Efficient methods and practical guidelines for simulating isotope effects

Michele Ceriotti and Thomas E. Markland

Citation: J. Chem. Phys. **138**, 014112 (2013); doi: 10.1063/1.4772676
View online: http://dx.doi.org/10.1063/1.4772676
View Table of Contents: http://jcp.aip.org/resource/1/JCPSA6/v138/i1
Published by the American Institute of Physics.

Efficient methods and practical guidelines for simulating isotope effects

Michele Ceriotti[1,a)] and Thomas E. Markland[2,b)]
[1]*Physical and Theoretical Chemistry Laboratory, University of Oxford, South Parks Road, Oxford OX1 3QZ, United Kingdom*
[2]*Department of Chemistry, Stanford University, Stanford, California 94305, USA*

(Received 26 October 2012; accepted 5 December 2012; published online 7 January 2013)

The shift in chemical equilibria due to isotope substitution is frequently exploited to obtain insight into a wide variety of chemical and physical processes. It is a purely quantum mechanical effect, which can be computed exactly using simulations based on the path integral formalism. Here we discuss how these techniques can be made dramatically more efficient, and how they ultimately outperform quasi-harmonic approximations to treat quantum liquids not only in terms of accuracy, but also in terms of computational cost. To achieve this goal we introduce path integral quantum mechanics estimators based on free energy perturbation, which enable the evaluation of isotope effects using only a single path integral molecular dynamics trajectory of the naturally abundant isotope. We use as an example the calculation of the free energy change associated with H/D and ^{16}O/^{18}O substitutions in liquid water, and of the fractionation of those isotopes between the liquid and the vapor phase. In doing so, we demonstrate and discuss quantitatively the relative benefits of each approach, thereby providing a set of guidelines that should facilitate the choice of the most appropriate method in different, commonly encountered scenarios. The efficiency of the estimators we introduce and the analysis that we perform should in particular facilitate accurate *ab initio* calculation of isotope effects in condensed phase systems. © *2013 American Institute of Physics*. [http://dx.doi.org/10.1063/1.4772676]

I. INTRODUCTION

Replacing an element with one of its isotopes is one of the most useful tools to study chemical mechanisms and equilibria. Since isotopes differ only by the number of neutrons, the potential energy surface on which they evolve, which arises from the electron-proton interactions, is unchanged. If the nuclei behaved as classical particles, moving on the Born-Oppenheimer potential energy surface, the thermodynamic properties of systems containing different isotopes would be identical. However, light nuclei exhibit nuclear quantum effects such as zero-point energy and tunnelling, which are more pronounced for lighter isotopes. As such, isotope effects provide an experimental window through which the influence of nuclear quantum effects can be probed.

Accurate and efficient simulations of isotope effects would enable the investigation of many important processes such as acid-base chemistry, shifts in phase boundaries, geological fractionation ratios and equilibrium constants of chemical reactions.[1,2] In addition, the shift in the free energy barrier along a particular reaction coordinate can be used to obtain the change in the rate of a chemical reaction via a quantum transition state theory[3–5] or, when combined with a dynamics scheme such as centroid molecular dynamics or ring polymer molecular dynamics, an approximation to the full dynamical rate constant.[6,7] Isotope effects in these processes show a range of interesting behaviors. For example, recent work on water and other hydrogen-bonded systems has demonstrated that quantum mechanical effects can act to strengthen short hydrogen bonds or to weaken long ones.[8–10] This concept of "competing quantum effects" has recently been shown to explain the inversion in the atmospherically important fractionation of hydrogen and deuterium between the liquid and vapor phases of water where at ambient temperatures D is favored in the liquid whereas at higher temperatures D is favored in the vapor.[11] Accurately describing the cancellation between the quantum effects in water requires the inclusion of anharmonic terms in the potential and therefore conclusions drawn from simple harmonic approximations can be misleading. In addition, the subtle nature of this cancellation means that fractionation ratios can be used as a sensitive test of a potential energy surface's ability to accurately predict quantum effects.

Modelling isotope effects requires the calculation of the free energy change upon isotope substitution. For some systems such as solids and simple molecular gases a reasonable approximation of the isotope effect can be gained from either a harmonic approximation or $\hbar$ perturbation.[12–14] However for liquids, which contain many anharmonic modes, or when treating more strongly quantum mechanical atoms such as H and D, these approximations can be inaccurate or inefficient.[11,15]

The exact calculation of isotope exchange free energies can be achieved by using the imaginary time path integral formalism of quantum mechanics.[16–18] This formalism exploits the isomorphism between a quantum mechanical system and a classical system in an extended "ring polymer" phase space. As such it can be applied to any system for which a classical simulation can be performed, albeit traditionally with a significant computational overhead.

a)michele.ceriotti@chem.ox.ac.uk.
b)tmarkland@stanford.edu.

A series of recent developments have dramatically decreased the computational cost of path integral molecular dynamics (PiMD) simulations. These have taken the form of the introduction of efficient integrators and thermostatting schemes to address non-ergodicity.[19–21] In addition, methods have been developed to reduce the number of force evaluations or points on the imaginary time path.[22–27] These developments along with improvements in the ability to perform computationally efficient *ab initio* calculations make on-the-fly first principles path integral molecular dynamics a feasible proposition.[28–31] Such a combination is natural, since *ab initio* calculations attempt to reproduce the exact Born-Oppenheimer potential energy surface. Whereas nuclear quantum effects are often implicitly accounted for by empirical potential energy surfaces, which are typically parameterized to reproduce experimental properties, they are completely absent in *ab initio* calculations unless they are treated explicitly.

In this work we introduce a series of developments to allow the extraction of free energy changes upon isotope substitution from path integral simulations. First, we show how the free energy change can be computed efficiently by performing a physically motivated transformation of the thermodynamic integration variables. Second, we address the failures of harmonic approximations when applied to liquids and other systems with anharmonic degrees of freedom in terms of both accuracy and, more remarkably, efficiency when compared to path integral simulations. Third, we introduce two path integral estimators, based on a free energy perturbation approach, that offer desirable features such as the ability to extract isotope free energy changes from a single path integral trajectory of the most abundant isotope. We benchmark these methods against the previously introduced thermodynamic integration approach[32] using H/D substitution in liquid water as a prototypical model of a hydrogen bonded system. In doing so, we demonstrate the strengths and pitfalls which each possess, and identify which approach may be most advantageous for a given application.

II. THEORY

Let us consider the thermodynamics of substituting an atom X of mass m in a system Φ with a different isotope of the same element, having mass m'. More specifically, we will consider the process of taking an atom in a phase, molecule, or compound, which we will label as ${}^{m}\mathrm{X}(\Phi)$, and exchanging it with an atom of the other isotope taken from a reservoir of non-interacting atoms, which we will label ${}^{m'}\mathrm{X}(\infty)$,

$$ {}^{m}\mathrm{X}(\Phi) + {}^{m'}\mathrm{X}(\infty) \rightleftharpoons {}^{m'}\mathrm{X}(\Phi) + {}^{m}\mathrm{X}(\infty). \tag{1}$$

The equilibrium is controlled by the free energy change corresponding to this process, ΔA. Here we will consider the transformation to be performed at constant volume, but the extension to constant pressure is straightforward. One can compute ΔA by thermodynamic integration, by first writing out the quantum mechanical free energy at inverse temperature β for the system Φ containing an isotope of mass μ

$$A(\mu, \Phi) = -\frac{1}{\beta} \ln \mathrm{Tr}[e^{-\beta H(\mu,\Phi)}] \tag{2}$$

and then differentiating with respect to mass, which yields an expression containing the expectation value of the quantum kinetic energy of the system,

$$\frac{\partial A(\mu, \Phi)}{\partial \mu} = -\frac{\langle T(\mu, \Phi)\rangle}{\mu}. \tag{3}$$

The thermodynamic integration for the free atoms is straightforward, and one is left with an expression for the free energy difference that reads

$$\Delta A(\Phi) = \frac{d}{2\beta} \log \frac{m'}{m} - \int_m^{m'} \frac{\langle T(\mu, \Phi)\rangle}{\mu} \mathrm{d}\mu. \tag{4}$$

Here d is the dimensionality of the problem and $\langle T(\mu, \Phi)\rangle$ is the expectation value of the kinetic energy of atom X with mass μ in the system Φ. Quantum effects stemming from particle-exchange statistics could also be included.[43] However, in this work we chose not to include them here and in the following, since they only play a role at cryogenic temperature and add considerable complexity to the calculation.

One observation that is immediately clear from this result is that if the atom behaved classically, its average kinetic energy would yield the classical equipartition value of $1/2\beta$ per degree of freedom and the free energy difference in Eq. (4) would amount to zero. In order to compute ΔA it is therefore necessary to evaluate the expectation value of the kinetic energy in a way that takes into account the quantum mechanical nature of the atoms.

A. Discretizing the free energy change

One can in practice discretize the integral in Eq. (4) and compute the integrand for a few values of the isotope mass, μ. In previous studies, a linear discretization in the isotope mass μ has been used.[11,32–34] However, one can accelerate the convergence considerably by performing a change of variables that takes into account the physical nature of the problem. Making such a change of variables acts to smooth the integrand so that a good approximation can be obtained with a very small number of integration points.

If the mass-dependence of $\langle T(\mu)\rangle$ was known, one could make the integrand $\langle T(\mu)\rangle / \mu$ constant by performing the transformation $y = f^{-1}(\mu)$ where $f(y)$ satisfies the differential equation

$$\langle T\,(f(y))\rangle\, f'(y)/f(y) = C, \tag{5}$$

where C is any constant. Clearly $\langle T(\mu)\rangle$ is not known exactly beforehand, so in practice one cannot solve Eq. (5). However, one can use physical intuition to make assumptions on the form of $\langle T(\mu)\rangle$ and derive a transformation which "flattens" the integral. A particularly useful limit for condensed phase systems is the limit where the problem is nearly harmonic and strongly quantized. In that case it can be shown that

$$\langle T(\mu)\rangle \propto \sum_i \sqrt{k_i/\mu}, \tag{6}$$

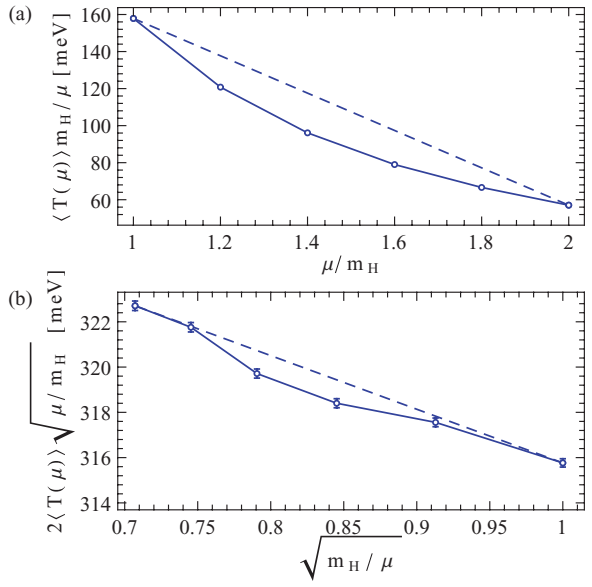

FIG. 1. Panel (a) shows the integrand of Eq. (4), as computed for a $P = 64$ PIMD simulation of liquid water with one hydrogen atom being transformed into deuterium. Panel (b) shows that the integrand becomes nearly constant when a transformed integration variable is used, as in Eq. (7). Lines are just guides for the eye, and statistical error bars are reported for each point. The dashed line shows the estimate from the two end points. Note the difference in the scale: in panel (a) the integrand varies by a factor of 2, whereas in panel (b) it varies by less than 3%.

where the k_i's are the spring constants of the various normal modes. With this assumption one obtains $f(y) \propto 4(\sum_i \sqrt{k_i})^2/(Cy + B\sum_i \sqrt{k_i})^2$, where B is an integration constant. In practice, by choosing $B = 0$ and $C = 2\sum_i \sqrt{k_i}$, the desired change of variable is simply $\mu = 1/y^2$. Equation (4) is therefore transformed into

$$\Delta A = \frac{d}{2\beta}\log\frac{m'}{m} - \int_{1/\sqrt{m'}}^{1/\sqrt{m}} 2\frac{\langle T(1/y^2)\rangle}{y}\mathrm{d}y. \tag{7}$$

Figure 1 shows the integrand of Eqs. (4) and (7) for the H/D substitution in liquid water. Using the coordinate transformation in Eq. (7) makes the integrand nearly constant, which simplifies greatly the integration. Using Eq. (4), one would obtain $\Delta A = -94.0 \pm 0.1$ meV when computing six integration points, and -107.5 ± 0.1 meV where only the two extrema are used. From Eq. (7) one would instead obtain the fully converged value $\Delta A = -93.4 \pm 0.1$ meV when using six points, and -93.5 ± 0.1 meV when using only two.

B. The pitfalls of harmonic approximations

For solids and gases nuclear quantum effects are usually accounted for using a harmonic approximation.[35] Within this approximation, the quantum kinetic energy can be estimated by first computing the normal mode frequencies ω_i and the corresponding normalized eigenvectors $\mathbf{u}^{(i)}$ and then evaluating the kinetic energy for a given atom k using[36–39]

$$T_k^{\mathrm{harm}} = \sum_{i\alpha} \left|u_{k\alpha}^{(i)}\right|^2 \frac{\hbar\omega_i}{4}\coth\frac{\beta\hbar\omega_i}{2}. \tag{8}$$

Here $u_{k\alpha}^{(i)}$ indicates the component of the eigenvector $\mathbf{u}^{(i)}$ corresponding to atom k in the Cartesian direction α. In solids, one would need to integrate over the Brillouin zone. For gas-phase molecules one can include the (classical) contributions from rotations and translations by also summing over the zero-frequency eigenvalues of the dynamical matrix. Including quantized rotations adds considerable additional complexity but is only necessary at very low temperatures.

For solids and simple gas-phase molecules this is often a good approximation which can be computed inexpensively. However, the cost of computing the Hessian grows rapidly with the number of atoms. Moreover, in systems such as liquids that contain many anharmonic modes it makes little sense to compute the Hessian for finite-temperature configurations.[40] However, one can compute the vibrational density of states for a given atom from the Fourier transform of the velocity-velocity correlation function,

$$D_k(\omega) \propto \int_{-\infty}^{\infty} \langle \mathbf{v}_k(t)\cdot\mathbf{v}_k(0)\rangle e^{-\mathrm{i}\omega t}\mathrm{d}t, \tag{9}$$

where $\mathbf{v}_k$ is the velocity of atom k. An approximation to the quantum kinetic energy can then be computed by normalizing $D_k(\omega)$ such that $\int_0^\infty D_k(\omega)\mathrm{d}\omega = 3$ and evaluating[38]

$$T_k^{\mathrm{DOS}} = \int_0^\infty D_k(\omega)\frac{\hbar\omega}{4}\coth\frac{\beta\hbar\omega}{2}\mathrm{d}\omega. \tag{10}$$

Figure 2 contrasts the values of the integrand that enter Eq. (7) for both liquid water and water vapor at 300 K, as computed exactly by direct substitution (see Sec. III B) and by quasi-harmonic analysis (Eq. (8) for the vapor and Eq. (10) for the liquid). For each phase, the harmonic approximation is close to the corresponding exact path integral result. There are however two important observations to be made.

First, one is generally interested in computing *differences* between isotope-substitution free energies such as isotopic fractionation ratios or acidity shifts, which are highly sensitive to subtle anharmonic effects. For instance the geochemi-

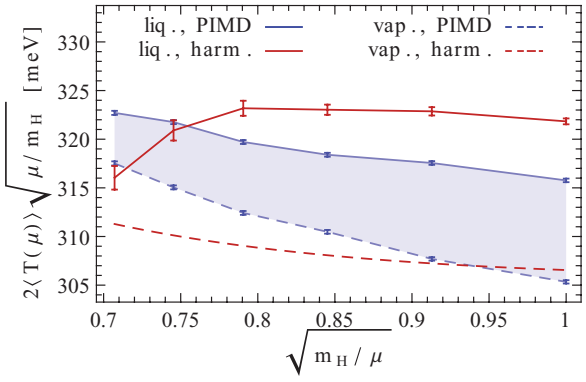

FIG. 2. The figure shows the integrand of the transformed thermodynamic integration (7) as a function of the mass of the isotope. The results from fully-converged PIMD ($P = 64$, blue lines) are compared with those from a harmonic/quasi-harmonic approximation (red lines). Curves are shown for both liquid water (full lines) and an isolated molecule in the gas phase (dashed lines). The harmonic approximation for the gas phase is based on static calculations and therefore has no error bars. For all other calculation lines are just guides for the eye, and statistical error bars are reported for each data point.

cally important process of isotopic fractionation between liquid water and its vapor corresponds to the free energy change associated with the process

$$H_2O(l) + HOD(v) \rightleftharpoons H_2O(v) + HOD(l), \tag{11}$$

where l denotes the liquid phase and v denotes the vapor phase. Typically this is expressed as a fractionation ratio,

$$\alpha_{l-v} = e^{-\beta(\Delta A(l) - \Delta A(v))}. \tag{12}$$

Using the data in Fig. 2 to compute the fractionation ratio the quasi-harmonic approximation yields a value of $10^3 \log \alpha_{l-v} = 156 \pm 10$ compared to the exact path integral result of 95 ± 1. A harmonic approximation would therefore substantially overpredict the amount to which H and D separate in the atmosphere.

Second, although one might view the quasi-harmonic approximation in Eq. (10) as a computationally inexpensive approach to obtain an estimate of the quantum kinetic energy of a liquid, this is not the case. The reason for its inefficiency is that the kinetic energy is obtained from the integral of the Fourier transform of the velocity autocorrelation function. Hence any statistical noise in the velocity autocorrelation function must be very small to avoid the appearance of large statistical errors in the quantum kinetic energy. For example, to obtain the exact results in Fig. 2 we performed a $P = 64$ bead path integral trajectory which in the most naive implementation is ~64 times more costly than a single classical simulation of the same length. Despite the fact that we computed the quasi-harmonic estimate in Eq. (10) from 64 independent classical trajectories of the same length—which thereby required the same computational effort as the path integral simulation—the statistical error of the PIMD results is 10 times smaller than that of the quasi-harmonic approximation. In other terms, for a given level of statistical accuracy, a quasi-harmonic approximation such as in Eq. (10) is 100 times more demanding than a fully converged PIMD simulation, even if we considered a number of beads twice as large as what is generally deemed converged and we did not exploit the efficient estimators we introduce in Sec. II D, ring-polymer contraction[23] or the PIGLET method.[27]

In summary, while harmonic approximations may be a viable way to crudely estimate the role of nuclear quantum effects and the change in isotope substitution free energies in the gas phase and in simple harmonic solids by means of Eq. (8) they should be treated with caution in terms of both efficiency and accuracy for liquids and systems with anharmonicity. In particular whenever something as subtle as isotope fractionation ratios are required or whenever the system is strongly anharmonic, the path integral methods we discuss in Sec. II C are not only more accurate, but paradoxically less computationally demanding than a quasi-harmonic approximation.

C. Imaginary time path integrals

To move beyond the harmonic approximation to the kinetic energy in Eq. (4) while including the quantum nature of the nuclei one can use the imaginary time path integral formulation of quantum mechanics. This approach allows one to exactly include the quantum nature of the nuclei for a system of N distinguishable particles at a finite temperature. For simplicity we shall use a notation for a single particle of mass m in one dimension, described by its position q and momentum p. The generalization to multiple degrees of freedom is routine and has been widely discussed elsewhere.[18,41] For a system with a Hamiltonian of the form $H(p, q) = p^2/2m + V(q)$, one can show that the quantum mechanical partition function $Z = \mathrm{Tr}[e^{-\beta H}]$ at inverse temperature β is isomorphic to the classical partition function of an extended system—the so-called *ring polymer*,

$$Z_P \propto \int \mathrm{d}\mathbf{p}\, \mathrm{d}\mathbf{q}\, e^{-\frac{\beta}{P} H_P(\mathbf{p},\mathbf{q})}. \tag{13}$$

Here H_P is the ring polymer Hamiltonian

$$H_P(\mathbf{p}, \mathbf{q}) = \sum_{i=0}^{P-1} \frac{p_i^2}{2m} + V(q_i) + \frac{1}{2} m\omega_P^2 (q_i - q_{i+1})^2, \tag{14}$$

where cyclic boundary conditions $i + P \equiv i$ are implied, and $\omega_P = P/\beta\hbar$. Equation (14) describes P replicas (beads) of the physical system, where replicas with adjacent indices are connected by harmonic springs to form a closed loop. As $P \to \infty$ the path integral partition function Z_P converges to the correct quantum mechanical one.

The momenta in the partition function of Eq. (13) simply provide a sampling tool and can be integrated out analytically. Hence all the quantum mechanical information is encoded in the configurational part of Z_P. Therefore, even a property which is directly related to the momentum such as the kinetic energy must be computed by means of a configurational average of an appropriate estimator. In what follows we will use $\langle O \rangle_m$ to indicate the configurational average of the estimator $O(\mathbf{q})$ over the canonical ensemble for a particle of mass m,

$$\langle O \rangle_m = \frac{\int \mathrm{d}\mathbf{q}\, O(\mathbf{q})\, e^{-\frac{\beta}{P}\sum_i V(q_i) + m\omega_P^2 (q_i - q_{i+1})^2/2}}{\int \mathrm{d}\mathbf{q}\, e^{-\frac{\beta}{P}\sum_i V(q_i) + m\omega_P^2 (q_i - q_{i+1})^2/2}}. \tag{15}$$

This phase-space average can be evaluated by generating a sequence of configurations consistent with the appropriate probability distribution, and averaging the values of O corresponding to the different snapshots. This set of configurations can be generated for instance by means of a Monte Carlo procedure, or by molecular dynamics.[17,18,42,43]

By differentiating the partition function with respect to β, one can derive the thermodynamic kinetic energy estimator,

$$\langle T_{\mathrm{TD}}(m) \rangle_m = \left\langle \frac{P}{2\beta} - \frac{1}{2P} m\omega_P^2 \sum_{i=0}^{P-1} (q_i - q_{i+1})^2 \right\rangle_m. \tag{16}$$

In practice however computing the kinetic energy from Eq. (16) is problematic, because the estimator $T_{TD}(m)$ exhibits a large variance that worryingly grows linearly with the number of beads, P. Hence the statistical error of a simulation of a given length grows as the number of beads is increased to reach convergence. It is however possible to derive a centroid-virial estimator,[42,44]

$$\langle T_{CV} \rangle_m = \left\langle \frac{1}{2\beta} + \frac{1}{2P} \sum_{i=0}^{P-1} (q_i - \bar{q}) \frac{\partial V}{\partial q_i} \right\rangle_m, \tag{17}$$

where $\bar{q} = \sum_i q_i / P$ is the centroid coordinate. This estimator yields the same average value as that in Eq. (16) but has a variance that does not grow with P.

D. Isotope substitution by free energy perturbation

As shown in Eq. (4) in order to evaluate the free energy change induced by isotope substitution, one has to calculate the change in the quantum kinetic energy upon the change in mass of one of the atoms in the system. The most direct approach is to perform a series of independent simulations with different values of the isotope mass and compute $\langle T_{CV} \rangle$ for each. Since in most cases one of the isotopes has very low natural abundance, this typically involves performing a simulation with a single particle of mass μ inside a large supercell containing several tens or hundreds of atoms of the naturally abundant isotope. This is far from ideal, because one has to perform multiple simulations, one for each value of μ, in order to evaluate the thermodynamic integration in Eq. (4), and because statistics are accumulated for just one atom out of the many present in the simulation. Hence a highly appealing solution would be to compute $\langle T_{CV} \rangle_\mu$ from a single simulation of *only* the most abundant isotope by a free energy perturbation (FEP) method. This approach involves evaluation of estimators that simultaneously compute the kinetic energy and correct the phase-space distribution to reproduce the statistics of the isotope-substituted system.

These estimators can be obtained by writing the full expressions for the expectation values $\langle T_{CV} \rangle_\mu$ and $\langle T_{CV} \rangle_m$ as phase-space averages. By comparing the two expressions, one sees that it is possible to write a "thermodynamic" FEP estimator for the kinetic energy at mass μ as the ratio between two phase-space averages computed at mass m,

$$\langle T_{CV} \rangle_\mu^{\mathrm{TD}} = \frac{\langle T_{CV}(\mathbf{q}) \exp[-h_{\mathrm{TD}}(\mu/m; \mathbf{q})] \rangle_m}{\langle \exp[-h_{\mathrm{TD}}(\mu/m; \mathbf{q})] \rangle_m}, \tag{18}$$

where we introduce the temperature-scaled difference Hamiltonian

$$h_{\mathrm{TD}}(\alpha; \mathbf{q}) = \frac{(\alpha - 1)\, \beta m \omega_P^2}{2P} \sum_{i=0}^{P-1} (q_i - q_{i+1})^2 . \tag{19}$$

These equations provide a recipe to compute the quantum kinetic energy for an arbitrary mass ratio $\alpha = \mu/m$ out of simulations performed for a single isotope mass. However, statistical re-weighting comes with its own set of problems which must be carefully scrutinized. The difference Hamiltonian Eq. (19) appears in the exponent of Eq. (18). If it exhibits large fluctuations the weights of different configurations will vary wildly and the average performed over many configurations will be dominated by a few outliers, resulting in very poor sampling efficiency. More quantitatively, recent work has shown that whenever one computes re-weighted averages of the form $\langle a e^{-h} \rangle / \langle e^{-h} \rangle$, the squared error in the mean for an average built out of M un-correlated samples will be of the order of $\sigma^2(a) e^{\sigma^2(h)} / M$.[45] Hence the error grows *exponentially* with the variance of the difference Hamiltonian, $\sigma^2(h)$. In addition re-weighted averages are affected by a slowly-decreasing systematic error. For re-weighted sampling to be efficient it is therefore necessary for the fluctuations of the temperature-scaled difference Hamiltonian to be small and certainly not much larger than one.

With this in mind Eq. (19) is worrisome in that it is closely reminiscent of the thermodynamic kinetic energy estimator in Eq. (16) which is known to exhibit a variance which grows with the number of beads, P. Due to this similarity we will refer to the estimator in Eq. (16) as the TD-FEP estimator. One may therefore wonder whether it is possible to derive an alternative expression which, just as the centroid-virial kinetic energy estimator in Eq. (17), yields the same average with a variance which does not grow with P. To this aim, we use a coordinate-scaling approach in the form of Yamamoto.[46] In our case, the appropriate scaled coordinates are

$$q_i'(\alpha) = \bar{q} + \frac{1}{\sqrt{\alpha}} (q_i - \bar{q}) . \tag{20}$$

This scaling corresponds to a contraction/dilation of the ring polymer. Starting from a configuration with a radius of gyration characteristic of the abundant isotope, it produces one that is more compatible with what would be expected for an isotope of mass μ. By performing such a change of variables in the configurational integral for $\langle T_{CV} \rangle_\mu$, and then using manipulations analogous to those in Ref. 46, we obtain

$$\langle T_{CV} \rangle_\mu^{\mathrm{SC}} = \frac{\langle T_{CV}(\mathbf{q}'(\mu/m)) \exp[-h_{\mathrm{SC}}(\mu/m; \mathbf{q})] \rangle_m}{\langle \exp[-h_{\mathrm{SC}}(\mu/m; \mathbf{q})] \rangle_m}. \tag{21}$$

Again, the average is performed in the ensemble that contains only the majority isotope, but now the centroid-virial estimator is evaluated at the scaled coordinates (Eq. (20)) and the temperature-scaled difference Hamiltonian is

$$h_{\mathrm{SC}}(\alpha; \mathbf{q}) = \frac{\beta}{P} \sum_{i=0}^{P-1} V(q_i'(\alpha)) - V(q_i) . \tag{22}$$

Unlike the TD-FEP estimator in Eq. (19), it can be seen that the variance of h_{SC} does not grow with P. Due to the scaled coordinate transformation we will refer to this estimator as the SC-FEP estimator. However, unlike the TD-FEP estimator which can be computed for an arbitrary mass ratio in a completely inexpensive way, the SC-FEP estimator requires an additional P energy and force evaluations for each mass ratio and tagged particle. Such force evaluations can be made inexpensive if the potential consists only of a sum of pairwise or otherwise local contributions, because in that case only the interactions involving the tagged particle need to be updated. However, in other cases using the SC-FEP estimator may entail a significant computational overhead due to the extra potential energy calculations.

Because of the subtle issues connected with re-weighting methods, in order to assess the relative merits of these two FEP estimators it is important to perform a careful quantitative analysis. In the Appendix we examine the case of a harmonic potential, where such analysis can be performed analytically. Let us consider here the resulting expressions for the fluctuations of the difference Hamiltonian, under the further simplifying assumptions that the large-P limit is being

attained and that we are dealing with a strongly quantized normal mode of frequency $\omega_{\max}$ (i.e., $P > \beta\hbar\omega_{\max} \gg 1$). For the TD-FEP estimator we find that $\sigma^2(h_{\mathrm{TD}})$ is proportional to P, as one might have expected,

$$\sigma^2(h_{\mathrm{TD}}) \sim \frac{(\alpha-1)^2}{2}\left(P - \frac{3}{4}\beta\hbar\omega_{\max}\right). \tag{23}$$

In contrast, the difference Hamiltonian for the SC-FEP estimator yields a variance which does not grow with P,

$$\sigma^2(h_{\mathrm{SC}}) \sim \frac{(\alpha^{-1}-1)^2}{2}\left(\frac{\beta\hbar\omega_{\max}}{4} - 1\right). \tag{24}$$

As discussed earlier, re-weighting approaches are only reliable when $\sigma^2(h) \lesssim 1$.[45] This imposes a limit on the range of values of the mass ratio, α, for which the TD-FEP estimator can be used. The number of beads required to converge a path integral simulation depends on the maximum frequency in the system and a reasonable convergence criterion is $P = 2\beta\hbar\omega_{\max}$.[22] Inserting this condition into (23) yields a constraint on the values of the mass ratio for which the TD-FEP estimator is reliable,

$$1 - \frac{\sqrt{8/5}}{\sqrt{\beta\hbar\omega_{\max}}} \leq \alpha_{TD} \leq 1 + \frac{\sqrt{8/5}}{\sqrt{\beta\hbar\omega_{\max}}}. \tag{25}$$

Proceeding similarly for the SC-FEP estimator yields a range of acceptable mass ratios of

$$\frac{1}{1+\sqrt{8/(\beta\hbar\omega_{\max}-4)}} \leq \alpha_{SC} \leq \frac{1}{1-\sqrt{8/(\beta\hbar\omega_{\max}-4)}}. \tag{26}$$

These analytic estimates allow us to predict how the two approaches might compare in a practical case. For example for liquid water at room temperature the highest frequency is the OH stretch which extends to frequencies around 3500 cm^{-1}. This gives $\beta\hbar\omega_{\max} \approx 16$. Inserting this into Eqs. (25) and (26) yields acceptable mass ratios of

$$0.7 < \alpha_{\mathrm{TD}} < 1.3 \tag{27}$$

and

$$0.5 < \alpha_{\mathrm{SC}} < 5.5 \tag{28}$$

for the TD-FEP and SC-FEP estimators, respectively. Hence the TD-FEP approach in Eq. (18) would not even allow one to perform a FEP that transforms ^1H into deuterium ($\alpha = 2$), whereas the SC-FEP approach in Eq. (21) would be reliable all the way to tritium ($\alpha = 3$). In Sec. III we assess the accuracy of these harmonic predictions of the applicability of the FEP estimators and compare them with the direct substitution method for the calculation of free energy changes in liquid water.

III. RESULTS

To assess the relative merits of the approaches introduced in Sec. II D we performed isotope substitution simulations in liquid water. Despite its apparent simplicity liquid water provides a challenging test-bed since it contains a wide range of frequencies and exhibits a competition between two quantum effects, one of which acts to strengthen its hydrogen bonding network and one which weakens it.[8,9] This competition is seen in a variety of other hydrogen bonded systems[10] and arises from the anharmonicity along the hydrogen bond-covalent bond coordinate.[9,11] As such water contains many of the features one may encounter in other liquids or hydrogen bonded systems and has an important but subtle anharmonic component which must be correctly captured in order to accurately compute the free energy change upon isotope substitution.

The free energy change we will focus on is the isotopic fractionation of H and D between the liquid and the gas phase for water at room temperature defined in Eq. (11) as well as the fractionation of ^{16}O and ^{18}O. The results are quoted as fractionation factors as defined in Eq. (12) which correspond to the ratio of isotope in the liquid to that in the vapor. The fractionation of isotopes is a frequently used technique to understand processes as diverse as the water cycle in the earth's climate, biological function and the origins of the interstellar medium.[47–49]

As can be seen from Eqs. (4) and (11) the liquid-vapor fractionation ratio probes the change in the quantum kinetic energy upon moving a water molecule from the liquid phase to the gas phase. An increase in the quantum kinetic energy occurs when a particle is confined by the forces exerted on it by the surrounding molecules. Computing fractionation ratios is particularly challenging, since their precise value depends on subtle cancellations of different quantum effects. As such, they could serve as a sensitive test of how well different potentials reproduce this delicate balance.

We performed our simulations using the flexible simple point charge q-TIP4P/F water model.[9] This model was chosen as it has been shown to accurately predict the H/D fractionation between liquid water and its vapor[11] as well as a variety of other properties of water such as its density maximum and melting point[9] when used in PIMD simulations.

We will compare three different approaches for computing the free energy change upon isotope substitution: (1) the previously introduced direct substitution approach; (2) the TD-FEP estimator in Eq. (18); and (3) the SC-FEP estimator in Eq. (21).

A. Generating path integral configurations

Evaluating the expectation values of the kinetic energy in Eq. (17) or the re-weighted kinetic energies in Eqs. (18) and (21) requires the generation of path integral trajectories. We will compare two approaches: an efficiently thermostatted implementation of PIMD[21] and the recently introduced PIGLET method.[26]

The difference in these approaches is in their convergence with respect to the number of beads employed. In conventional PIMD[18] in order to converge a simulation for a system whose fastest vibrational mode has a frequency $\omega_{\max}$, the number of beads, P, must be two or more times the

parameter $\beta\hbar\omega_{\max}$. For water at room temperature typically $P = 32$ yields results with an error of a few percent. In most simulations the computational effort is dominated by the calculation of the forces, and so the total cost of a conventional PIMD simulation scales linearly with beads. Hence a PIMD simulation of liquid water is ∼32 times more expensive than a classical one. The PIMD simulations were performed using the PILE-G thermostatting scheme[21] that has previously been demonstrated to be an efficient approach to simulate liquids. Since the global thermostat minimally hinders diffusion[50] we used a short thermostat correlation time of $\tau_{\text{thermo}} = 25$ fs.

Recently introduced colored-noise Langevin dynamics approaches can reduce the cost of path integral simulations by enforcing certain pre-determined bead-bead correlations to be exact in the harmonic limit.[26] This dramatically reduces the number of beads needed to converge many properties, even in anharmonic systems. Of these colored-noise approaches the PIGLET method[27] is particularly relevant to the present work since it was designed to ensure fast convergence of the quantum kinetic energy, $\langle T_{CV}\rangle_\mu$ which is all that is required to compute the free energy change by explicit isotope substitution. The fast convergence of the kinetic energy using PIGLET is demonstrated in Figure 3(a). Using PIMD with $P = 32$ yields an estimate of the kinetic energy with an error of about 2%. PIGLET yields results with similar accuracy to this using only 6 beads. Figure 3(b) shows the fractionation ratio evaluated from Eq. (12). This property is the small difference of two large quantities and is therefore very sensitive to errors in the description of anharmonicities, which is apparent in the very poor performance of the quasi-harmonic approximation (see Sec. II B). Nevertheless, PIGLET performs remarkably well: from 4 to 12 beads, the values of $\ln\alpha_{l-v}$ differ by less than 10% from the asymptotic value, an error which is much lower than the one stemming from the use of an empirical potential.[11]

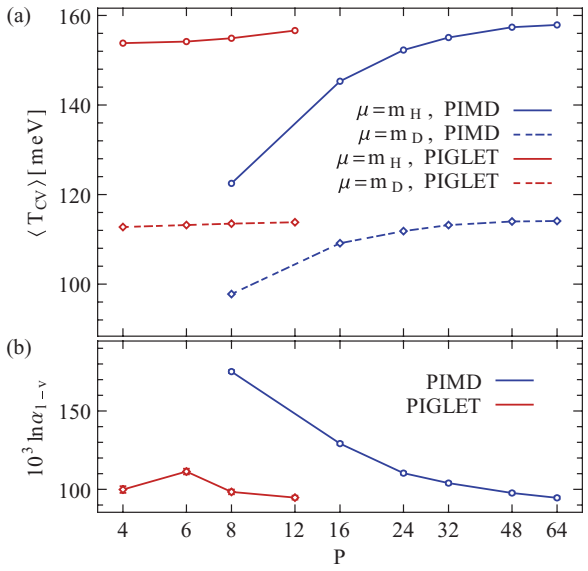

FIG. 3. (a) Expectation values for the centroid-virial kinetic energy estimator for hydrogen in liquid water as a function of the number of beads. Results are shown for both PIMD and PIGLET simulations, and for the tagged particle having mass set to that of hydrogen (m_H) or deuterium (m_D). (b) Convergence of the isotope fractionation ratio α_{l-v} between the liquid and the gas phase of water for the H/D substitution. Results are reported as a function of the number of beads, using conventional path integral MD (blue circles) and PIGLET (red lozenges). Lines are just guides for the eye, and statistical error bars are smaller than the size of the points.

However, even in the harmonic limit, the PIGLET approach does not enforce all the bead-bead correlations needed for the expectation values of the TD-FEP and SC-FEP estimators in Eqs. (18) and (21). Therefore, there is no guarantee that PIGLET will converge much faster than conventional PIMD when used together with one of the FEP estimators. On the other hand, the earlier PI+GLE colored noise approach[26] was shown to enhance convergence of the kinetic energy even though it was not developed to enforce all the required correlations. This suggests that PIGLET, even in its current form, might be applied to calculate the FEP estimators as we investigate in Secs. III C and III D.

B. Direct computation of $\langle T_{CV}\rangle_\mu$

The direct technique to compute isotope effects involves evaluating the kinetic energy $\langle T_{CV}\rangle_\mu$ in multiple simulations where the mass of one of the atoms is changed between those of the two isotopes in order to perform the thermodynamic integration in Eq. (4). As demonstrated in Sec. II A, with a physically motivated change of integration variable it is possible to accurately evaluate the integral in Eq. (7) by computing just the extremal points. Despite the need to replicate the system this can be much cheaper than simply using a quasi-harmonic approximation from a classical trajectory.

One reason for this is the efficiency in de-correlating the kinetic energy in a path integral simulation. The statistical uncertainty of a property computed out of a simulation of length t_{sim} decreases in proportion to $\sqrt{\tau/t_{\text{sim}}}$ where τ is the correlation time of that property. Hence, for a given length of the simulation, observables with short correlation times will exhibit smaller statistical error. In this regard the PILE-G thermostat we used for PIMD calculations and PIGLET are both very efficient. Our simulations show the autocorrelation time of the centroid-virial kinetic energy estimator to depend weakly on P and μ, varying between 1.5 and 2 fs for all the simulations we performed. We used a time step of 0.25 fs and hence 6–8 force evaluations are enough to obtain a configuration where the kinetic energy estimator is uncorrelated. In addition, because of the small number of beads required for convergence, PIGLET simulations could have been performed with a 0.5 fs time step, which enhances the efficiency even further. This explains why despite the additional cost of path integral approaches in terms of force evaluations they are still cheaper than using the quasi-harmonic approach, as we demonstrated in Sec. II B. In our case we used an inexpensive empirical potential, and therefore we ran simulations that were as long as 500 ps, resulting in minuscule statistical errors. If one were prepared to accept a ±5% uncertainty on the H/D fractionation ratio, liquid water simulations of the order of 10 ps starting from an equilibrated configuration would suffice, which makes performing *ab initio* fractionation calculations for liquids highly plausible.

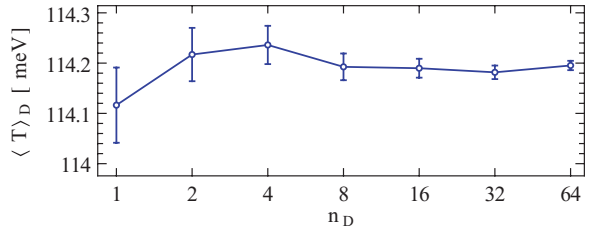

FIG. 4. Average kinetic energy (blue circles) computed from simulations in which n_D hydrogen atoms have been substituted with deuterium, in a simulation box containing 64 water molecules. At most one atom per molecule was substituted.

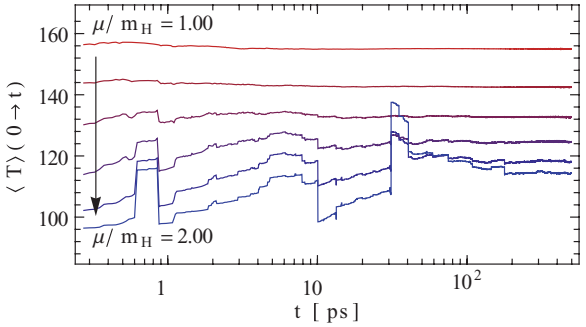

FIG. 5. Cumulative average of $\langle T_{CV}\rangle_\mu^{\mathrm{TD}}$, as computed for hydrogen isotopes in a single PIMD trajectory of liquid water, with $P = 32$. Lines, from top to bottom, correspond to different mass ratios ranging from $\mu/m_H = 1$ to $\mu/m_H = 2$.

Since minority isotopes have very low natural abundance, fractionation ratios are typically required in the low-dilution limit. In order to simulate this limit exactly the mass has to be changed for a single atom in a system that contains hundreds of instances of the naturally-abundant isotope, and the kinetic energy can be computed for that particle only. However, in a homogeneous system such as water it is interesting to explore how sensitive the kinetic energy, and hence the computed fractionation ratio, is to multiple isotope substitutions in the system. The advantage of making these multiple isotope substitutions is that one can then gather statistics from several atoms at the same time. Momentum-dependent observables tend to be very local in nature, and one can hope that the error introduced by having an unnaturally high concentration of isotope substituted species will be small. For instance, when performing open-path PIMD in liquid water, one can simultaneously open the path for one hydrogen atom per water molecule without introducing significant systematic errors.[51,52] Figure 4 shows how $\langle T_{CV}\rangle_{m_D}$ changes when multiple atoms in a simulation of light water are transmuted into deuterium, with the precaution of never substituting two atoms on the same molecule. Even when half the H's in the system are converted, the kinetic energy of the D's changes by less than 0.3%, a difference that is not statistically significant even with these relatively long calculations. When performing more demanding, and therefore shorter, *ab initio* simulations, the statistical error will almost certainly be dominant, making it particularly attractive to gather horizontal statistics by multiple substitutions. One can also observe that when multiple substitutions are performed the statistical uncertainty of $\langle T_{CV}\rangle_{\mu=m_D}$ decreases with the square root of the number of substituted atoms. This implies that the instantaneous values of T_{CV} for atoms in different molecules are almost perfectly un-correlated—another indication of the local nature of the quantum kinetic energy. However, we stress that one should be careful when performing multiple substitutions, in particular for inhomogeneous systems, and that whenever possible one should benchmark this approximation against the highly-diluted case.

Even exploiting our efficient transformation of the integral which requires only simulations at the two extrema the need to perform multiple simulations to obtain free energy changes using direct substitution is undesirable. We now consider our FEP approaches which allow the free energy change to be obtained from a single trajectory of the most abundant isotope.

C. "Thermodynamic" free energy perturbation

The TD-FEP estimator in Eq. (18) provides an appealing approach to the computation of the kinetic energy of an isotope-substituted system by just post-processing the data from a simulation that contains only the naturally-abundant isotope. For a homogenous system such as water this "virtual substitution" can be performed one atom at a time for every atom in the system allowing for full horizontal statistics to be collected. However, as discussed in Sec. II the re-weighting procedure which underlies free-energy perturbation methods is potentially dangerous, as it can introduce systematic errors and a difficult-to-detect statistical inefficiency.[45] The presence of unusual statistical behavior, dominated by outliers, is demonstrated in Fig. 5, which shows the cumulative average of the kinetic energy predicted from Eq. (18) for different mass ratios as the simulation progresses. For large values of μ/m_H the cumulative average does not converge with regularity to a mean value, but exhibits a characteristic sequence of plateaus, interrupted by abrupt jumps when a new outlier is encountered.

A more quantitative approach to monitor the statistical quality of the TD-FEP estimates of $\langle T_{CV}\rangle_\mu^{\mathrm{TD}}$ involves computing the variance of the temperature scaled difference Hamiltonian in Eq. (19). As discussed in Sec. II, a value of $\sigma^2(h)$, smaller than one is required for re-weighting to be of practical use. Figure 6 shows the variance obtained from our water simulations for different combinations of the number of beads, P, and mass ratio, α. As expected from our analysis in the Appendix the variance of the TD-FEP difference Hamiltonian grows rapidly with both P and α so in practice this estimator becomes useless for mass ratios greater than about 1.2. Importantly the simulated behavior of the variance is in excellent agreement with the analytic prediction derived for the harmonic limit in Eq. (27) which predicts a maximum mass ratio of 1.3 for water simply from the knowledge of the highest frequency present. Hence, the analytic estimates provide an excellent *a priori* way of deciding whether one should employ TD-FEP for a given system.

A maximum mass ratio of 1.2 using conventional PIMD simulations implies that TD-FEP is not recommended for studying isotope substitution for hydrogen in room temperature water but this mass ratio is within the range needed to

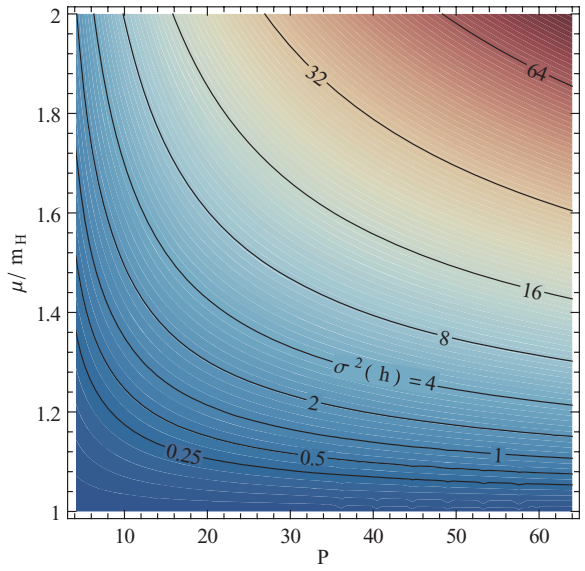

FIG. 6. Contour plot of the variance of the difference Hamiltonian for the thermodynamic free energy perturbation h_{TD}, as a function of the number of beads P and the mass ratio $\alpha = \mu/m_H$. For $\sigma^2(h_{\mathrm{TD}}) \gtrsim 1$ computing $\langle T_{CV}\rangle_\mu^{\mathrm{TD}}$ becomes impractical.

perform ^{16}O/^{18}O substitution, which involves a mass ratio of only $\alpha = 18/16$. Figure 7 shows the isotope fractionation ratio for ^{16}O and ^{18}O between the liquid and the gas phases of water at ambient conditions. The isotope substitution free energies were computed by TD-FEP from a single simulation of pure ^{16}O water. Despite being a heavy atom the number of beads required to converge the kinetic energy of O in the liquid and gas phase is ∼32 since it is bonded to a light atom. The fractionation ratio converges with the number of beads more rapidly than the kinetic energy estimator does, because of a cancellation of errors between the liquid and the gas phases. We find $10^3 \ln \alpha_{l-v} = 12.5 \pm 0.5$, for the q-TIP4P/F

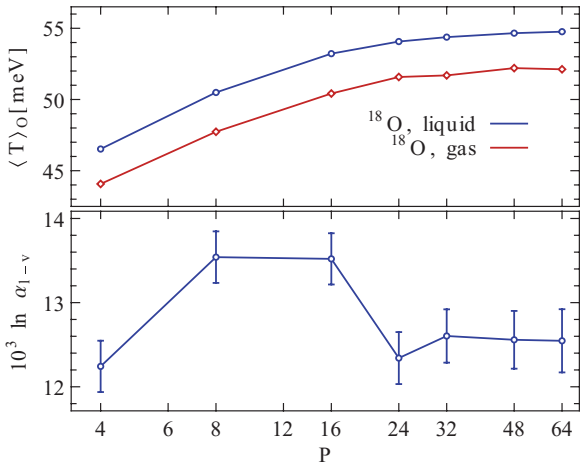

FIG. 7. Convergence of the isotope fractionation ratio α_{l-v} between the liquid and the gas phases of water at ambient conditions for the ^{16}O/^{18}O substitution. Upper panel depicts the convergence of the ^{18}O kinetic energy in the two phases, as computed from thermodynamic free energy perturbation, while the lower panel reports the fractionation ratio. Lines are just guides for the eye.

model which is in reasonable agreement with the experimental value of 9.2.[53] This is in error by an amount similar to a fall of 30 K in the temperature which is roughly the same as the error in the H/D liquid vapor fractionation when compared to experiment[11] suggesting this model mildly overpredicts fractionation in water.

As we discussed in Sec. III A PIGLET does not enforce the bead correlations necessary to ensure that it will accelerate the convergence of the TD-FEP estimator even in the harmonic limit. However, there are two reasons why such a combination could be highly beneficial and hence is worth investigating. The first is simply the reduction in the number of beads, which implies that fewer force evaluations are required to evolve the system resulting in a speed-up of ∼6 times in water. The second is that the variance of the difference Hamiltonian that exponentially determines the rise in the error of TD-FEP depends on the number of beads (see Fig. 6). Hence any reduction in the number of beads by using PIGLET would offer exponential benefits in the efficiency of the TD-FEP approach.

Figure 8(a) shows the value of the kinetic energy of $\langle T_{CV}\rangle_{m_D}$ estimated from a single simulation of pure liquid H_2O, using TD-FEP with PIGLET. Without PIGLET, at least 32 beads would be needed to converge the kinetic energy, which as discussed above would restrict the maximum acceptable mass ratios to about 1.2. On the other hand, as shown in Fig. 8(b), PIGLET reaches convergence with about 6 beads for both the kinetic energy and the fractionation ratio, which is just on the boundary of the region for which performing TD-FEP up to $\alpha = 2$ would become impractical. The fact that PIGLET was not parameterized to enforce the particular correlation needed for TD-FEP does not appear to limit its utility in converging it and indeed remarkably it even seems

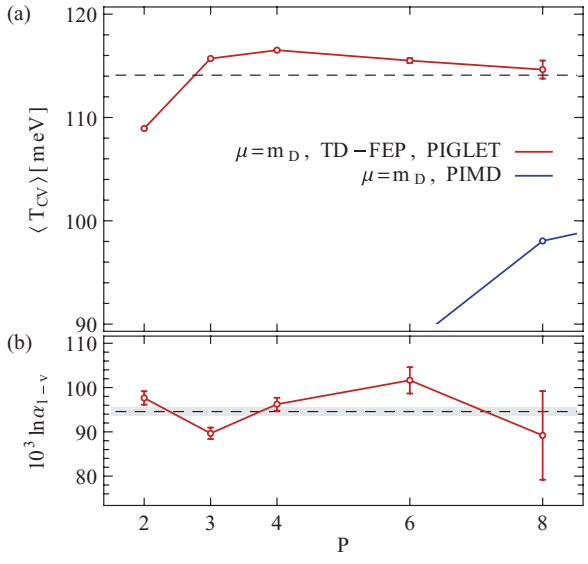

FIG. 8. The figure demonstrates the performance of the TD-FEP estimator used together with PIGLET. Panel (a) represents $\langle T_{CV}\rangle_{m_D}$ as a function of the number of beads. PIMD results are shown for comparison, and the black line indicates the reference value (direct substitution, PIMD, $P = 64$). In Panel (b) the isotope fractionation ratio between liquid and vapor is shown, for water at room temperature, as estimated from TD-FEP and PIGLET. The black line indicates the reference value (direct substitution, PIMD, $P = 64$).

to converge as rapidly as kinetic energy, which is rigorously enforced in the harmonic limit. Although by reducing the required number of beads PIGLET delays the failure of TD-FEP, beyond 6–8 beads the statistical error blows up rendering the estimator unusable (see Figure 6). Still, PIGLET makes it possible to use TD-FEP profitably over a broader range of mass ratios and temperatures. This is incredibly useful, since it allows H/D isotope substitution to be computed at room temperature and for ubiquitous aqueous and hydrogen-bonded systems, performing just a single trajectory and using as few as 6 beads.

D. Scaled coordinates free energy perturbation

In Sec. III C TD-FEP was shown to be limited due to possessing a statistical error that grows with the number of beads. Based on our harmonic analysis the SC-FEP approach should allow one to obtain a much wider range of mass ratios α from a single trajectory. In addition, the analytic predictions suggest that, unlike TD-FEP, the variance of the difference Hamiltonian for SC-FEP will show no dependence on the number of beads allowing the path integral simulation to be performed to high precision without the statistical inefficiency rising to unacceptable levels.

To test these predictions, Fig. 9 shows the variance of the difference Hamiltonian h_{SC} for SC-FEP as a function of P and μ/m_H, over the same range of masses and numbers of beads we explored in Figure 6 for TD-FEP. In striking contrast with the case of the TD-FEP, $\sigma^2(h_{\mathrm{SC}})$ is nearly independent of P, and even up to a mass ratio of $\alpha = 2$ is well below the threshold value of one. The SC-FEP estimator is therefore suitable for computing H/D substitution using conventional PIMD using either $P = 32$ or 64 beads depending on the accuracy required. Based on the scaling of the variance with temperature

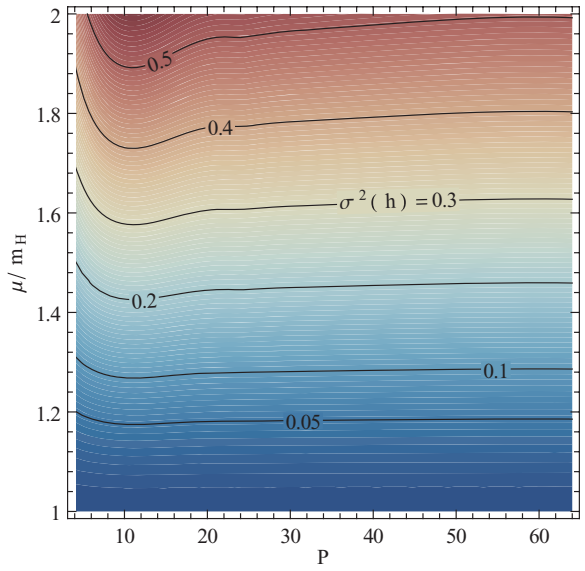

FIG. 9. Contour plot of the variance of the difference Hamiltonian for the coordinate-scaling free energy perturbation h_{SC}, as a function of the number of beads P and the mass ratio $\alpha = \mu/m_H$. Throughout the range we considered, $\sigma^2(h_{\mathrm{SC}}) < 1$.

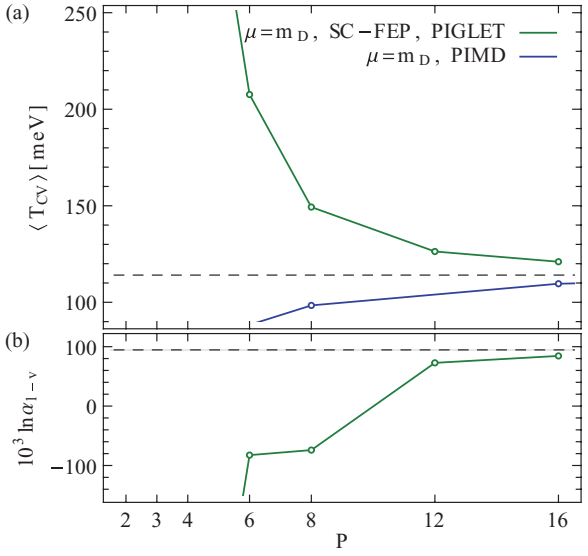

FIG. 10. The figure demonstrates the poor convergence of the SC-FEP estimator when used together with PIGLET. (a) $\langle T_{CV}\rangle_{m_D}$ as a function of the number of beads. PIMD results are shown for comparison, and the black line indicates the reference value (direct substitution, PIMD, $P = 64$). (b) Isotope fractionation ratio between liquid and vapor for water at room temperature, estimated from SC-FEP and PIGLET. The black line indicates the reference value (direct substitution, PIMD, $P = 64$).

predicted by Eq. (26) one would expect this to be the case even at temperatures considerably lower than 300 K.

Given SC-FEP's much greater range of applicability than TD-FEP when used in conjunction with conventional PIMD one might wonder if SC-FEP could be reduced in cost by using PIGLET which was shown to be highly effective in accelerating convergence of TD-FEP. Figure 10 demonstrates that in this case the bead-bead correlations needed to converge the SC-FEP estimator are clearly not included in the current PIGLET parameterization leading to very poor convergence. When using PIGLET with 6 beads, which is essentially converged for TD-FEP or direct substitution, the SC-FEP estimator gives a kinetic energy estimate for deuterium with an error of 100% and a fractionation ratio which does not even possess the correct sign. It is reassuring to note that, by further increasing the number of beads, convergence will be eventually reached. This underlines the robustness of the PIGLET approach: when a sufficiently large number of beads are used, the non-equilibrium nature of the colored noise naturally vanishes and these methods reduce to conventional PIMD. It is also worth mentioning that in principle one could exploit the philosophy of the PIGLET method to enforce in the harmonic limit the bead-bead correlations that are necessary to rigorously converge both TD-FEP and SC-FEP. Such a reparameterization of the colored noise is a non-trivial task and hence is left for future work.

E. Recommendations

From the discussion above it is clear that each approach for obtaining the free energy change upon isotope exchange has benefits and deficits. As such there can be no single recommendation but rather, based on our simulations and ana-

lytic results, one can suggest the most appropriate approach to efficiently tackle typical simulation scenarios.

The direct substitution approach provides a baseline since it has been extensively used in previous studies.[11,32–34] In the present paper we demonstrate a number of strategies that greatly improve its efficiency. For a start, by performing a physically motivated smoothing of the integral (Sec. II A) one can usually perform the thermodynamic integration with only two simulations for the extremal values of the isotope mass. In addition, the PIGLET approach can be used to reduce the number of force evaluations required for each simulation. Finally, our results suggest that in a homogeneous system performing multiple substitutions is a reliable way to obtain horizontal statistics—although this should of course be treated with caution and tested carefully before it is applied to other problems. With these improvements, direct substitution becomes a very competitive choice even in comparison to approximate methods.

The free energy perturbation approaches we introduce, however, are even more appealing, in that they allow one to obtain the isotope substitution free energy from a single simulation containing only the majority isotope, and allow one to gather horizontal statistics in an exact way. However, one must be careful of the statistical efficiency of re-weighting whenever large fluctuations occur in the difference Hamiltonian. In Secs. I–III we have presented both analytical and simulation evidence to determine when FEP approaches are effective. This allows us to provide some simple practical guidelines:

- For heavy isotope substitution at temperatures above 100 K, one should perform a PIGLET simulation containing only the most abundant isotope and use the TD-FEP estimator to extract the free energy change.
- For H/D substitution in "biological" conditions (room temperature, aqueous systems) and at higher temperatures, one should also use TD-FEP and PIGLET. However, this more challenging regime sits on the edge of the TD-FEP estimator's applicability when combined with PIGLET. Hence TD-FEP should not be used with PIMD in this regime since the additional beads required for convergence cause the statistical efficiency to become unacceptable.
- For isotope substitution below room temperature for H/D or at room temperature when making a substitution with a larger mass ratios than H/D (e.g., Muonium/H or H/Tritium) a direct substitution calculation is the best option. The efficiency of direct substitution can be greatly enhanced by using our transformed integration variable, PIGLET and, where appropriate, using multiple substitutions.

While the three cases above cover most practical applications, there are circumstances in which the SC-FEP estimator becomes competitive. Due to the requirement to evaluate the potential on a scaled ring polymer its main advantage lies in cases where the potential can be evaluated efficiently upon a single particle change such as for empirical potentials. As we demonstrated the convergence of the SC-FEP is not improved by PIGLET and so PIMD must be used to simulate the trajectory. However, for empirical potentials PIMD trajectories can be made extremely cheap by using ring-polymer contraction techniques.[22–24] SC-FEP is also more attractive in problems where only a single atom in the system is substituted such as for a particular proton in an enzyme active site.

Since our analytic predictions of the fluctuations in the difference Hamiltonian for the FEP estimators were seen to be reliable when compared to liquid water simulations we can use them to derive simple expressions that predict the efficiency of TD-FEP and SC-FEP relative to direct simulation. These efficiency measures, E_{TD} and E_{SC} gauge the cost of obtaining the free energy change from the TD-FEP and SC-FEP estimators, respectively, compared to obtaining it from direct substitution. If E_{TD} or E_{SC} comes out much greater than one then the FEP approach is preferable and the FEP variant with the highest value should be used. If both values come out much less than one then direct substitution should be employed. Values around one suggest that either FEP or direct substitution is of comparable efficiency for the problem. For an isotope substitution with a mass ratio $\alpha = m'/m$ between the most and least abundant isotope at inverse temperature β in a system with a maximum vibrational frequency ω_{max} one can show that for the TD-FEP estimator the efficiency is

$$E_{\mathrm{TD}} = n_{\mathrm{int}}\sqrt{\frac{N_X}{M_X}}\exp\left[-\frac{(\alpha-1)^2}{4}\left(P-\frac{3}{4}\beta\hbar\omega_{\mathrm{max}}\right)\right]. \tag{29}$$

Here n_{int} is the number of integration points one needs to converge the thermodynamic integration (2 or 3 when using Eq. (7)), N_X is the number of equivalent atoms for which the substitution can be performed, and M_X the number of multiple substitutions one can perform in a direct calculation ($M_X = 1$ is the only completely safe bet). P is the number of beads required for convergence, which is of the order of $2\beta\hbar\omega_{\mathrm{max}}$ with conventional PIMD and typically about $\beta\hbar\omega_{\mathrm{max}}/3$ with PIGLET. From this equation one can immediately see that for systems which require large numbers of integration points, n_{int} or posses a large number of exchangeable atoms then TD-FEP holds an advantage. However, at low temperature (high β) or a large mass ratio or high frequencies ω_{max}, the efficiency can quickly shift in favor of direct substitution. One can also immediately see why PIGLET is so useful when used with TD-FEP since the relative efficiency is exponentially sensitive to the number of beads used.

The efficiency of the SC-FEP estimator is

$$E_{\mathrm{SC}} = n_{\mathrm{int}}\sqrt{\frac{\tau N_X P_{\mathrm{GLE}}}{\Delta t(1+KN_X)M_X P}} \times \exp\left[-\frac{\left(\alpha^{-1}-1\right)^2}{4}\left(\frac{\beta\hbar\omega_{\mathrm{max}}}{4}-1\right)\right]. \tag{30}$$

Besides the quantities present in Eq. (29), this expression also contains τ, the correlation time of T_{CV}, Δt, the time step of the simulation when using PIGLET, and K, the cost of computing the scaled-coordinates potential relative to the cost of a simulation step (e.g., $K = 1$ for *ab initio*, and virtually $K = 0$ if the potential can be evaluated locally). Since SC-FEP

cannot be used together with PIGLET, the expression also contains both the number of beads required to converge PIGLET, P_{GLE} and the number of beads required to converge conventional PIMD, P. Here it can be seen that for systems with long correlation times τ, cheap scaled-coordinate potential evaluations K or low numbers of exchangeable sites N_X, SC-FEP dominates.

IV. CONCLUSIONS

In this paper we have introduced a series of improvements which enhance the ease and efficiency of calculating isotope exchange free energies. This includes two new estimators, based on free energy perturbation techniques as well as a number of improvements to the existing direct substitution approach. We then critically discussed the advantages and shortcomings of each method predicating our reasoning on both analytic arguments and on extensive simulations of room-temperature water.

The balance of pros and cons is fundamentally related to the statistical efficiency of sampling of the different techniques. This, in turn, depends on the "quantumness" of the system, as expressed by the ratio between the highest quantum of vibrational energy and the thermal energy, $\beta\hbar\omega_{\max}$ and on the ratio of the masses of the minority isotope and of the naturally abundant one.

When this mass ratio is close to one, as is the case whenever one is dealing with relatively heavy elements such as oxygen, a "thermodynamic" FEP estimator allows one to compute inexpensively the effect of isotope substitution based on a single simulation which only contains the majority isotope. This both simplifies and makes less computationally demanding the evaluation of quantities of great interest for geochemistry and climatology.

When instead the mass ratio is large, such as when one needs to substitute deuterium for hydrogen, or when the temperature is considerably below room temperature, the scenario is less clear-cut and one has to weight carefully a number of factors. The statistical efficiency of the "thermodynamic" estimator degrades dramatically with the number of beads, so it becomes mandatory to use PIGLET to ensure that convergence is achieved with a reduced number of beads. With this caveat, H/D substitution is still manageable with TD-FEP in most systems of interest at room temperature. At cryogenic temperatures an alternative estimator can be used, which is based on a scaling of coordinates relative to the centroid of the path and exhibits more robust statistical properties. However, it requires multiple evaluations of the forces, and is therefore more computationally demanding than TD-FEP. For simulations using empirical potentials it is possible to reduce this computational cost by exploiting the locality of the coordinate scaling procedure. For an *ab initio* simulation, however, there might be corner cases in which a carefully-performed direct substitution simulation is more advantageous that either TD-FEP or SC-FEP.

We provide analytical expressions that allow one to estimate *a priori* the different factors that determine the sampling efficiency, so that one can make the optimal choice for a given system without having to perform time-consuming benchmarks. This analysis and the estimators we introduce will greatly simplify the evaluation of observables connected with isotope substitution that can be measured experimentally with great precision and provide precious insight into a multitude of natural phenomena. Obtaining the corresponding theoretical predictions will be useful both to elucidate and predict the results of experiments, and to validate and benchmark empirical and *ab initio* predictions of condensed-phase systems.

ACKNOWLEDGMENTS

We would like to thank David Manolopoulos for insightful discussion and many precious suggestions and Timothy Berkelbach for a thorough reading of the manuscript. M.C. acknowledges funds from the EU Marie Curie IEF No. PIEFGA-2010-272402. T.E.M acknowledges funding from a Terman fellowship and Stanford start-up funds.

APPENDIX: REWEIGHTING STATISTICS FOR FEP ESTIMATORS

When developing new estimators in a path integral context, analyzing their properties in the harmonic limit often allows one to obtain analytic results and useful insight into their limitations. This is particularly important when performing re-weighted sampling, because of the risk of dramatically degraded sampling statistics. In what follows we will consider a one-dimensional harmonic oscillator of frequency ω and mass m, simulated at temperature $1/k_B\beta$ and using P beads. Consider for instance the expectation value of the difference Hamiltonian for the TD-FEP estimator (19). One can first transform into the normal-mode coordinates of the path $\tilde{q}_k$,[41] and get

$$\langle h_{\mathrm{TD}}\rangle_m = \frac{(\alpha-1)m\beta}{2P}\sum_k \omega_k^2\langle\tilde{q}_k^2\rangle_m. \tag{A1}$$

One can then compute analytically the average fluctuations of the normal mode coordinates, and take the limit of an infinite number of beads analytically, which yields

$$\begin{aligned}\langle h_{\mathrm{TD}}\rangle_m &= \frac{\alpha-1}{2}\left(P-\sum_k \frac{1}{1+\omega_k^2/\omega^2}\right)\underset{P\to\infty}{=}\\ &= \frac{\alpha-1}{2}\left(P-\frac{x}{2}\coth\frac{x}{2}\right).\end{aligned} \tag{A2}$$

Here we introduced the frequency of the kth normal mode of the free ring polymer, $\omega_k = 2\omega_P \sin k\pi/P$, and the shorthand $x = \beta\hbar\omega$. Proceeding in a similar way, one can obtain the fluctuations of h_{TD},

$$\begin{aligned}\sigma^2(h_{\mathrm{TD}}) &= \left[\frac{(\alpha-1)m\beta}{2P}\right]^2 2\sum_k\left[\omega_k^2\langle\tilde{q}_k^2\rangle_m\right]^2\\ &= \frac{(\alpha-1)^2}{2}\left(P+\sum_k\frac{1-2\left(1+\omega_k^2/\omega^2\right)}{\left(1+\omega_k^2/\omega^2\right)^2}\right)\underset{P\to\infty}{=}\\ &= \frac{(\alpha-1)^2}{2}\left[P-\frac{3}{4}x\coth\frac{x}{2}+\frac{1}{2}\left(\frac{x/2}{\sinh x/2}\right)^2\right].\end{aligned} \tag{A3}$$

As one could have imagined, fluctuations exhibit a term growing linearly with P, which in the context of re-weighted sampling has particularly disruptive consequences.

The procedure for the SC-FEP estimator (21) is analogous. One evaluates the expectation value of the difference Hamiltonian

$$\begin{aligned}\langle h_{\mathrm{SC}}\rangle_m &= \frac{(\alpha^{-1}-1)m\beta}{2P}\sum_{k>0}\omega^2\langle\tilde{q}_k^2\rangle_m \\ &= \frac{\alpha^{-1}-1}{2}\sum_{k>0}\frac{1}{1+\omega_k^2/\omega^2} \underset{P\to\infty}{=} \\ &= \frac{\alpha^{-1}-1}{2}\left(\frac{x}{2}\coth\frac{x}{2}-1\right),\end{aligned} \tag{A4}$$

which does not contain an explicit P dependence in the asymptotic limit. This is also true of the fluctuations that evaluate to

$$\begin{aligned}\sigma^2(h_{\mathrm{SC}}) &= \left[\frac{(\alpha^{-1}-1)m\beta}{2P}\right]^2 2\sum_{k>0}\left[\omega^2\langle\tilde{q}_k^2\rangle_m\right]^2 \\ &= \frac{(\alpha^{-1}-1)^2}{2}\sum_{k>0}\frac{1}{\left(1+\omega_k^2/\omega^2\right)^2} \underset{P\to\infty}{=} \\ &= \frac{(\alpha^{-1}-1)^2}{2}\left[\frac{x}{4}\coth\frac{x}{2}+\frac{1}{2}\left(\frac{x/2}{\sinh x/2}\right)^2-1\right].\end{aligned} \tag{A5}$$

Equations (A3) and (A5) capture the essential statistical properties of the TD-FEP and SC-FEP estimators, and can be profitably used to assess the range of applicability of the two methods by computing analytically the order of magnitude of the statistical errors one may expect for a system with known normal mode frequencies.

[1] T. Chacko, D. R. Cole, and J. Horita, Rev. Mineral. Geochem. **43**, 1 (2001).
[2] M. Wolfsberg, W. Van Hook, P. Paneth, and L. Rebelo, *Isotope Effects: In the Chemical, Geological, and Bio Sciences* (Springer, 2009).
[3] M. J. Gillan, Phys. Rev. Lett. **58**, 563 (1987).
[4] G. A. Voth, D. Chandler, and W. H. Miller, J. Chem. Phys. **91**, 7749 (1989).
[5] G. Mills, G. Schenter, D. Makarov, and H. Jónsson, Chem. Phys. Lett. **278**, 91 (1997).
[6] J. Cao and G. A. Voth, J. Chem. Phys. **101**, 6168 (1994).
[7] I. R. Craig and D. E. Manolopoulos, J. Chem. Phys. **121**, 3368 (2004).
[8] B. Chen, I. Ivanov, M. L. Klein, and M. Parrinello, Phys. Rev. Lett. **91**, 215503 (2003).
[9] S. Habershon, T. E. Markland, and D. E. Manolopoulos, J. Chem. Phys. **131**, 024501 (2009).
[10] X. Z. Li, B. Walker, and A. Michaelides, Proc. Natl. Acad. Sci. U.S.A. **108**, 6369 (2011).
[11] T. E. Markland and B. J. Berne, Proc. Natl. Acad. Sci. U.S.A. **109**, 7988 (2012).
[12] E. Wigner, Phys. Rev. **40**, 749 (1932).
[13] J. G. Kirkwood, Phys. Rev. **44**, 31 (1933).
[14] K. F. Herzfeld and E. Teller, Phys. Rev. **54**, 912 (1938).
[15] A. A. Chialvo and J. Horita, J. Chem. Phys. **130**, 094509 (2009).
[16] R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York, 1964).
[17] D. Chandler and P. G. Wolynes, J. Chem. Phys. **74**, 4078 (1981).
[18] M. Parrinello and A. Rahman, J. Chem. Phys. **80**, 860 (1984).
[19] M. E. Tuckerman, B. J. Berne, G. J. Martyna, and M. L. Klein, J. Chem. Phys. **99**, 2796 (1993).
[20] G. J. Martyna, A. Hughes, and M. E. Tuckerman, J. Chem. Phys. **110**, 3275 (1999).
[21] M. Ceriotti, M. Parrinello, T. E. Markland, and D. E. Manolopoulos, J. Chem. Phys. **133**, 124104 (2010).
[22] T. E. Markland and D. E. Manolopoulos, J. Chem. Phys. **129**, 024105 (2008).
[23] T. E. Markland and D. E. Manolopoulos, Chem. Phys. Lett. **464**, 256 (2008).
[24] G. S. Fanourgakis, T. E. Markland, and D. E. Manolopoulos, J. Chem. Phys. **131**, 094102 (2009).
[25] M. Ceriotti, G. Bussi, and M. Parrinello, Phys. Rev. Lett. **103**, 030603 (2009).
[26] M. Ceriotti, D. E. Manolopoulos, and M. Parrinello, J. Chem. Phys. **134**, 084104 (2011).
[27] M. Ceriotti and M. Manolopoulos, Phys. Rev. Lett. **109**, 100604 (2012).
[28] D. Marx and M. Parrinello, J. Chem. Phys. **104**, 4077 (1996).
[29] D. Marx, M. E. Tuckerman, and G. J. Martyna, Comput. Phys. Commun. **118**, 166 (1999).
[30] I. Ufimtsev and T. Martinez, Comput. Sci. Eng. **10**, 26 (2008).
[31] T. Fujita, H. Watanabe, and S. Tanaka, J. Phys. Soc. Jpn. **78**, 104723 (2009).
[32] J. Vanicek and W. H. Miller, J. Chem. Phys. **127**, 114309 (2007).
[33] F. Paesani and G. A. Voth, J. Phys. Chem. B **113**, 5702 (2009).
[34] A. Perez and O. A. von Lilienfeld, J. Chem. Theory Comput. **7**, 2358 (2011).
[35] P. Pavone and S. Baroni, Solid State Commun. **90**, 295 (1994).
[36] M. Ceriotti, G. Miceli, A. Pietropaolo, D. Colognesi, A. Nale, M. Catti, M. Bernasconi, and M. Parrinello, Phys. Rev. B **82**, 174306 (2010).
[37] L. Lin, J. A. Morrone, R. Car, and M. Parrinello, Phys. Rev. B **83**, 220302 (2011).
[38] M. Krzystyniak and F. Fernandez-Alonso, Phys. Rev. B **83**, 134305 (2011).
[39] A. G. Seel, M. Ceriotti, P. P. Edwards, and J. Mayers, J. Phys.: Condens. Matter **24**, 365401 (2012).
[40] A. Pohorille, L. R. Pratt, R. A. LaViolette, M. A. Wilson, and R. D. MacElroy, J. Chem. Phys. **87**, 6070 (1987).
[41] B. J. Berne and D. Thirumalai, Annu. Rev. Phys. Chem. **37**, 401 (1986).
[42] M. F. Herman, E. J. Bruskin, and B. J. Berne, J. Chem. Phys. **76**, 5150 (1982).
[43] D. M. Ceperley, Rev. Mod. Phys. **67**, 279 (1995).
[44] J. Cao and B. J. Berne, J. Chem. Phys. **91**, 6359 (1989).
[45] M. Ceriotti, G. A. R. Brain, O. Riordan, and D. E. Manolopoulos, Proc. R. Soc. London, Ser. A **468**, 2 (2012).
[46] T. M. Yamamoto, J. Chem. Phys. **123**, 104101 (2005).
[47] J. Worden, D. Noone, and K. Bowman, Nature (London) **445**, 528 (2007).
[48] R. A. Berner, S. T. Petsch, J. A. Lake, D. J. Beerling, B. N. Popp, R. S. Lane, E. A. Laws, M. B. Westley, N. Cassar, F. I. Woodward *et al.*, Science **287**, 1630 (2000).
[49] R. C. Paniello, J. M. D. Day, and F. Moynier, Nature (London) **490**, 376 (2012).
[50] G. Bussi, D. Donadio, and M. Parrinello, J. Chem. Phys. **126**, 014101 (2007).
[51] J. A. Morrone, V. Srinivasan, D. Sebastiani, and R. Car, J. Chem. Phys. **126**, 234504 (2007).
[52] J. A. Morrone and R. Car, Phys. Rev. Lett. **101**, 017801 (2008).
[53] J. Horita and D. J. Wesolowski, Geochim. Cosmochim. Acta **58**, 3425 (1994).