JCTC Journal of Chemical Theory and Computation

ARTICLE

pubs.acs.org/JCTC

# Path Integral Computation of Quantum Free Energy Differences Due to Alchemical Transformations Involving Mass and Potential

Alejandro Pérez*

Department of Chemistry, New York University, New York, New York 10003, United States, Nano-bio spectroscopy group, Centro Joxe Mari Korta, Avenida de Tolosa, 72, E-20018 Donostia-San Sebastian, Spain, and Institute of Pure and Applied Mathematics, University of California Los Angeles, Los Angeles, California 90095, United States

O. Anatole von Lilienfeld*

Surface and Interface Sciences Department, Sandia National Laboratories, Albuquerque, New Mexico 87185, United States, Argonne Leadership Computing Facility, Argonne National Laboratory, Argonne, Illinois 60439, United States, and Institute of Pure and Applied Mathematics, University of California Los Angeles, Los Angeles, California 90095, United States

**ABSTRACT:** Thermodynamic integration, perturbation theory, and $\lambda$-dynamics methods were applied to path integral molecular dynamics calculations to investigate free energy differences due to "alchemical" transformations. Several estimators were formulated to compute free energy differences in solvable model systems undergoing changes in mass and/or potential. Linear and nonlinear alchemical interpolations were used for the thermodynamic integration. We find improved convergence for the virial estimators, as well as for the thermodynamic integration over nonlinear interpolation paths. Numerical results for the perturbative treatment of changes in mass and electric field strength in model systems are presented. We used thermodynamic integration in *ab initio* path integral molecular dynamics to compute the quantum free energy difference of the isotope transformation in the Zundel cation. The performance of different free energy methods is discussed.

## I. INTRODUCTION

The quantum nature of nuclei plays an important role in the prediction of many properties of systems containing light elements such as hydrogen. A successful approach to investigating such effects relies on first principles (*ab initio*) based path-integral molecular dynamics (AI-PIMD).[1,2] In AI-PIMD, the nuclei are quantized following Feynman's path integral (PI) formalism,[3,4] while the many-body potential of the system is computed "on-the-fly" from electronic structure calculations.[5,6] Proton transfer events are now routinely investigated using AI-PIMD in various hydrogen-bonded systems in gas,[7−10] condensed phases,[11,12] or even on surfaces.[13,14]

Nuclear quantum effects in biological systems have attracted considerable attention in recent years. Various methods have been developed and successfully deployed to investigate nuclear quantum effects in biological systems.[15,16] For example, using AI-PIMD, we recently identified nuclear quantum effects to be crucial for kinetically stabilizing the canonical Watson−Crick base pair configuration by rendering their rare enol tautomer form metastable.[17]

In this study, we explore "alchemical" changes within the path integral formalism. By alchemical changes, we refer to changes of parameters that define a given Hamiltonian, such as atomic numbers, atomic masses, number of particles, or the interaction potential. Studying changes in composition (atomic numbers) holds great promise for future applications in the arena of rational computational materials design, i.e., approaches that attempt to outperform a mere screening of combinatorial sets of materials candidates.[18−23] The computation of free energy differences[24] between compounds based on thermodynamic integration,[25] or free-energy perturbations,[26] has only recently been applied using electronic structure methods.[27−29] For example, a related framework was used to successfully determine the chemical composition of the Earth's core.[30] To the best of our knowledge, similar alchemical changes have not been hitherto investigated within PIMD.

Alchemical transformations involving changes in atomic numbers could one day lead to the engineering of tunneling barriers in proton-conducting materials. For changes involving only the mass, such an approach permits the direct evaluation of kinetic isotope effects. Only very recently, and while we carried out our study, an adaptation of AI-PIMD for the computation of isotope effects in intramolecular hydrogen transfer *via* thermodynamic integration was published.[31] More recently, Ramirez and Herrero applied thermodynamic integration in path integrals to investigate the isotope effects on the melting temperature of ice Ih.[32] Finally, we also note a recent application of the Jarzynski equality within PIMD.[33]

In this study, we discuss analytically solvable model systems subjected to changes in mass and potential. Free energy differences were computed numerically using thermodynamic integration and $\lambda$-dynamics methods.[34] We furthermore consider the perturbative treatment of changes in mass and static electric fields. While ultimately AI-PIMD will be used to investigate such alchemical changes in more realistic systems, our focus here is on

**Received:** January 24, 2011
**Published:** June 06, 2011

formulating and assessing the methods for simple model systems. To the best of our knowledge, a path integral scheme to generally estimate free energy changes due to alchemical transformations has yet to be presented.

This paper is organized as follows: In the Theory section, we first recall Feynman's path integral expression for the Helmholtz free energy. Then, we describe thermodynamic integration, perturbation theory, and $\lambda$-dynamics methods applied to PIMD. In the subsequent Results section, the investigated model systems and numerical results are discussed. The paper ends with the Conclusions section and the Appendix.

## II. THEORY

**A. Path Integral Expression for the Helmholtz Free Energy.** Feynman's PI formulation[3,4] of quantum statistical mechanics describes each nucleus by a ring polymer of $P$ (known as the Trotter number) interacting quasi-particles, colloquially referred to as "beads". In the usual formulation, these beads are connected by harmonic springs and experience the effect of an attenuated classical potential.[35] Path configurations can then be sampled either stochastically via Monte Carlo methods[36] or deterministically using molecular dynamics.[37] Thus, path integral molecular dynamics (PIMD) is a straightforward way to include nuclear quantum effects in ordinary molecular dynamics (MD) simulations.

The discrete path integral expression for the canonical quantum partition function of a single particle in 3D is[34]

$$Z(\lambda) \approx C(\lambda) \int \mathrm{d}\mathbf{r}_1 \ldots \mathrm{d}\mathbf{r}_P \, \mathrm{e}^{-\beta\Phi(\{r\};\lambda)},$$

with

$$C(\lambda) = \left(\frac{m(\lambda)P}{2\pi\beta\hbar^2}\right)^{3P/2} \tag{1}$$

where $P$ is the Trotter number, i.e., the number of replicas or beads[35] representing the quantum particle, and $\Phi$ is the path integral effective potential,

$$\Phi(\{\mathbf{r}\};\lambda) = \sum_{s=1}^{P} \frac{m(\lambda)}{2} \omega_P^2 (\mathbf{r}_s - \mathbf{r}_{s+1})^2 + \sum_{s=1}^{P} \frac{V(\mathbf{r}_s;\lambda)}{P} \tag{2}$$

where $\omega_P{}^2 = P/(\beta\hbar)^2$, $\{\mathbf{r}\}$ denotes the bead coordinates, and $V$ is the classical potential. The paths obey the cyclic condition $\mathbf{r}_{P+1} = \mathbf{r}_1$ imposed by the quantum-mechanical trace operation $Z = Tr[\exp(-\beta\hat{H})]$. Equation 1 shows explicitly the parametric dependence on $\lambda$ driving the transformation in mass $m(\lambda)$ and potential (see next section).

The canonical average of any observable $O$ at a fixed $\lambda$ value, $\langle O\rangle_\lambda$, is given by the following integral over the path configurations:

$$\langle O\rangle_\lambda \approx \frac{C(\lambda)}{Z(\lambda)} \int \mathrm{d}\mathbf{r}_1 \ldots \mathrm{d}\mathbf{r}_P \, O(\mathbf{r}_1, \cdots, \mathbf{r}_P;\lambda) \, \mathrm{e}^{-\beta\Phi(\{\mathbf{r}\};\lambda)} \tag{3}$$

In practical path integral calculations, one performs a suitable coordinate transformation of eq 2 to either normal modes or staging coordinates.[37] To carry out PIMD, a Hamiltonian is constructed by adding a set of momenta to this transformed effective path integral potential. The dynamical masses of these pseudoparticles are then chosen so as to bring all free-particle modes to the same time scale. Finally, thermostats are attached to each bead to improve ergodic canonical sampling.[37]

**B. Thermodynamic Integration and Path Integrals.** The quantum free energy change of a system due to an alchemical transformation is, by definition,

$$\Delta F = F_\mathrm{f} - F_\mathrm{i} = -\frac{1}{\beta}\ln\left(\frac{Z_\mathrm{f}}{Z_\mathrm{i}}\right) \tag{4}$$

where $Z_\mathrm{i}(F_\mathrm{i})$ and $Z_\mathrm{f}(F_\mathrm{f})$ denote the initial and final canonical quantum partition functions (Helmholtz free energies), respectively. Equation 4 can be evaluated numerically using the thermodynamic integration (TI) method:[25,38,39]

$$\Delta F = \int_0^1 \mathrm{d}\lambda \left(\frac{\mathrm{d}F(\lambda)}{\mathrm{d}\lambda}\right)_\lambda \tag{5}$$

where $\lambda$ is a coupling parameter that drives the transformation from an initial state i to a final state f. We remark that while it is well-known that eq 5 can be identified with $\int \mathrm{d}\lambda\langle\partial_\lambda V\rangle$ in classical statistical mechanics ($V$ being the potential energy),[34] this is not the case for PI (see below).

Consider now a single quantum particle undergoing a change of identity, i.e., a simultaneous change of mass and potential energy function. Since the Helmholtz free energy $F$ is a state function, $\Delta F$ is independent of how the two states are interpolated using the $\lambda$ variable, provided that the two end points are met at $\lambda = 0$ and 1. For simplicity, consider a linear interpolation path and a single $\lambda$ parameter driving both transformations

$$\begin{aligned} V(\lambda) &= \lambda V_\mathrm{f} + (1-\lambda)V_\mathrm{i} \\ m(\lambda) &= \lambda m_\mathrm{f} + (1-\lambda)m_\mathrm{i} \end{aligned} \tag{6}$$

with $0 \le \lambda \le 1$. Obviously, the intermediate $\lambda$ states do not correspond to any real system but are only a (perfectly rigorous) way of obtaining the corresponding free energy difference. Although in eq 6 a single parameter $\lambda$ is used, it may be more efficient to use two independent control parameters, $\lambda_\mathrm{m}$ and $\lambda_\nu$, for the mass and potential transformation, respectively.

As noted above, any interpolation scheme is permissible provided that the end point condition $Z(\lambda = 1) = Z_\mathrm{f}$ and $Z(\lambda = 0) = Z_\mathrm{i}$ is satisfied. Although a linear interpolation is the simplest choice, it is not necessarily the most efficient. An optimal choice is the scheme that samples many function values in the region of the $\lambda$ domain where the function to be integrated varies rapidly. In the Results and Discussion section, we explore in detail the relative performance of linear and nonlinear interpolations. Specifically, a simple quartic interpolation in $\lambda$ will be discussed

$$\begin{aligned} V(\lambda) &= \lambda^4 V_\mathrm{f} + (1-\lambda^4)V_\mathrm{i} \\ m(\lambda) &= \lambda^4 m_\mathrm{f} + (1-\lambda^4)m_\mathrm{i} \end{aligned} \tag{7}$$

As expected, different interpolation functions lead to different convergence properties for thermodynamic integration using PIMD (TI-PIMD).

**C. TI-PIMD: Changes in Mass.** The TI-PIMD evaluation of the quantum free energy difference associated with a change of mass enables us to predict isotope effects. For changes being restricted to the mass, i.e., $V$ being independent of $\lambda$, the first

derivative of eq 1 with respect to $\lambda$ yields the average of the free energy derivative:

$$\left(\frac{\mathrm{d}F}{\mathrm{d}\lambda}\right)^{\mathrm{prim}} = -\frac{1}{\beta}\frac{\partial_\lambda Z(\lambda)}{Z(\lambda)} \approx -\left(\frac{\partial_\lambda m(\lambda)}{m(\lambda)}\right) \times\left[\frac{3P}{2\beta}-\frac{m}{2}\omega_P^2\sum_{s=1}^{P}\langle(\mathbf{r}_s-\mathbf{r}_{s+1})^2\rangle_\lambda\right] \quad (8)$$

Here, the term $3P/(2\beta)$ comes from the normalization prefactor $C(\lambda)$ in eq 1 . Equation 8 corresponds to the *primitive* estimator of the first-order derivative of the free energy with respect to mass. This expression is expected to have the same statistical problems as the *primitive* estimator for the quantum kinetic energy.[40] Better convergence, however, can be achieved using the *virial* relation for bound systems in path integrals:[34,40]

$$\left\langle\frac{3P}{2\beta}-\frac{m\omega_P^2}{2}\sum_{s=1}^{P}(\mathbf{r}_s-\mathbf{r}_{s+1})^2\right\rangle = \left\langle\frac{1}{2P}\sum_{s=1}^{P}\mathbf{r}_s\cdot\frac{\partial V}{\partial\mathbf{r}_s}\right\rangle \quad (9)$$

Insertion in eq 8 yields the virial estimator for a bound particle

$$\left(\frac{\mathrm{d}F}{\mathrm{d}\lambda}\right)^{\mathrm{vir}} \approx -\left(\frac{\partial_\lambda m(\lambda)}{m(\lambda)}\right)\left\langle\frac{1}{2P}\sum_{s=1}^{P}\mathbf{r}_s\cdot\frac{\partial V(\mathbf{r}_1\ldots\mathbf{r}_P)}{\partial\mathbf{r}_s}\right\rangle_\lambda \quad (10)$$

For a homogeneous system of $N$ particles (all undergoing a simultaneous change in mass), the virial expression becomes

$$\left(\frac{\mathrm{d}F}{\mathrm{d}\lambda}\right)^{\mathrm{vir}} \approx -\left(\frac{\partial_\lambda m(\lambda)}{m(\lambda)}\right) \times\left\langle\frac{3N}{2\beta}+\frac{1}{2P}\sum_{I=1}^{N}\sum_{s=1}^{P}(\mathbf{r}_{I,s}-\bar{\mathbf{r}}_I)\cdot\frac{\partial V}{\partial\mathbf{r}_{I,s}}\right\rangle_\lambda \quad (11)$$

where $\bar{\mathbf{r}}_I = \sum_{s=1}^{P}\mathbf{r}_{I,s}/P$ is the centroid of the ring polymer describing particle $I$.

If only the atoms belonging to a certain element (usually, hydrogen) are undergoing the isotope transformation, then

$$\left(\frac{\mathrm{d}F}{\mathrm{d}\lambda}\right)^{\mathrm{vir}} \approx -\sum_{I=1}^{N_c}\left(\frac{\partial_\lambda m}{m_I}\right) \times\left[\frac{3}{2\beta}+\left\langle\frac{1}{2P}\sum_{s=1}^{P}(\mathbf{r}_{I,s}-\bar{\mathbf{r}}_I)\cdot\nabla_{\mathbf{r}_I}V(\mathbf{r}_s)\right\rangle_\lambda\right] \quad (12)$$

where $N_c$ is the number of atoms in the subset. Similarly, the primitive estimators is

$$\left(\frac{\mathrm{d}F}{\mathrm{d}\lambda}\right)^{\mathrm{prim}} \approx -\sum_{I=1}^{N_c}\left(\frac{\partial_\lambda m}{m_I}\right) \times\left[\frac{3P}{2\beta}-\left\langle\frac{m_I}{2}\omega_P^2\sum_{s=1}^{P}(\mathbf{r}_{I,s}-\mathbf{r}_{I,s+1})^2\right\rangle_\lambda\right] \quad (13)$$

It is well-known that *virial* estimators exhibit less fluctuations than their *primitive* counterparts.[34,40] In the Results and Discussion section, we confirm the favorable statistical properties of the virial estimator.

**D. TI-PIMD: Changes in Potential.** In the same vein as isotope transformations, changes in the free energy due to varying the underlying classical potential can be also evaluated using TI-PIMD. The path integral estimator for the free energy derivative with respect to changes in the potential $V$ is given by

$$\frac{\mathrm{d}F}{\mathrm{d}\lambda} = -\frac{1}{\beta}\left(\frac{\partial_\lambda Z}{Z}\right) \approx \left\langle\sum_{s=1}^{P}\frac{\partial_\lambda V(\mathbf{r}_s;\lambda)}{P}\right\rangle_\lambda \quad (14)$$

In contrast to changes in the mass, there is generally only one version of the estimator for changes in the potential.

As an illustrative example, the free energy dependence on the force constant $\kappa(\lambda)$ of the harmonic potential ($V(x;\lambda) = \kappa(\lambda)x^2/2$) can be evaluated according to

$$\frac{\mathrm{d}F(\lambda)}{\mathrm{d}\lambda} \approx \left\langle\frac{\partial\kappa(\lambda)}{\partial\lambda}\sum_{s=1}^{P}\frac{x_s^2}{2P}\right\rangle_\lambda \quad (15)$$

The numerical result of this expression can be compared to the analytical free energy derivative with respect to force constant for the harmonic potential,

$$\frac{\mathrm{d}F}{\mathrm{d}\lambda} = \frac{\hbar}{4}\frac{\partial_\lambda\kappa(\lambda)}{\sqrt{m\kappa(\lambda)}}\coth\left(\frac{\beta\hbar}{2}\sqrt{\frac{\kappa(\lambda)}{m}}\right) \quad (16)$$

Numerical results will be presented and discussed below.

**E. TI-PIMD: Changes in Mass and Potential.** We now consider a simultaneous transformation in mass and potential state function using TI-PIMD. The first derivative of the quantum free energy (see eq 1) with respect to parameter $\lambda$ is simply the combination of eq 8 and eq 14

$$\frac{\mathrm{d}F}{\mathrm{d}\lambda} = -\frac{1}{\beta}\left(\frac{\partial_\lambda Z}{Z}\right) \approx -\frac{1}{\beta}\langle\mathscr{T}\rangle_\lambda \quad (17)$$

where the function to be averaged is

$$\mathscr{T} = \frac{3P}{2m(\lambda)}\frac{\partial m(\lambda)}{\partial\lambda}-\beta\sum_{s=1}^{P}\frac{\partial_\lambda V(\{\mathbf{r}_s\};\lambda)}{P} -\beta\left[\frac{\partial_\lambda m(\lambda)}{2}\omega_P^2\sum_{s=1}^{P}(\mathbf{r}_s-\mathbf{r}_{s+1})^2\right]$$

In eq 17, the primitive estimator (eq 8) was used for the contribution of the isotope transformation to the free energy change. Using eq 10, we can also obtain a "virial" expression for eq 17.

The analytical expression for the harmonic oscillator undergoing a simultaneous change in mass $m(\lambda)$ and force constant $\kappa(\lambda)$ is

$$\frac{\mathrm{d}F}{\mathrm{d}\lambda} = \frac{\hbar}{4\sqrt{\kappa(\lambda)/m(\lambda)}}\left[\frac{m(\lambda)\,\partial_\lambda\kappa(\lambda)-\kappa(\lambda)\,\partial_\lambda m(\lambda)}{m(\lambda)^2}\right] \times\coth\left(\frac{\beta\hbar}{2}\sqrt{\kappa(\lambda)/m(\lambda)}\right) \quad (18)$$

Numerical results and comparison will be presented and discussed below.

**F. Perturbative Treatment of Changes in Mass.** Path integral calculations are often very expensive computationally. Therefore, it would be desirable to have a method capable of estimating the isotope effect with less or even without any of the intermediate calculations required by TI-PIMD.

If the isotope transformation is not too large, perturbation theory[41] (PT) can be used to estimate the free energy change from a single PIMD simulation. The combination of free energy

perturbation and PIMD is termed here as PT-PIMD. The free energy difference due to a (small) change in mass of all atoms of a certain element (leaving the others unchanged) is

$$\Delta F \approx -\frac{1}{\beta}\ln\left[\left(\frac{m_f}{m_i}\right)^{3N_cP/2}\frac{\int d^{NP}\mathbf{r}\, e^{-\beta(\Phi_f-\Phi_i)}\, e^{-\beta\Phi_i}}{\int d^{NP}\mathbf{r}\, e^{-\beta\Phi_i}}\right]$$

$$= -\frac{1}{\beta}\left[\frac{3N_cP}{2}\ln\left(\frac{m_f}{m_i}\right)+\ln\langle e^{-\beta(m_f-m_i)A}\rangle_i\right],$$

$$\text{with } A = \frac{\omega_P^2}{2}\sum_{I=1}^{N_c}\sum_{s=1}^{P}(\mathbf{r}_{I,s}-\mathbf{r}_{I,s+1})^2 \quad (19)$$

where $N_c$ is the number of atoms undergoing the isotope transformation, $\Phi$ denotes the effective path integral potential similar to eq 2, and the angular brackets denote the canonical sampling over path configurations of a system where the mutating elements have fixed mass $m_i$.

From eq 19, it is clear that if the difference in mass is very large, there will be very little overlap between states, and the sampling will be very inefficient. However, even in this case, one can split the interval of interest into *nw* windows and perform PT-PIMD on each individual window. Then, the free energy difference is given by the sum over all PT-PIMD windows,

$$\Delta F = \sum_{i=0}^{nw-1}\Delta F(i\Delta\lambda \rightarrow (i+1)\Delta\lambda) \quad (20)$$

As long as the free energy change in each window is around $\sim 2k_BT$, we can efficiently sample eq 19 using PT-PIMD methods. For further statistics, the $nw - 2$ interior intervals can be sampled forward and backward within *double-wide sampling*.[38]

**G. λ-Dynamics Approach for Quantum Free Energy Differences: λ-PIMD.** The free energy difference between two chemical states can be also computed using the so-called λ-dynamics (LD) approach.[38,42−44] LD follows the same spirit as thermodynamic integration, but instead of running separate simulations for each fixed λ value, this parameter is allowed to vary continuously in a single (albeit long) MD simulation. Here, we propose to combine LD with PIMD. To this end, the λ parameter is endowed with dynamical character according to the following extended PIMD Hamiltonian:

$$H = \sum_{s=1}^{P}\left[\frac{p_s^2}{2m_s}+\frac{m}{2}\omega_P^2(\mathbf{r}_s-\mathbf{r}_{s+1})^2\right]+\frac{p_\lambda^2}{2m_\lambda}$$
$$+\sum_{s=1}^{P}[\lambda V_f(\mathbf{r}_s)+(1-\lambda)V_i(\mathbf{r}_s)]/P \quad (21)$$

which allows the continuous interpolation between initial ($V_i$) and final ($V_f$) potential energy functions. In LD, the associated mass of the λ particle ($m_\lambda$) has to be large enough to allow for proper equilibration of the other degrees of freedom. We term the scheme given by eq 21 as λ-PIMD.

The λ path chosen in eq 21 corresponds to a linear path. The coupling parameter obeys the constraint $\lambda \in [0, 1]$. Usually, this condition is enforced in MD via Lagrange multipliers. However, an elegant alternative (adopted here) is to use the trigonometric parametrization presented in ref 38. Using the change of variable $\lambda = \cos^2\theta$, with $\theta \in [0,\pi/2]$, in eq 21 yields

$$H = \sum_{s=1}^{P}\left[\frac{p_s^2}{2m_s}+\frac{m}{2}\omega_P^2(\mathbf{r}_s-\mathbf{r}_{s+1})^2\right]+\frac{p_\theta^2}{2m_\theta}$$
$$+\sum_{s=1}^{P}[V_f(\mathbf{r}_s)\cos^2\theta+V_i(\mathbf{r}_s)\sin^2\theta]/P \quad (22)$$

The velocity of the λ particle, $v_\theta = p_\theta/m_\theta$, is reversed whenever its position θ hits the boundaries.

The free energy change for this transformation is then obtained from numerical integration of the λ-derivative

$$\Delta F = F_f - F_i = \int_0^1 \langle\Delta V\rangle_\lambda\, d\lambda$$
$$= \int_0^{\pi/2}\langle\Delta V\rangle_\theta \sin(2\theta)\, d\theta \quad (23)$$

where $\Delta V(\theta) = (1/P)\sum_{s=1}^{P}[V_f(\mathbf{r}_s;\theta) - V_i(\mathbf{r}_s;\theta)]$. In eq 23, configurations from the λ-dynamics trajectory are histogrammed according to their $\lambda(\theta)$ value.[45]

We note in passing that in other λ-dynamics methods such as the one proposed in ref 46, λ switching functions are designed so as to induce a free energy barrier between the end points with the goal of enhancing the sampling of these states. Then, the λ particle is kept at a higher temperature than the system to promote barrier crossing along the λ trajectory.[46] We note that under the physical conditions of our calculations, we did not encounter significant free energy barriers, and the λ particle was thermalized at the same temperature as the physical system.

The λ-PIMD scheme presented here is general, and it could be used for instance to estimate quantum free energy changes for proton transfer, ligand binding affinities, or acid dissociation constants ($pK_a$). As an illustrative example, in the next section, we present an application of λ-PIMD to study changes in the force constant in the harmonic oscillator. By declaring the mass variable, it could also be used to investigate the isotope effect.

Finally, λ-dynamics could be easily extended by using $P$ independent λ particles. That is, an independent λ particle is associated with each bead of the ring polymer according to the Hamiltonian

$$H = \sum_{s=1}^{P}\left[\frac{p_s^2}{2m_s}+\frac{p_{\lambda,s}^2}{2m_{\lambda,s}}+\frac{m}{2}\omega_P^2(\mathbf{r}_s-\mathbf{r}_{s+1})^2\right]$$
$$+\sum_{s=1}^{P}[\lambda_s V_f(\mathbf{r}_s)+(1-\lambda_s)V_i(\mathbf{r}_s)]/P \quad (24)$$

## III. RESULTS AND DISCUSSION

Natural units ($\hbar = k_B = m = 1$) are used for the model systems investigated, except for the Zundel cation, hydrogen molecule, and Morse potential, where we used atomic units (au).

**A. TI-PIMD: Changes in Mass.** *1. Quantum Harmonic Oscillator.* Using TI-PIMD, the isotope effect was investigated for the harmonic potential $V(x) = m(\lambda)\omega^2x^2/2$. For this system, the difference in quantum free energy is known analytically

$$\Delta F = F(m_f) - F(m_i) = \frac{1}{\beta}\ln\left[\frac{\sinh(\beta\hbar\omega_f/2)}{\sinh(\beta\hbar\omega_i/2)}\right] \quad (25)$$

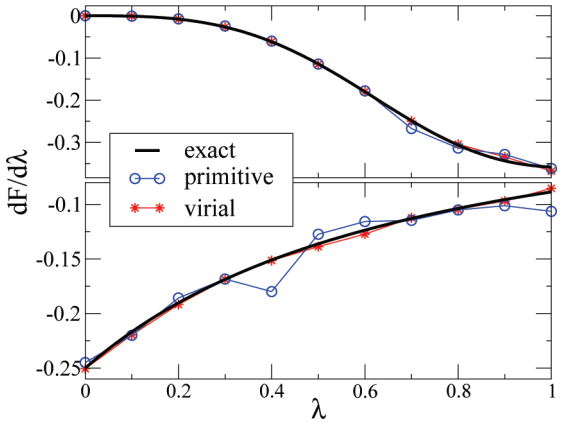


**Figure 1.** TI-PIMD: First derivative of the quantum free energy with respect to mass for the harmonic potential $V = m(\lambda)\,\omega^2 x^2/2$. The mass $m(\lambda)$ is a linear (bottom) or quartic (top) interpolation between $m_i = 1 \rightarrow m_f = 2$. Black: exact result (eq 26). Red: virial estimator (eq 10). Blue: primitive estimator (eq 8). Natural units ($\hbar = k_B = 1$) are used. The temperature was set to $\beta\hbar\omega = 10$, and PIMD (with $P = 64$) was conducted for 400 000 steps.

where the oscillator frequencies are ${\omega_\lambda}^2 = \kappa/m(\lambda)$ and $\kappa$ is the (fixed) force constant. The derivative of the free energy with respect to $\lambda$ is also known analytically,

$$\frac{dF}{d\lambda} = -\frac{\hbar}{4}\left(\frac{\partial_\lambda m(\lambda)}{m(\lambda)}\right)\omega_\lambda \coth(\beta\hbar\omega_\lambda/2) \tag{26}$$

Note that the value of eq 26 depends on the interpolation function adopted for $m(\lambda)$.

Figure 1 (bottom) shows the primitive, virial, and exact first-order derivative of the free energy as a function of mass using a linear path. Using the same number of PIMD steps (400 000), the convergence of the primitive estimator (blue) is worse than the virial estimator (red), which is nearly identical to the exact value given by eq 26. The corresponding change of free energy as obtained from thermodynamic integration (eq 5) yields −0.154 and −0.148 for the primitive and virial estimators, respectively. The exact value from eq 25, $\Delta F = -0.147$, compares favorably with these numerical results using eqs 8 and 10. The average error bar for the primitive estimator is 0.0091, which is 5 times larger than the error bar for the virial estimator (0.0018), both for the linear path. This result confirms that the virial expression usually has less statistical error and is preferable for efficient computation of quantum free energy differences.

Figure 1 (top) shows numerical results for the virial and primitive estimators of the free energy derivative for the same system but with the quartic interpolation in mass, $m(\lambda) = \lambda^4 m_f + (1-\lambda^4)m_i$. Using the same number of normal mode PIMD steps (400 000) as Figure 1 (bottom), the nonlinear interpolation (top) clearly improves the sampling for both estimators. In particular, the statistical error of the primitive estimator (0.0054) is greatly reduced, whereas the virial remains the same: 0.0018. Both estimators yield results nearly indistinguishable (and therefore converged) from the exact values, eq 26 . In this case, the change of free energy integrates to −0.151 for the primitive and to −0.150 for the virial estimator, the exact value being −0.147.

*2. Double Quantum Harmonic Oscillator.* We investigated a three-body system that resembles the highly relevant proton transfer scenario in many naturally occurring systems, such as in ref 17. This model system consists of three collinear masses $S_1{-}S_2\cdots S_3$, where $S_1$ and $S_3$ sites represent identically electronegative "heavy" atoms (e.g., oxygens) with equal masses ($m_1 = m_3$). The extension to the asymmetric, heterogeneous ($m_1 \neq m_3$) case is straightforward. The intermediate atom $S_2$ is supposed to represent the light atom that will undergo a change in mass (from hydrogen to deuterium). The particles are assumed to interact harmonically. The Hamiltonian operator of this system is

$$\hat{\mathscr{H}} = -\frac{\hbar^2}{2}\sum_{i=1}^{3}\frac{1}{m_i}\frac{\partial^2}{\partial_{x_i}^2} + \frac{\kappa}{2}\left[(x_2 - x_1 - x_{eq})^2 + (x_3 - x_2 - x_{eq})^2\right] \tag{27}$$

where $\kappa$ is the force constant and $x_{eq}$ is the equilibrium distance. This many-body system can be solved exactly using Jacobi coordinates

$$\begin{aligned} x_s &= \frac{1}{\sqrt{2}}(x_3 - x_1 - 2x_{eq}) \\ x_a &= \frac{1}{\sqrt{2}}(x_1 + x_3 - 2x_2) \\ X &= [m_1(x_1 + x_3) + m_2 x_2]/(2m_1 + m_2) \end{aligned}$$

where $x_a$ and $x_s$ represent the asymmetric and symmetric stretching modes, respectively, and $X$ is the center of mass position. In terms of these Jacobi coordinates, the Hamiltonian becomes separable

$$\begin{aligned} \hat{\mathscr{H}} = &\left[-\frac{\hbar^2}{2}\frac{\partial_X^2}{m_X}\right] + \left[-\frac{\hbar^2}{2}\frac{\partial_s^2}{m_s} + \frac{m_s}{2}\omega_s^2 x_s^2\right] \\ &+ \left[-\frac{\hbar^2}{2}\frac{\partial_a^2}{m_a} + \frac{m_a}{2}\omega_a^2 x_a^2\right] \end{aligned} \tag{28}$$

where the effective masses $m_a = m_1 m_2/(2m_1 + m_2)$ and $m_s = m_1$ and corresponding frequencies $w_n = (\kappa/m_n)^{1/2}$, $n$ = a, s, have been introduced. In eq 28, the first term represents the center of mass motion, the second term is the symmetric vibrational mode, and the third term is the asymmetric stretch. Note that for this simple system, $x_a$ is the only mode that involves motion of the light atom $S_2$, thereby exclusively contributing to the isotope effect. The symmetric mode $x_s$ (only the end point masses move in opposite directions) does not contribute to the isotope effect.

The first derivative of the free energy with respect to the mass of $S_2$ can be solved analytically,

$$\frac{dF}{d\lambda} = -\frac{\hbar\omega_a}{4}\left(\frac{\partial_\lambda m_a}{m_a}\right)\coth\left(\frac{\beta\hbar\omega_a}{2}\right) - \frac{1}{2\beta}\left(\frac{\partial_\lambda m_2}{M}\right) \tag{29}$$

where $M = 2m_1 + m_2$ is the total mass of the system and $\partial_\lambda m_a$ and $\partial_\lambda m_2$ are the interpolation-function-dependent partial derivatives of the effective asymmetric and transforming masses, respectively. The first and second terms in eq 29 are the contributions from the asymmetric stretching mode and from the free motion of the center of mass, respectively.

For this model system, the numerical and analytic results are presented in Figure 2 for the quartic interpolation of the mass in $\lambda$, $m_2(\lambda) = \lambda^4 m_D + (1-\lambda^4)m_H$. The partial derivatives for this quartic interpolation path are $\partial_\lambda m_a(\lambda) = 2{m_1}^2 m_2(\lambda)/(2m_1 + m_2(\lambda))^2$ and $\partial_\lambda m_2(\lambda) = 4\lambda^3$. The initial hydrogen mass $m_H = 1$ was changed to the mass of deuterium $m_D = 2$. The equilibrium

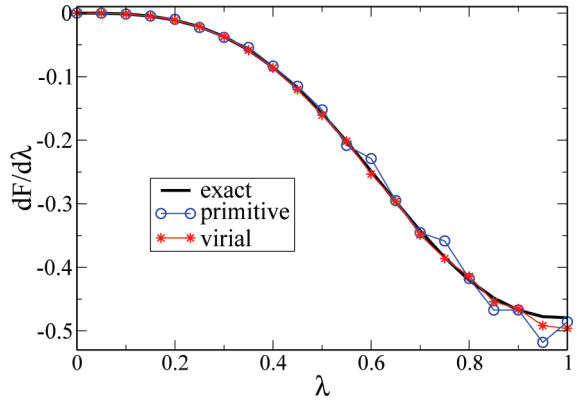


**Figure 2.** TI-PIMD: First order derivative of the quantum free energy with respect to mass of the central atom in the double oscillator model (eq 27). A quartic nonlinear interpolation path $m(\lambda) = \lambda^4 m_D + (1 - \lambda^4) m_H$ between the initial mass of hydrogen ($m_H = 1$) and final mass of deuterium ($m_D = 2$) was used. Natural units are used. Black: exact result (eq 29). Red: virial estimator (eq 10). Blue: primitive estimator (eq 8).

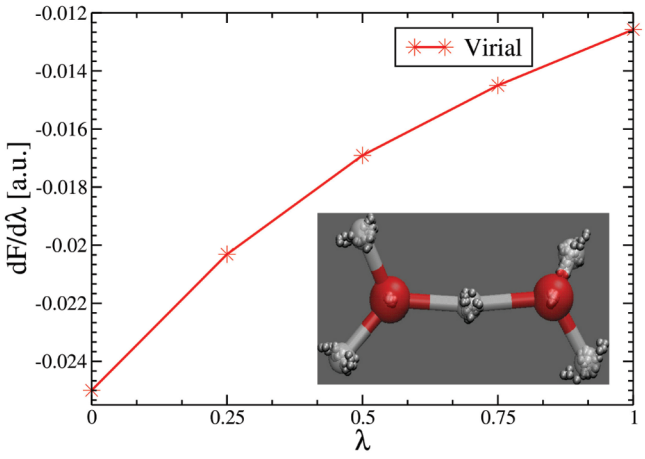


**Figure 3.** Results for the virial estimator for the *ab initio* TI-PIMD simulation ($P = 32$) of the Zundel cation at 300 K. The predicted change in free energy for the all-H to all-D transformation is $\Delta F = -0.01755$ au ($-11.02$ kcal/mol). The inset shows a typical snapshot (created with the visualization program VMD[59,60]) of the Zundel cation from our PIMD simulations.

distance $x_{eq}$ in eq 27 was set to 1. The inverse temperature $\beta = 10$ and Trotter number $P = 64$ were used. The numerical estimators were computed using 100 000 normal mode PIMD steps per window ($dt = 0.01$) for a total of 21 trajectories.

Figure 2 shows again the superior convergence behavior of the virial estimator. The corresponding integrated quantum free energy changes are $-0.204$ for the primitive estimator (eq 8) and $-0.205$ for the virial estimator (eq 10). These values are in excellent agreement with the exact result ($-0.204$). The average error bar for the primitive estimator is 0.012, which is 5 times larger than the error for the virial estimator (0.002). We note that the overall shape of this derivative is similar to the corresponding single harmonic oscillator in Figure 1, suggesting that the first term in eq 29 is dominating.

In more general situations (heterogeneous system A–B···C and/or anharmonic potentials), all vibrational modes are coupled and are therefore expected to contribute to the isotope effect. Such scenarios could be modeled using an antiparallel double Morse or Lennard-Jones potential.

*3. Isotope Effect on the Zundel Cation from First Principles.* The Zundel cation ($H_5O_2^+$), which consists of a proton shared by two water molecules (see inset of Figure 3), is a model for strongly hydrogen-bonded systems and is of paramount importance in condensed phases and in enzymatic reactions. Using a combination of centroid molecular dynamics (CMD) and empirical valence bond (EVB) theory, Schmitt and Voth found that the deuterated Zundel complex in liquid water exhibits a free energy barrier increase of about 0.2 kcal/mol compared to the all-hydrogen complex, leading to a decrease of 1.4 times in the hopping rate of proton transfer.[47] Here, we investigated the stability of $H_5O_2^+$ versus its deuterated counterpart ($D_5O_2^+$) with *ab initio* path-integral molecular dynamics. The virial estimator of eq 13 was implemented in the density functional theory code CPMD (version 3.13.2)[48] for the PIMD thermodynamic integration. The details of the *ab initio* PIMD simulation are as follows. The exchange and correlation were approximated with the BLYP[49−51] functional. Numerical norm-conserving Trouillier–Martins pseudopotentials[52] were used to represent core electronic orbitals. These pseudopotentials were augmented with a dispersion energy correction as described in refs 53 and 54. A kinetic energy cutoff of 100 Ry was employed for the plane wave basis set expansion of the electronic valence orbitals. The Zundel cation was placed in a cubic box of volume $12^3$ Å$^3$. Isolated molecule boundary conditions were imposed using the algorithm of Martyna and Tuckerman.[55] The Car–Parrinello (CP) algorithm[5] was used with a fictitious electronic mass of 340 au and a time step of 2 au ($\approx 0.0484$) fs. These parameters were carefully chosen to ensure the stability of the simulations. With these parameters, the adiabaticity of the CP scheme was preserved, and the fictitious electronic kinetic energy was stable and always less than 0.003 au. Normal mode variables and 32 beads were used for the discrete path integral. Massive Nosé–Hoover chain thermostats[56] of length 3 and a characteristic frequency of 4400 cm$^{-1}$ were employed to ensure adequate canonical sampling at 300 K.[57] A total of five independent PIMD trajectories, each of 3.7 ps, were simulated corresponding to ($\lambda = 0$, 0.25, 0.5, 0.75, 1.0) values in the thermodynamic integration. A linear path was used for the mass interpolation between hydrogens and deuteriums. We checked the correctness of our CPMD implementation by first simulating the hydrogen molecule at 300 K using a setup similar to that described above. An estimate of the free energy change of the isotopic transformation $H_2 \rightarrow D_2$ is available by other quantum chemistry methods ($\approx -2$ kcal/mol) and was reproduced numerically (see details in the Appendix).

Figure 3 shows the virial TI-PIMD results for the change in mass of the hydrogen atoms to deuterium in the Zundel cation at 300 K. The system is expected to be more stable upon deuteration mostly due to the suppression in the zero point energy. The predicted change in quantum free energy, $\Delta F = -0.01755$ au ($-11.02$ kcal/mol), was obtained from the numerical integration of the TI-PIMD curve in Figure 3 using Simpson's rule. This value agrees well with the thermochemistry result ($-10.16$ kcal/mol) computed using the quantum chemistry code Firefly[58] at the BLYP/6-311++G** level under the harmonic and rigid rotor approximation. We remark that converging error bars below 1 kcal/mol (our average standard deviation was 1.36 kcal/mol) would require very long PIMD simulations and extremely well equilibrated path integral configurations. Finally, we mention that

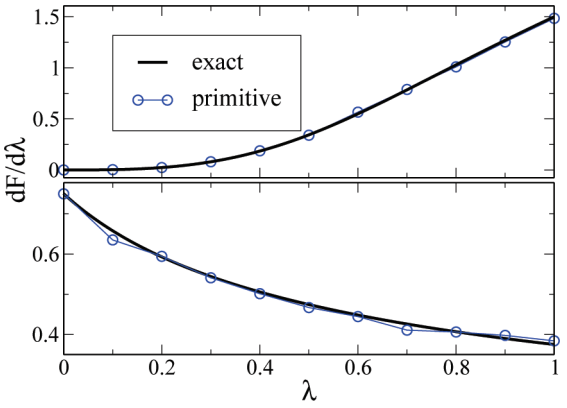


**Figure 4.** TI-PIMD: Free energy derivative with respect to the force constant ($\kappa_i = 1 \rightarrow \kappa_f = 4$) in the harmonic oscillator. The inverse temperature was $\beta = 10$. Natural units are used. PIMD (with $P = 64$) was conducted for 400 000 steps. Bottom: linear path $\kappa(\lambda) = \lambda\kappa_f + (1 - \lambda)\kappa_i$. Top: nonlinear path $\kappa(\lambda) = \kappa_f\lambda^4 + (1 - \lambda^4)\kappa_i$.

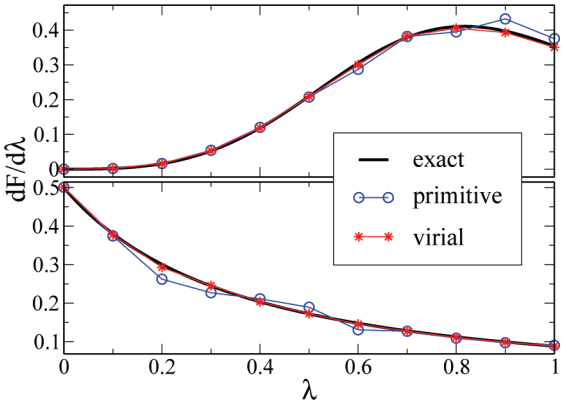


**Figure 5.** TI-PIMD: First order free energy derivative with respect to simultaneous changes in force constant and mass in the harmonic oscillator: $m_i = 1$, $\kappa_i = 1 \rightarrow m_f = 2$, $\kappa_f = 4$. Bottom: linear interpolation, eq 6. Top: nonlinear path, quartic order in $\lambda$. A total of 400 000 PIMD (64 beads) steps were simulated at inverse temperature $\beta = 10$.

computing changes in quantum kinetic energy due to isotope transformation may be useful to compare with deep inelastic neutron scattering experiments.

**B. TI-PIMD: Changes in Potential.** Free energy differences due to variations in the force constant $\kappa$ of the harmonic potential $V(x) = \kappa x^2/2$ were computed using TI-PIMD. Figure 4 compares numerical values from eq 16 with analytical results according to eq 15 for transformation $\kappa_i = 1 \rightarrow \kappa_f = 4$ in the force constant. Two interpolation paths have been investigated, first an interpolation of the force constant that is linear in $\lambda$ (bottom), then an interpolation that is quartic in $\lambda$, $\kappa(\lambda) = \kappa_f\lambda^4 + (1 - \lambda^4)\kappa_i$ (top). In analogy to the results for changing masses, the results displayed in Figure 4 suggest that also for changing potentials a quartic interpolation path leads to superior convergence behavior. The typical error bars for the primitive estimator are 0.0039 (quartic) and 0.0044 (linear), while for the virial they are 0.0022 (linear) and 0.0022 (quartic). However, as in the previous example and not without irony, due to statistical errors, the numerical integrated result of the less converged linear interpolation (0.5028) compares fortuitously better to the exact value (0.5) than the result of the more converged quartic interpolation (0.5214).

**Table 1. Morse Potential $V(x) = D[1 - e^{-a(x-x_{eq})}]^2$ Parameters Describing the Diatomic Molecules[a]**

| Morse parameters | $H_2$ | HCl |
|---|---|---|
| $r_{eq}$ | 1.4014 | 2.4086 |
| $D$ | 0.1745 | 0.1676 |
| $a$ | 1.0213 | 0.9684 |
| $\mu$ | 918.0584 | 1785.6425 |
| $\omega$ | 0.0199 | 0.013268 |

[a] Also shown is the reduced mass $\mu$ and the harmonic frequency $\omega = a((2D/\mu))^{1/2}$. All values are in atomic units. The thermal energy at 300 K is $k_BT = 0.00095$ au.

**C. TI-PIMD: Simultaneous Changes in Mass and Potential.** Using TI-PIMD, we have investigated simultaneous changes in mass and potential not only for the quantum harmonic oscillator, but also for transforming $H_2$ into HCl.

*1. Quantum Harmonic Oscillator.* Figure 5 shows the TI-PIMD results for virial and primitive estimators of the quantum free energy of the harmonic oscillator shown together with exact results. Linear (bottom) and quartic (top) interpolation functions in $\lambda$ (eq 7) were considered. As for the previous cases, the latter exhibits better convergence with the exception of the last two windows. After TI, the change in free energy according to the linear interpolation yields 0.211 for the primitive and 0.215 for the virial. The nonlinear interpolation slightly outperforms the linear interpolation when compared to the exact value given by eq 18 ($\Delta F = 0.207$): The primitive estimator predicts 0.206, whereas the virial leads to 0.203. The typical error bar for the primitive estimator is 0.0088, which is nearly 4 times larger than the error bar for the virial estimator (0.0022), both for the linear path. Similarly, for the quartic path, the average error bar for the primitive estimator is 0.066, while the error for the virial estimator is 0.0097.

We note that both derivatives bear more resemblance with the example discussed in the previous section where only the potential was changed, and not the mass. Furthermore, the effect of changing the mass in addition to changing the potential clearly introduces further curvature to the free energy derivative profile of the quartic interpolation. This raises the question of whether the increase in curvature leads to finer resolution requirements for the thermodynamic integration, thereby removing the advantage gained from better convergence of the sampled estimator.

*2. Transmutation of a Diatomic Molecule.* In this section, the alchemical transformation of molecular hydrogen into hydrogen chloride at 300 K is investigated using TI-PIMD. We approximate the internuclear potential of the diatomic molecules by a Morse potential:

$$V(x) = D[1 - e^{-a(x - x_{eq})}]^2$$

where $x_{eq}$ is the internuclear equilibrium distance, $D$ is the dissociation energy measured from the potential minimum, and $a$ is related to the curvature of the potential at the equilibrium position. Table 1 displays all of the Morse parameters employed for these calculations. The reduced mass and all of the parameters ($a$, $D$, $x_{eq}$) were linearly interpolated in $\lambda$ between initial ($H_2$) and final (HCl) states. The free energy difference was obtained using normal mode TI-PIMD ($P = 64$) and 4 million steps (dt = 0.1 au). Since at 300 K $H_2$ and HCl are mostly ground-state-dominated, we can approximate the

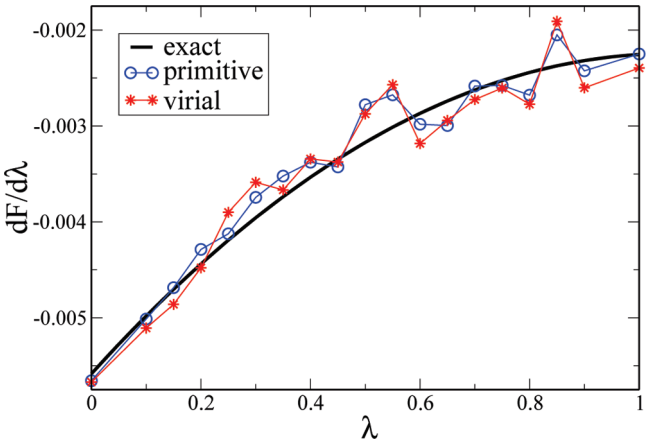


**Figure 6.** TI-PIMD: First-order free energy derivative with respect to simultaneous changes in Morse parameters and mass corresponding to the alchemical transformation of $H_2 \rightarrow$HCl at $T = 300$ K. Black: Exact analytical estimate using a harmonic approximation to the Morse potential. Blue: Primitive estimator. Red: Virial estimator. Atomic units are used.

Morse potential harmonically to get an estimate of the change in free energy,

$$\Delta F = F(\mathrm{HCl}) - F(\mathrm{H_2}) \approx k_B T \ln\left[\frac{\sinh(\beta\hbar\omega_{\mathrm{HCl}}/2)}{\sinh(\beta\hbar\omega_{\mathrm{H_2}}/2)}\right] = -0.003324 \text{ au} \quad (30)$$

TI-PIMD results (eq 17) for the primitive and virial estimators of the free energy derivative with respect to transforming the Morse potential of $H_2$ into HCl are presented in Figure 6, together with the exact result for the harmonically approximated Morse potential. In contrast to all the previous examples, and despite the large number of PIMD steps taken, the remaining statistical error (on the order of 0.00091 au) is still significant. This behavior is exarcerbated as approaching the end point ($\lambda = 1$) when the new physical parameters instantly appear where other interactions were previously present. Such numerical instabilities are commonly ameliorated in classical simulations by softening the underlying potential and making it $\lambda$-dependent. Nonetheless, the resulting integrated free energy differences, $-0.003326$ au and $-0.003320$ au, for the primitive and virial estimators, respectively, both compare very favorably to the analytical estimate in eq 30 ($-0.003324$ au).

The shape of Figure 6 is reminiscent of Figure 1 (bottom), suggesting that the linear transformation of $H_2$ into HCl is dominated by the change in mass, rather than by the change in potential.

**D. Mass Transformation in the Harmonic Oscillator via PT-PIMD.** In this section, we use PT-PIMD to recompute the free energy difference of the harmonic oscillator, which was already investigated with TI-PIMD in section III.A.1. Figure 7 illustrates the convergence of the running average of the free energy difference with the number of PT-PIMD steps. Two cases are considered: a "large" change in mass, $m_i = 1 \rightarrow m_f = 2$ (top), and a "small" one, $m_i = 1 \rightarrow m_f = 1.5$ (bottom). Unlike TI-PIMD, these results were obtained using a single, albeit substantially longer, PT-PIMD simulation. As one would expect, the smaller change in mass (bottom) converges more rapidly to the analytical

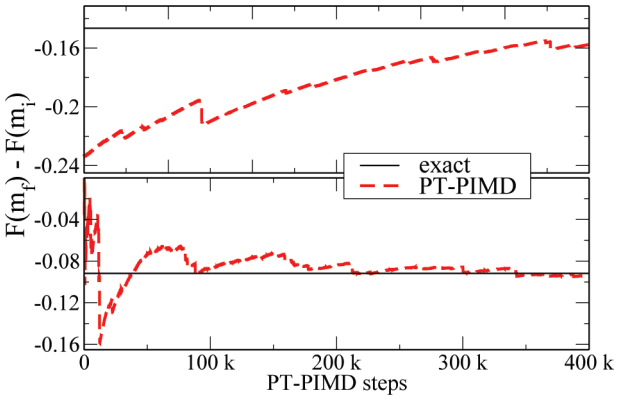


**Figure 7.** PT-PIMD: Running average of the free energy difference for an isotope transformation in the harmonic potential computed using perturbation theory, eq 19. Top: $m_i = 1 \rightarrow m_f = 2$. Bottom: $m_i = 1 \rightarrow m_f = 1.5$. Natural units are used. Red dashes: PT-PIMD using eq 19. Black solid: Exact result. The inverse temperature is $\beta = 10$, and the Trotter number is $P = 64$.

reference result than the larger (top) for which the number of steps (400 000) is clearly insufficient. The error bar for the latter (0.0424) is more than 5 times larger than the error for the "small" transformation (0.0075), indicating that the calculation suffers from slow convergence. Clearly, for the large transformation, the end point states are too disparate (there is little overlap) and PT methods do not converge. A remedy for such a situation would be to perform a series of PT calculations on mutually overlapping intermediate states between the end points.

The sawtooth pattern in Figure 7 is a common feauture when averaging exponentials.[61] The spikes indicate that the averages are dominated by rare events. As mentioned before, a possible remedy to dampen these sudden fluctuations in $\Delta F$ consists of further reducing the interval, i.e., to use a multistage PT-PIMD approach. Despite the abrupt initial transient, the running average for the small mass transformation converges nicely toward the exact value as the simulation goes on, Figure 7 (bottom).

**E. PT-PIMD: Effect of a Uniform Static Electric Field.** Perturbation theory can also be used to compute electric polarizabilities, which are very important in understanding not only weakly bonded intermolecular and long-range interactions[62] but also the ferro/paraelectric transition in some hydrogen-bonded materials. In this context, Sebastiani and Srinivasan recently investigated the phase transition in potassium dihydrogen phosphate (KDP), which is known to exhibit strong nuclear quantum effects.[63] In this subsection, we suggest an estimator that could be used to further characterize these phase transitions. The change in free energy of a single particle of charge $q$ due to a uniform static electric field of strength $E = |\mathbf{E}|$ is

$$\Delta F = F(E) - F(0) = -\frac{1}{\beta}\ln\langle\exp[-\beta(q\mathbf{E}\cdot\bar{\mathbf{r}})]\rangle_{E=0} \quad (31)$$

where $\bar{\mathbf{r}} = \sum_{s=1}^{P}\mathbf{r}_s/P$ is the centroid of the ring polymer. For bulk systems, eq 31 must be modified to account for periodicity.

Figure 8 shows the numerical results for the running average of the free energy difference in the harmonic oscillator with and without an homogeneous static electric field of varying strength. When the magnitude of the electric field is not too large (bottom), the converged estimate of the corresponding change

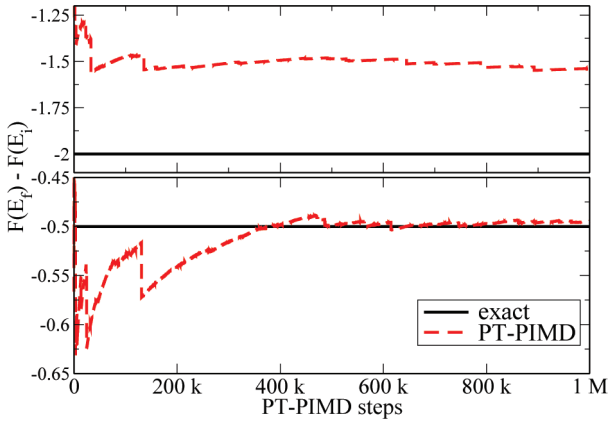


**Figure 8.** PT-PIMD: Running averages of the free energy difference of the harmonic potential $m\omega x^2/2$ due changes in the static external electric field. Top: $E_i = 0 \rightarrow E_f = 2$. Bottom: $E_i = 0 \rightarrow E_f = 1$ (natural units are used). Red dashes: PT-PIMD using eq 31. Black solid: Exact. The inverse temperature is $\beta = 10$, and the Trotter number is $P = 64$.

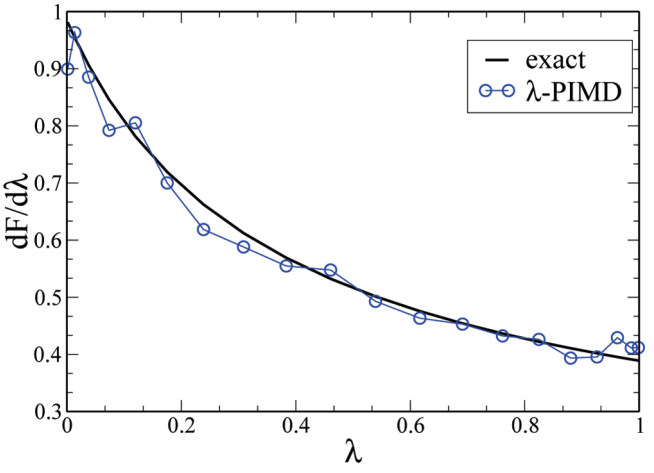


**Figure 9.** $\lambda$-PIMD: First order free energy derivative versus $\lambda$ for harmonic potential $\kappa(\lambda)x^2/2$ at $T = 0.5$. Black solid: Exact, eq 16. Blue solid: Numerical using $\lambda$-PIMD, i.e. $\langle \Delta V \rangle_\lambda$. One million PIMD steps were accumulated with $P = 64$ and dt = 0.01 at the inverse temperature $\beta = 10$ (natural units are used).

in free energy (−0.4947) is in very good agreement with the exact value (−0.5), as given by $\Delta F = -((q^2E^2)/(2m\omega^2))$, where $q = 1$ is the charge and $m = \omega = 1$. However, when the applied electric field is twice as large (top), the estimator eq 31 converges to the wrong value (−1.5411), which deviates significantly from the exact value of −1.1250. We note that in both cases the error bar is similar (0.042 for large and 0.031 for small transformation), so it is not a reliable indicator of the accuracy of the final result. This example, as with the mass transformation in the previous section, highlights the limitations of the PT-PIMD approach.

**F. Changes in Potential of the Quantum Harmonic Oscillator Revisited: λ-PIMD.** In this final section, we revisit changes of the force constant in the quantum harmonic oscillator in analogy to section III.B. Instead of TI-PIMD, however, we used $\lambda$-PIMD. Specifically, we numerically tested eq 22 for the same linear change in force constant as in section III.B, i.e., $V(x) = \kappa(\lambda)x^2/2$, with $\kappa(\lambda) = \lambda\kappa_f + (1 - \lambda)\kappa_i$ and $\kappa_i = 1$ and $\kappa_f = 4$.

The corresponding free energy change for this transformation is obtained from numerical integration of the $\lambda$-derivative over $\lambda$,

$$\Delta F = F_f - F_i = \int_0^1 d\lambda \langle \Delta V \rangle_\lambda = \int_0^{\pi/2} d\theta \langle \Delta V \rangle_\theta \sin(2\theta) \tag{32}$$

where $\Delta V = (1/P)(\kappa_f - \kappa_i)\Sigma_{s=1}^P x_s^2/2$ for the harmonic oscillator. Numerical results for the first-order derivative of the free energy with respect to $\lambda$ at $T = 0.5$ are shown in Figure 9. The integrated predicted free energy difference amounts to 0.555 (average error bar 0.009) after 1 million PIMD steps ($P = 64$ beads). This value is in good agreement with the exact value for the harmonic oscillator (0.5635).

## IV. CONCLUSIONS

New path integral estimators have been formulated for the calculation of free energy changes due to transformations in the mass and/or potential and changes in the external electric field. Perturbation theory and $\lambda$-dynamics are found to be useful alternatives to thermodynamic integration for the prediction of the free energy differences of alchemical transformations from a single simulation. All free energy methods were implemented within path integral molecular dynamics (TI-PIMD, PT-PIMD, $\lambda$-PIMD) and applied to various solvable model systems, including the single and double harmonic oscillator and the Morse potential. The methods introduced are rather general and valid for any force-field-based PIMD calculation of an alchemical change. The estimators were also applied to investigate the isotope transformation in the Zundel cation using *ab initio* path integral molecular dynamics.

Future work will deal with the extension of these ideas to *ab initio* PIMD for the simulation of more realistic systems. For example, it will be interesting to investigate a complete change of identity of a molecule from first principles. While this is feasible in force-field-based PIMD calculations, it is more challenging in *ab initio* PIMD due to the use of pseudopotentials, or atomic basis functions. As an example, we would first investigate the isoelectronic transformation of the Zundel cation to the protonated water−ammonia complex cation. Finally, it will also be interesting to devise a way to predict changes in the quantum free energy barrier and not just on traced quantitites such as the overall free energy.

## APPENDIX A. ISOTOPE EFFECT ON INTERNAL ENERGY

The formalism presented so far can be easily extended to compute changes of any thermodynamic quantity due to alchemical changes. In this section, we derive an estimator for the change of internal energy corresponding to the isotope effect.

The path integral expression for the internal energy according to the *primitive* form is

$$\langle E \rangle = \lim_{P \rightarrow \infty} \frac{C}{Z} \int d\mathbf{r}_1 \ldots d\mathbf{r}_P \, \varepsilon_{\text{prim}}(\lambda) \, e^{-\beta\Phi(\{r\};\lambda)} \tag{A1}$$

where $C(\lambda) = ((m(\lambda)P)/(2\pi\beta\hbar^2))^{3P/2}$ is a normalization constant,

$$\Phi(\{\mathbf{r}\};\lambda) = \sum_{s=1}^{P} \left[ \frac{m(\lambda)}{2} \omega_P^2 (\mathbf{r}_s - \mathbf{r}_{s+1})^2 + \frac{V(\mathbf{r}_s)}{P} \right]$$

is the effective path integral potential and

$$\varepsilon_{\mathrm{prim}}(\lambda) = \frac{3P}{2\beta} + \sum_{s=1}^{P}\left[\frac{V(\mathbf{r}_s)}{P} - \frac{m(\lambda)}{2}\omega_P^2(\mathbf{r}_s - \mathbf{r}_{s+1})^2\right]$$

is the so-called primitive energy estimator.

Taking the derivative of eq A1 with respect to the parameter $\lambda$ gives

$$\frac{\mathrm{d}E(\lambda)}{\mathrm{d}\lambda} = \langle \partial_\lambda \varepsilon_{\mathrm{prim}}\rangle + \beta[\langle \varepsilon_{\mathrm{prim}}\rangle\langle \partial_\lambda \Phi\rangle - \langle (\partial_\lambda \Phi)\varepsilon_{\mathrm{prim}}\rangle] \quad \text{(A2)}$$

where we note that $\partial_\lambda \varepsilon_{\mathrm{prim}} = -\partial_\lambda \Phi$. Rearranging into an expression more amenable for numerical computation gives

$$\frac{\mathrm{d}E(\lambda)}{\mathrm{d}\lambda} = \left(\frac{\partial_\lambda m}{m}\right)\{\langle \mathcal{J}\rangle + \beta[\langle \mathcal{J}\varepsilon_{\mathrm{prim}}\rangle - \langle \varepsilon_{\mathrm{prim}}\rangle\langle \mathcal{J}\rangle]\} \quad \text{(A3)}$$

where $\mathcal{J} = -(m/2)\omega_P^2\sum_{s=1}^{P}(\mathbf{r}_s - \mathbf{r}_{s+1})^2$.

As a test, we can compare $E(m_{\mathrm{f}}) - E(m_{\mathrm{i}}) = \int_0^1 \mathrm{d}\lambda(\mathrm{d}E^{\mathrm{prim}}/\mathrm{d}\lambda)_\lambda$ to the exact expression for a model system. The analytical expression for the internal energy of the harmonic oscillator is $U(\beta) = (\hbar\omega/2)\coth(\beta\hbar\omega/2)$. Thus, the analytic change of internal energy is

$$\Delta E(\beta) = \frac{\hbar}{2}[\omega_{\mathrm{f}}\coth(\beta\hbar\omega_{\mathrm{f}}/2) - \omega_{\mathrm{i}}\coth(\beta\hbar\omega_{\mathrm{i}}/2)] \quad \text{(A4)}$$

where the oscillator frequencies are $\omega_{\mathrm{i(f)}} = (\kappa/m_{\mathrm{i(f)}})^{1/2}$. The exact analytical expression for the first derivative of the total energy with respect to $\lambda$ is for the harmonic oscillator is

$$\frac{\mathrm{d}E(\lambda)}{\mathrm{d}\lambda} = \frac{\hbar\omega\partial_\lambda m}{2m}\left[\frac{(\beta\hbar\omega/2)}{\sinh^2(\beta\hbar\omega/2)} - \frac{\coth(\beta\hbar\omega/2)}{2}\right] \quad \text{(A5)}$$

Finally, the derivative of the entropy with respect to $\lambda$ is simply $\partial_\lambda S = (\partial_\lambda(U - F))/T$. Thus, the isotope effect on the entropy can also be estimated.

## APPENDIX B. SECOND-ORDER DERIVATIVES

In 1994, Smith and van Gunsteren published a paper on predictions of classical free energy differences using only the information of free energy derivatives computed at the initial state.[64] On the basis of that work, we have investigated extrapolations of the free energy from the initial state within PIMD.

From Figure 1, it is clear that the change in free energy is not linear. Using only the first-order derivative information at the initial state ($\lambda = 0$) leads to inaccurate predictions of the overall free energy difference. Thus, it is necessary to include second- and higher-order derivatives. Expanding the free energy difference in a Taylor series to second order,

$$\Delta F_{\mathrm{if}} \approx (\lambda_{\mathrm{f}} - \lambda_{\mathrm{i}})\left(\frac{\mathrm{d}F}{\mathrm{d}\lambda}\right)_{\lambda_{\mathrm{i}}} + \frac{1}{2!}(\lambda_{\mathrm{f}} - \lambda_{\mathrm{i}})^2\left(\frac{\mathrm{d}^2F}{\mathrm{d}\lambda^2}\right)_{\lambda_{\mathrm{i}}} \quad \text{(B1)}$$

we could in principle estimate the difference in free energies using only the information from the initial state. To accomplish this, higher derivatives of the free energy with respect to parameter $\lambda$ are needed. The primitive estimator for the second derivative of the free energy with respect to mass reads

$$\frac{\mathrm{d}^2F}{\mathrm{d}\lambda^2} = \left\langle (\partial_\lambda m + \partial_\lambda^2 m)\frac{\omega_P^2}{2}\mathcal{J} - \frac{3P}{2\beta}\left(\frac{\partial_\lambda m}{m} + \frac{\partial_\lambda^2 m}{m} + \frac{(\partial_\lambda m)^2}{m^2}\right)\right\rangle + \frac{1}{\beta}\left\langle \frac{P\partial_\lambda m}{2m} - \beta\frac{\partial_\lambda m}{2}\omega_P^2\mathcal{J}\right\rangle \quad \text{(B2)}$$

where the sum $\mathcal{J}$ denotes $\sum_{s=1}^{P}(\mathbf{r}_s - \mathbf{r}_{s+1})^2$ with $\mathbf{r}_{P+1} = \mathbf{r}_1$. Note that the last term in this expression vanishes if we assume a linear interpolation path for the mass ($\partial_\lambda^2 m = 0$). This expression is problematic because it involves a significant cancellation of terms. The numerical results can be checked with the analytical expression for the second-order derivative of the quantum harmonic oscillator with respect to changes in mass

$$\left(\frac{\mathrm{d}^2F}{\mathrm{d}\lambda^2}\right) = -\left(\frac{1}{2}\right)\left(\frac{\partial_\lambda m}{m}\right)^2\frac{\hbar\omega}{2} \times \left[\frac{\beta\hbar\omega/2}{2\sinh^2(\beta\omega\hbar/2)} - \frac{3}{2\tanh(\beta\hbar\omega/2)}\right] - \left(\frac{\hbar\omega}{4}\right)\left(\frac{\partial_\lambda^2 m}{m}\right)\coth(\beta\hbar\omega/2) \quad \text{(B3)}$$

Our numerical tests (not shown) suggest, however, that convergence of higher order derivatives is difficult to achieve.

Fortunately, an empirical way to circumvent this nonlinearity issue was recently presented in the context of potential energy differences by one of us.[65] In that work, higher-order terms in the Taylor expansion were absorbed into a correction factor accompanying the first derivative

$$\Delta F_{\mathrm{if}} = (\lambda_{\mathrm{f}} - \lambda_{\mathrm{i}})\left(\frac{\mathrm{d}F}{\mathrm{d}\lambda}\right)_{\lambda_{\mathrm{i}}} + \frac{1}{2!}(\lambda_{\mathrm{f}} - \lambda_{\mathrm{i}})^2\left(\frac{\mathrm{d}^2F}{\mathrm{d}\lambda^2}\right)_{\lambda_{\mathrm{i}}} + \ldots \approx (\lambda_{\mathrm{f}} - \lambda_{\mathrm{i}})\mathcal{C}(\lambda_{\mathrm{i}})\left(\frac{\mathrm{d}F}{\mathrm{d}\lambda}\right)_{\lambda_{\mathrm{i}}} \quad \text{(B4)}$$

where $\mathcal{C}$ is a renormalization factor that accounts for the nonlinear terms. The hope is to calculate $\mathcal{C}$ numerically for a reference system (say $\mathcal{C}^{\mathrm{ref}}$) and then use this coefficient (without modification) to estimate the isotope free energy change in other systems. This approach is admittedly somewhat *ad hoc* because it assumes transferability of $\mathcal{C}$ to other systems. In other words, the effect of higher-order terms is similar in the reference system and in the system of interest. Nonetheless, there might be some justification to this method if the leading term in the Taylor expansion is the first derivative (and the higher terms are small and cancel among each other). Work is underway to find numerically stable *virial* expressions for the second-order derivatives.

## APPENDIX C. CONVERGENCE WITH THE NUMBER OF STEPS

We have carried out a convergence test on the free energy change versus the number of steps for the isotope transformation in the harmonic oscillator. Figure 10 shows that the convergence is achieved well within the first 400 000 steps for this toy model.

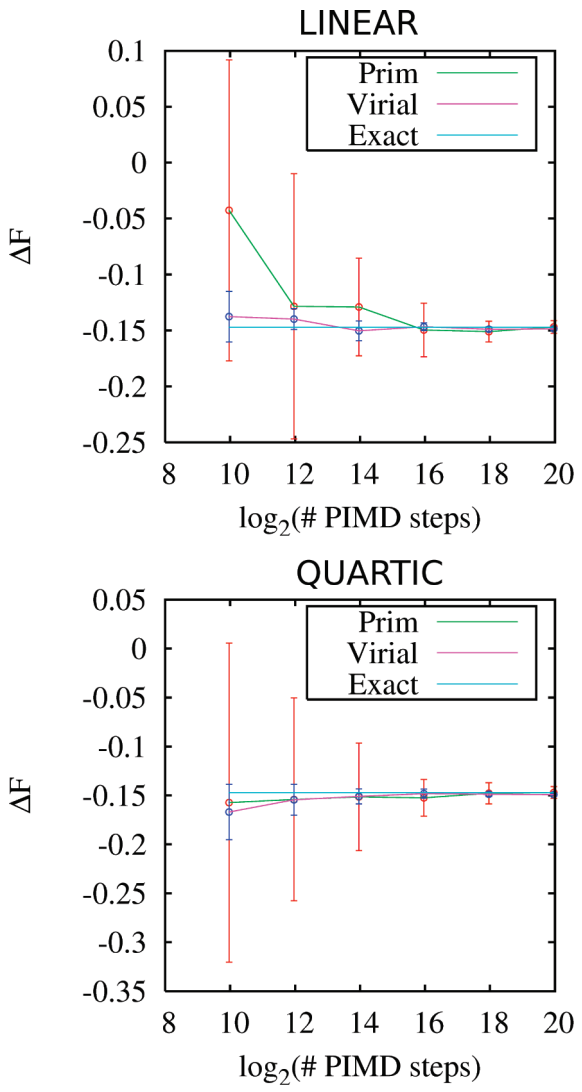


**Figure 10.** Convergence of the free energy difference (in natural units) with the number of PIMD steps for the mass transformation in the harmonic oscillator at $\beta\hbar\omega = 10$. Top: linear path. Bottom: quartic path.

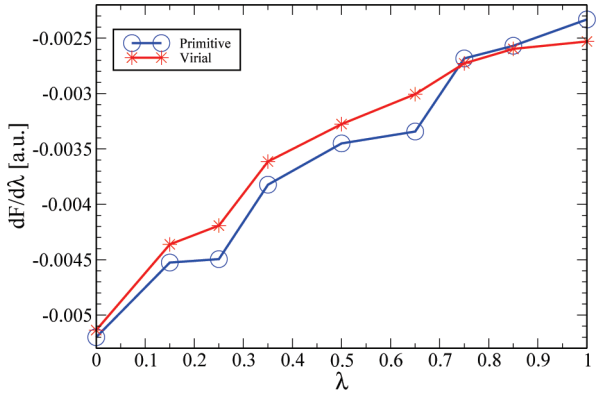


**Figure 11.** Numerical results (in au) from the *ab initio* TI-PIMD simulation ($P = 8$) of the hydrogen molecule at 300 K.

The figure also confirms that the quartic path leads to a faster convergence than the linear path.

## APPENDIX D. ISOTOPE TRANSFORMATION ON THE HYDROGEN MOLECULE: *AB INITIO* APPROACH

To test our numerical implementation in CPMD, we simulated the $H_2 \rightarrow D_2$ isotope transformation. The hydrogen molecule was placed in a cubic box of volume $9^3$ Å$^3$. The kinetic energy cutoff for the plane wave expansion of electronic orbitals was 75 Ry. Eight beads were used for the path integral discretization. The rest of the parameters were identical to the ones employed for the Zundel ion. The results of this test are shown in Figure 11. The predicted change in free energy is $\Delta F = -0.00358$ au ($-2.25$ kcal/mol) for the primitive and $\Delta F = -0.00345$ au ($-2.17$ kcal/mol) for the virial estimator. These values are in good agreement with the values from other quantum chemistry methods ($\approx -2$ kcal/mol). The average standard deviation was 1.72 kcal/mol for the primitive and 0.56 kcal/mol for the virial estimator. The perturbation theory approach of eq 19 yields a value of $-0.00375$ au ($-2.36$ kcal/mol), which agrees very well with TI-PIMD results.

## AUTHOR INFORMATION

### Corresponding Authors

*E-mail: ap1484@nyu.edu (A.P.), anatole@alcf.anl.gov (O.A.v.L.). Website: http://www.alcf.anl.gov/~anatole (O.A.v.L.).

## ACKNOWLEDGMENT

We gratefully acknowledge stimulating discussions with Mark E. Tuckerman. We thank the hospitality of UCLA/IPAM where this work was completed and the City University of New York (CUNY) High Performance Computing Center (HPCC) at College of Staten Island for computational resources. A.P. is grateful for support from the Horizon fellowship at New York University, the SNL's LDRD project No. 117866, and the SNL's summer student internship program at the Computer Science Research Institute. O.A.v.L. acknowledges SNL's LDRD Truman project No. 120209. SNL is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000. This research also used resources of the Argonne Leadership Computing Facility at Argonne National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under contract DE-AC02-06CH11357.

## REFERENCES

(1) Marx, D.; Parrinello, M. *J. Chem. Phys.* **1996**, *104*, 4077–4082.
(2) Tuckerman, M. E.; Marx, D.; Klein, M. L.; Parrinello, M. *J. Chem. Phys.* **1996**, *104*, 5579–5588.
(3) Feynman, R. P. *Rev. Mod. Phys.* **1948**, *20*, 367–387.
(4) Feynman, R. P.; Hibbs, A. R. *Quantum Mechanics and Path Integrals*; McGraw-Hill: New-York, 1965.
(5) Car, R.; Parrinello, M. *Phys. Rev. Lett.* **1985**, *55*, 2471–2474.
(6) Iftimie, R.; Minary, P.; Tuckerman, M. E. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6654–6659.
(7) Tuckerman, M. E.; Marx, D. *Phys. Rev. Lett.* **2001**, *86*, 4946–4949.
(8) Miura, S.; Tuckerman, M. E.; Klein, M. L. *J. Chem. Phys.* **1998**, *109*, 5290–5299.
(9) Morrone, J. A.; Car, R. *Phys. Rev. Lett.* **2008**, *101*, 017801.
(10) Tuckerman, M. E.; Marx, D.; Klein, M. L.; Parrinello, M. *Science* **1997**, *275*, 817–820.

(11) Marx, D.; Tuckerman, M. E.; Hutter, J.; Parrinello, M. *Nature* **1999**, *397*, 601–604.
(12) Tuckerman, M. E.; Marx, D.; Parrinello, M. *Nature* **2002**, *417*, 925–929.
(13) Li, X.-Z.; Probert, M. I. J.; Alavi, A.; Michaelides, A. *Phys. Rev. Lett.* **2010**, *104*, 066102.
(14) Davidson, E. R. M.; Alavi, A.; Michaelides, A. *Phys. Rev. B* **2010**, *81*, 153410.
(15) Pu, J.; Gao, J.; Truhlar, D. G. *Chem. Rev.* **2006**, *106*, 3140–3169.
(16) Hwang, J. K.; Warshel, A. *J. Am. Chem. Soc.* **1996**, *118*, 11745.
(17) Pérez, A.; Tuckerman, M. E.; Hjalmarson, H. P.; von Lilienfeld, O. A. *J. Am. Chem. Soc.* **2010**, *132*, 11510–11515.
(18) von Lilienfeld, O. A.; Lins, R.; Rothlisberger, U. *Phys. Rev. Lett.* **2005**, *95*, 153002.
(19) Wang, M.; Hu, X.; Beratan, D. N.; Yang, W. *J. Am. Chem. Soc.* **2006**, *128*, 3228.
(20) Marcon, V.; von Lilienfeld, O. A.; Andrienko, D. *J. Chem. Phys.* **2007**, *127*, 064305.
(21) Rinderspacher, B. C.; Andzelm, J.; Rawlett, A.; Dougherty, J.; Beratan, D. N.; Yang, W. *J. Chem. Theory Comput.* **2009**, *5*, 3321.
(22) Froemming, N. S.; Henkelman, G. *J. Chem. Phys.* **2009**, *131*, 234103.
(23) Sheppard, D.; Henkelman, G.; von Lilienfeld, O. A. *J. Chem. Phys.* **2010**, *133*, 084104.
(24) van Gunsteren, W. F.; Daura, X.; Mark, A. E. *Helv. Chim. Acta* **2002**, *85*, 3113.
(25) Kirkwood, J. G. *J. Chem. Phys.* **1935**, *3*, 300–313.
(26) Oostenbrink, C. *J. Comput. Chem.* **2009**, *30*, 212.
(27) von Lilienfeld, O. A.; Tuckerman, M. E. *J. Chem. Phys.* **2006**, *125*, 154104.
(28) von Lilienfeld, O. A.; Tuckerman, M. E. *J. Chem. Theory Comput.* **2007**, *3*, 1083.
(29) Leung, K.; Rempe, S. B.; von Lilienfeld, O. A. *J. Chem. Phys.* **2009**, *130*, 204507.
(30) Alfè, D.; Gillan, M. J.; Price, G. D. *Nature* **2000**, *405*, 172.
(31) Zimmermann, T.; Vanicek, J. *J. Chem. Phys.* **2009**, *131*, 024111.
(32) Ramírez, R.; Herrero, C. P. *J. Chem. Phys.* **2010**, *133*, 144511.
(33) de la Peña, L. H.; Peslherbe, G. H. *J. Phys. Chem. B* **2010**, *114* (16), 5404–5411.
(34) Tuckerman, M. E. *Statistical mechanics: Theory and molecular simulation*; Oxford University Press: Oxford, U. K., 2010.
(35) Chandler, D.; Wolynes, P. G. *J. Chem. Phys.* **1981**, *74*, 4078–4095.
(36) Zamalin, V. M.; Norman, G. E. *USSR Comp. Math. Math. Phys.* **1973**, *13* (2), 408–420.
(37) Tuckerman, M.; Berne, B.; Martyna, G.; Klein, M. *J. Chem. Phys.* **1993**, *99*, 2796–2808.
(38) *Free Energy Calculations: Theory and Applications in Chemistry and Biology*; Chipot, C., Pohorille, A., Eds.; Springer: New York, 2007; Springer Series in Chemical Physics, Vol. 86.
(39) Frenkel, D.; Smit, B. *Understanding Molecular Simulation*; Academic Press: New York, 2002.
(40) Herman, M. F.; Bruskin, E. J.; Berne, B. J. *J. Chem. Phys.* **1982**, *76*, 5150.
(41) Zwanzig, R. W. *J. Chem. Phys.* **1954**, *22*, 1420–1426.
(42) Liu, Z.; Berne, B. J. *J. Chem. Phys.* **1993**, *99*, 6071–6077.
(43) Kong, X.; Brooks, C. L., III. *J. Chem. Phys.* **1996**, *105*, 2414–2423.
(44) Guo, Z.; Brooks, C. L.; Kong, X. *J. Phys. Chem. B* **1998**, *102*, 2032–2036.
(45) Bitetti-Putzer, R.; Yang, W.; Karplus, M. *Chem. Phys. Lett.* **2003**, *377*, 633–641.
(46) Abrams, J. B.; Rosso, L.; Tuckerman, M. E. *J. Chem. Phys.* **2006**, *125*, 074115.
(47) Schmitt, U. W.; Voth, G. A. *Chem. Phys. Lett.* **2000**, *329*, 36–41.
(48) *CPMD*, version 3.13.2; IBM Corp: Armonk, New York, 1990−2008; MPI für Festkörperforschung Stuttgart: Stuttgart, Germany, 1997−2001. http://www.cpmd.org/ (accessed June 2011).
(49) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098–3100.
(50) Colle, R.; Salvetti, D. *Theor. Chim. Acta* **1988**, *37*, 329–334.
(51) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785–789.
(52) Troullier, N.; Martins, J. L. *Phys. Rev. B* **1991**, *43*, 1993–2006.
(53) von Lilienfeld, O. A.; Tavernelli, I.; Rothlisberger, U.; Sebastiani, D. *Phys. Rev. Lett.* **2004**, *93*, 153004.
(54) Lin, I.-C.; Coutinho-Neto, M. D.; Felsenheimer, C.; von Lilienfeld, O. A.; Tavernelli, I.; Rothlisberger, U. *Phys. Rev. B* **2007**, *75*, 205131.
(55) Martyna, G. J.; Tuckerman, M. E. *J. Chem. Phys.* **1999**, *110*, 2810.
(56) Martyna, G. J.; Klein, M. L.; Tuckerman, M. E. *J. Chem. Phys.* **1992**, *97*, 2635–2643.
(57) *Quantum Simulations of Complex Many-Body Systems: from theory to algorithms*; Grotendorst, J., Marx, D., Muramatsu, A., Eds.; John von Neumann Institute for Computing: Princeton, NJ, 2002.
(58) Granovsky, A. A. Firefly, version 7.1.G. http://classic.chem.msu.su/gran/firefly/index.html (accessed June 2011).
(59) Kohlmeyer, A. VMD scripts for visualization of CPMD data. http://klein-group.icms.temple.edu/cpmd-vmd/files.html (accessed June 2011).
(60) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33–38.
(61) Pohorille, A.; Jarzynski, C.; Chipot, C. *J. Phys. Chem. B* **2010**, *114*, 10235–10253.
(62) French, R. H.; Parsegian, V. A.; Podgornik, R.; Rajter, R. F.; et al. *Rev. Mod. Phys.* **2010**, *82*, 1887.
(63) Srinivasan, V.; Sebastiani, D. *J. Phys. Chem. C* **2011**, *115*, 12631.
(64) Smith, P. E.; van Gunsteren, W. F. *J. Chem. Phys.* **1994**, *100*, 577–585.
(65) von Lilienfeld, O. A. *J. Chem. Phys.* **2009**, *131*, 164102.