648.

25 A. Witt, S. D. Ivanov, M. Shiga, H. Forbert and D. Marx, *J. Chem. Phys.*, 2009, **130**, 194510.

26 F. Paesani and G. A. Voth, *J. Chem. Phys.*, 2010, **132**, 014105.

27 S. Fritsch, R. Potestio, D. Donadio and K. Kremer, *J. Chem. Theory Comput.*, 2014, **10**, 816–824.

28 C. John, T. Spura, S. Habershon and T. D. Kühne, *Phys. Rev. E*, 2016, **93**, 043305.

29 T. Spura, C. John, S. Habershon and T. D. Kühne, *Mol. Phys.*, 2015, **113**, 808–822.

30 F. Paesani, W. Zhang, D. A. Case, T. E. Cheatham and G. A. Voth, *J. Chem. Phys.*, 2006, **125**, 184507.

31 M. Rossi, M. Ceriotti and D. E. Manolopoulos, *J. Chem. Phys.*, 2014, **140**, 234116.

32 S. Habershon, G. S. Fanourgakis and D. E. Manolopoulos, *J. Chem. Phys.*, 2008, **129**, 074501.

33 M. Parrinello and A. Rahman, *J. Chem. Phys.*, 1984, **80**, 860–867.

34 T. D. Hone, P. J. Rossky and G. A. Voth, *J. Chem. Phys.*, 2006, **124**, 154103.

35 S. Jang and G. A. Voth, *J. Chem. Phys.*, 1999, **111**, 2357–2370.

36 R. W. Hall and B. J. Berne, *J. Chem. Phys.*, 1984, **81**, 3641–3643.

37 M. Ceriotti, M. Parrinello, T. E. Markland and D. E. Manolopoulos, *J. Chem. Phys.*, 2010, **133**, 124104.

38 H. Partridge and D. W. Schwenke, *J. Chem. Phys.*, 1997, **106**, 4618–4639.

39 M. Fritz, M. Fernández-Serra and J. M. Soler, *J. Chem. Phys.*, 2016, **144**, 224101.

40 A. P. Bartók, M. J. Gillan, F. R. Manby and G. Csányi, *Phys. Rev. B*, 2013, **88**, 054104.

41 M. Tuckerman, B. J. Berne and G. J. Martyna, *J. Chem. Phys.*, 1992, **97**, 1990–2001.

42 M. Tuckerman, *Statistical Mechanics: Theory and Molecular Simulation*, Oxford University Press, UK, 2008.

43 N. Luehr, T. E. Markland and T. J. Martínez, *J. Chem. Phys.*, 2014, **140**, 084116.

44 J. A. Morrone, T. E. Markland, M. Ceriotti and B. J. Berne, *J. Chem. Phys.*, 2011, **134**, 014103.

45 T. E. Markland and D. E. Manolopoulos, *J. Chem. Phys.*, 2008, **129**, 024105.

46 G. S. Fanourgakis, T. E. Markland and D. E. Manolopoulos, *J. Chem. Phys.*, 2009, **131**, 094102.

47 O. Marsalek and T. E. Markland, *J. Chem. Phys.*, 2016, **144**, 054112.

48 O. Marsalek and T. E. Markland, *J. Phys. Lett.*, 2017, **8**, 1545–1551.

49 E. Anglada, J. Junquera and J. M. Soler, *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.*, 2003, **68**, 055701.

50 X. Cheng, J. D. Herr and R. P. Steele, *J. Chem. Theory Comput.*, 2016, **12**, 1627–1638.

51 V. Kapil, J. VandeVondele and M. Ceriotti, *J. Chem. Phys.*, 2016, **144**, 054111.

52 H. Y. Geng, *J. Comput. Phys.*, 2015, **283**, 299–311.

53 S. J. Buxton and S. Habershon, *J. Chem. Phys.*, 2017, **147**, 224107.

54 A. B. Poma and L. Delle Site, *Phys. Rev. Lett.*, 2010, **104**, 250201.

55 K. Kreis, K. Kremer, R. Potestio and M. E. Tuckerman, *J. Chem. Phys.*, 2017, **147**, 244104.

56 R. Ramírez, T. López-Ciudad, P. Kumar and D. Marx, *J. Chem. Phys.*, 2004, **121**, 3973–3983.

57 M. A. Witt, PhD thesis, Ruhr University, 2012.

58 C. Vega, *Mol. Phys.*, 2015, **113**, 1145–1163.

59 J. E. Bertie and Z. Lan, *Appl. Spectrosc.*, 1996, **50**, 1047–1057.

60 G. S. Fanourgakis and S. S. Xantheas, *J. Chem. Phys.*, 2008, **128**, 074506.

61 J. P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.*, 1996, **77**, 3865–3868.

62 K. Berland and P. Hyldgaard, *Phys. Rev. B*, 2014, **89**, 035412.

63 M. Dion, H. Rydberg, E. Schröder, D. C. Langreth and B. I. Lundqvist, *Phys. Rev. Lett.*, 2004, **92**, 246401.

64 A. B. McCoy, *J. Phys. Chem. B*, 2014, **118**, 8286–8294.

65 D. C. Elton, PIMD-F90, https://github.com/delton137/PIMD-F90.

66 K. Yao, J. E. Herr and J. Parkhill, *J. Chem. Phys.*, 2017, **146**, 014106.

67 H. Wang, L. Zhang, J. Han and W. E, *Comput. Phys. Commun.*, 2018, **228**, 178–184.

This journal is © the Owner Societies 2019

Phys. Chem. Chem. Phys., 2019, **21**, 409–417 | 417