

PREDICTING ISOTOPIC FRACTIONATION IN WATER FROM QUANTUM SIMULATIONS

DAVID GABRIEL SELASSIE

CONTENTS

Acronyms	2
1. Introduction	3
2. Theoretical Methods	6
2.1. Path Integral Molecular Dynamics	6
2.2. PIGLET	18
2.3. Fractionation Ratios	24
3. Simulation Results	26
3.1. Hydrogen Bonding Effects on Quantum Kinetic Energy	26
3.2. Sampling of the Dilute Deuterium Limit	30
3.3. Quantum Kinetic Energy Distribution at Liquid–Vapor Interfaces	35
3.4. Quantum Kinetic Energy Distribution in Isolated Water Clusters	38
4. Future Directions	41
References	42

ACRONYMS

QKE: quantum kinetic energy

MD: molecular dynamics

PBC: periodic boundary conditions

PIMD: path integral molecular dynamics

GLE: generalized Langevin equation

PI+GLE: path integral with generalized Langevin equation thermostat

PIGLET: PI+GLE on the normal modes of ring polymers

PRO: thermostat

1. INTRODUCTION

The second thing a middle-school science student learns after the very basics of the periodic table is the existence of isotopes, sets of atoms with the same electronic structure that only vary in the number of neutrons in their nucleus and thus the mass of the atom. Hydrogen’s (H) isotopes are the most famous, and have their own names: deuterium (D), with one more neutron, and tritium (T), with two more neutrons. Since their chemical behavior is identical to a first approximation, after their initial introduction, they are promptly forgotten. If that student continues their chemical education, they’ll see isotopes crop up in a handful of other places in their undergraduate curriculum: peak identification based on isotopic ratios in mass spectrometry, using deuterated solvents to mask signals in nuclear magnetic resonance spectroscopy, radioactive isotopic labeling and tracking of atoms in reactions, kinetic effects, and using the OD stretch as a spectroscopic probe in hydrogen bonded substances.

All of these uses still do not touch upon, and often assume, the fact that isotopes do not have identical chemical behavior. Each nucleus is still a quantum mechanical object, and thus manifests all the qualitative behaviors that discrete energy states entail: hydrogen atoms in a molecule have quantum kinetic energy (QKE) through zero-point energy, a minimum energy that can not be removed, and have tunneling behavior, where quantum particles can seemingly go “through” potential energy barriers to classically forbidden regions. Since the specifics of those qualitative quantum behaviors depend on the mass of the nucleus, there *is* an energetic and chemical difference between isotopes. Since lighter, colder, and more

confined particles show more quantum mechanical behavior, and hydrogen and its isotopes are the lightest of nuclei, hydrogen’s nuclear quantum effects will be the easiest to detect.

There are techniques that hinge on this sometimes subtle isotopic chemical difference. Isotopic fractionation is a phenomena where molecules containing different isotopes will phase segregate themselves at equilibrium. Geologists and geochemists use hydrogen–deuterium liquid–vapor fractionation ratios to aid in determining from what temperature, location, altitude, or environment a water sample was obtained [8]. Since hydrogen–deuterium isotopic fractionation is an equilibrium property, we need to be able to predict the free energy difference between hydrogen and deuterium existing in the same configuration.

In order to make theoretical predictions about fractionation, we require advanced simulation and calculation methods that can capture QKE. Classical molecular dynamics (MD) or Monte Carlo simulations can be used to find equilibrium properties but the nuclei are expressed as classical particles, and so show no free energy dependence on mass. Standard mixed quantum mechanical–molecular mechanical techniques, like Born–Oppenheimer *ab-initio* MD or Carr–Parrinello MD [2], although quantum is in the name, usually treat only the electrons quantum mechanically and use this quantum electronic potential to evolve *classical* nuclear trajectories, again missing out on nuclear QKE. One set of methods that does capture nuclear QKE is path integral molecular dynamics (PIMD). Since fractionation is a function of the difference in QKE between hydrogen and deuterium between two environments, examining patterns in QKE is a proxy for directly examining fractionation ratios.

Due to the immense practical interest in water chemistry, we are interested in answering a few specific questions about nuclear quantum effects in water. Quantum effects have been previously explored in water [11], with two competing effects understood: the weakening of hydrogen bonds through hydrogen delocalization disrupts bulk structure, and the strengthening of hydrogen bonds through QKE elongating the OH bond which allows shorter and stronger hydrogen bonds.

Now that we have this outline of how nuclear quantum effects manifest themselves in the bonding of water, we can move on to addressing the following questions:

- Hydrogen’s QKE gives an indication of what chemical environment the atom is in; how does this map onto the hydrogen bond structure of water?
- Are there systematic effects with increasing deuterium concentration on observed QKE, fractionation ratios, or water properties, and what are the origin of these effects?
- Since the OD stretch in HOD is often used as a spectroscopic probe, is there any predicted difference in equilibrium distribution of HOD versus H₂O between phases or at interfaces that could change experimental interpretation?
- If an interface does affect QKE, do small clusters of tens to hundreds of water molecules, which can be a major contributing environment in atmospheric studies, show behavior that deviates from either the bulk or larger interfaces?

The rest of this thesis will set up the theoretical methods used in section 2, provide a start at answering the questions above in section 3, then close with some suggestions for further investigation in section 4.

2. THEORETICAL METHODS

This section will provide an overview of the theoretical methods used to simulate condensed phase water and capture quantum effects, take advantage of thermostating to decrease the computational cost, calculate QKE from simulation trajectory, and compute fractionation ratios.

2.1. Path Integral Molecular Dynamics. Here is a summary of the full derivation of PIMD based on the explanations of Markland [12] and Schulman [17].

2.1.1. Path Integral Quantum Mechanics. The expectation value A of an observable \hat{A} in a single quantum state with density matrix $\boldsymbol{\rho}$ can be given by taking the trace “tr” over the operator’s value in the component eigenstates of that state; this is done by summing the expectation value of the observable in each eigenstate i by the density of that eigenstate ρ_i .

$$A = \text{tr}[\boldsymbol{\rho}\hat{A}] = \sum_i \rho_i \langle \psi_i | \hat{A} | \psi_i \rangle \quad (1)$$

If we want the thermally averaged expectation value, we need to weight the contribution of each eigenstate in the trace by a Boltzmann factor, which takes into account the energy of each state through the Hamiltonian \hat{H} , and then normalize by the partition function Z ; k_B

is Boltzmann's constant and T is the temperature.

$$\langle A \rangle = \frac{1}{Z} \text{tr} \left[\exp(-\beta \hat{H}) \hat{A} \right] \quad \text{where} \quad \begin{aligned} Z &= \text{tr} \left[\exp(-\beta \hat{H}) \right] \\ \beta &= 1/k_B T \end{aligned} \quad (2)$$

One can calculate this trace by integrating over all position eigenstates via the state variable x_1 , yielding

$$\langle A \rangle = \frac{1}{Z} \int dx_1 \langle x_1 | \exp(-\beta \hat{H}) | x_1 \rangle A(x_1) \quad (3)$$

We can then use the Lie-Trotter operator product formula [18]

$$\exp(\hat{A} + \hat{B}) = \lim_{n \rightarrow \infty} \left[\exp \left(\frac{\hat{A}}{n} \right) \exp \left(\frac{\hat{B}}{n} \right) \right]^n \quad (4)$$

along with the fact the Hamiltonian is composed of kinetic and potential energy terms

$\hat{H} = \hat{T} + \hat{V}$ to break up the Hamiltonian above into

$$\langle A \rangle = \frac{1}{Z} \lim_{n \rightarrow \infty} \int dx_1 \langle x_1 | \left[\exp(-\beta_n \hat{T}) \exp(-\beta_n \hat{V}) \right]^n | x_1 \rangle A(x_1) \quad (5)$$

where $\beta_n = \beta/n$.

This form is not convenient since we would like an expression for the expectation value that does not depend on operators but instead depends solely on the parameters of the system, potential V , temperature T , and mass of the particle m .

We insert a position identity operator $\hat{1}_x$ that will integrate over all position eigenstates,

$$\hat{1}_x = \int dx |x\rangle \langle x| \quad (6)$$

$n - 1$ times between each of the n exponential operators. None of these position integrals can be evaluated independently since each of the identity operators is chaining together two terms to form sets of complete brackets. Thus all of the integrals must remain independent and are brought out front. We already have one integral over all position eigenstates via x_1 , so we will call these $n - 1$ other integrating variables x_2 through x_n . All of the following sums and products in k are cyclic, with $n + 1 \rightarrow 1$ to simplify the written expression. The cyclic property is from the trace as the first bra and last ket in the series are the same position eigenstate.

$$\langle A \rangle = \frac{1}{Z} \lim_{n \rightarrow \infty} \int dx_1 \cdots \int dx_n \prod_{k=1}^n \langle x_k | \exp(-\beta_n \hat{T}) \exp(-\beta_n \hat{V}) | x_{k+1} \rangle A(x_1) \quad (7)$$

The potential energy operator \hat{V} acting on any of the position eigenstates $x_1 \cdots x_n$ yields the potential function value at that position $V(x)$ as the eigenvalue. Thus, each of the n potential function values come out of the brackets and can be collected in front of the product

in an exponential.

$$\begin{aligned} \langle A \rangle = \frac{1}{Z} \lim_{n \rightarrow \infty} \int dx_1 \cdots \int dx_n \exp \left(-\beta_n \sum_{k=1}^n V(x_k) \right) \\ \times \prod_{k=1}^n \langle x_k | \exp(-\beta_n \hat{T}) | x_{k+1} \rangle A(x_1) \end{aligned} \quad (8)$$

To work on the kinetic energy operator, we insert just after each kinetic energy term a momentum identity operator $\hat{1}_p$

$$\hat{1}_p = \int dp |p\rangle \langle p| \quad (9)$$

in the product to give

$$\begin{aligned} \langle A \rangle = \frac{1}{Z} \lim_{n \rightarrow \infty} \int dx_1 \cdots \int dx_n \exp \left(-\beta_n \sum_{k=1}^n V(x_k) \right) \\ \times \prod_{k=1}^n \int dp_k \langle x_k | \exp(-\beta_n \hat{T}) | p_k \rangle \langle p_k | x_{k+1} \rangle A(x_1) \end{aligned} \quad (10)$$

We will keep the integral over momentum eigenstates inside of the product for now, unlike the position integrals which we split and pulled out, so that we can use an identity in a moment.

The kinetic energy operator \hat{T} acts upon any of the momentum eigenstates p to yield $p^2/2m$ in front of each bracket as the eigenvalue.

$$\begin{aligned} \langle A \rangle = \frac{1}{Z} \lim_{n \rightarrow \infty} \int dx_1 \cdots \int dx_n \exp \left(-\beta_n \sum_{k=1}^n V(x_k) \right) \\ \times \prod_{k=1}^n \int dp_k \exp \left(\frac{-\beta_n p_k^2}{2m} \right) \langle x_k | p_k \rangle \langle p_k | x_{k+1} \rangle A(x_1) \end{aligned} \quad (11)$$

One can show [17] that $\langle x | p \rangle$ is identical to

$$\langle x | p \rangle = \sqrt{\frac{1}{2\pi\hbar}} \exp(+ipx/\hbar) \quad \langle p | x \rangle = \sqrt{\frac{1}{2\pi\hbar}} \exp(-ipx/\hbar) \quad (12)$$

by using the definition of a momentum eigenstate $|p\rangle$ and left multiplying by $\langle x|$ and working out the integral. Plugging that in and rearranging gives

$$\begin{aligned} \langle A \rangle = \frac{1}{Z} \lim_{n \rightarrow \infty} \int dx_1 \cdots \int dx_n \exp \left(-\beta_n \sum_{k=1}^n V(x_k) \right) \\ \times \prod_{k=1}^n \frac{1}{2\pi\hbar} \int dp_k \exp \left(\frac{-\beta_n p_k^2}{2m} \right) \exp(ip_k(x_k - x_{k+1})/\hbar) A(x_1) \end{aligned} \quad (13)$$

then using the fact that these are Gaussian integrals with the form

$$\int_{-\infty}^{\infty} dy \exp(-a y^2 + b y) = \sqrt{\frac{\pi}{a}} \exp\left(\frac{b^2}{4a}\right) \quad (14)$$

we can show

$$\begin{aligned} \langle A \rangle = \frac{1}{Z} \lim_{n \rightarrow \infty} \int dx_1 \cdots \int dx_n \exp \left(-\beta_n \sum_{k=1}^n V(x_k) \right) \\ \times \prod_{k=1}^n \sqrt{\frac{m}{2\pi\beta_n \hbar^2}} \exp(-\beta_n m \omega_n^2 (x_k - x_{k+1})^2 / 2) A(x_1) \end{aligned} \quad (15)$$

where $\omega_n = 1/(\hbar\beta_n)$. Rewriting this in summation form yields

$$\begin{aligned} \langle A \rangle = \frac{1}{Z} \lim_{n \rightarrow \infty} \left(\frac{m}{2\pi\beta_n \hbar^2} \right)^{n/2} \\ \times \int dx_1 \cdots \int dx_n \exp \left(-\beta_n \sum_{k=1}^n \left[\frac{m\omega_n^2}{2} (x_k - x_{k+1})^2 + V(x_k) \right] \right) A(x_1) \end{aligned} \quad (16)$$

Equation 16 is the *classical isomorphism* of Chandler and Wolynes [6].

2.1.2. Interpretation of the Path Integral. Let's take a moment to physically interpret Equation 16. In this equation, the relationship between x_i and x_{i+1} is not one of a step in real time, but a step along the *imaginary* time path.

We can make the exponent in Equation 16 look like the Riemann sum definition of an integral (sum of infinitely many, infinitesimally thin rectangles under a curve)

$$\int_0^b dx f(x) = \lim_{n \rightarrow \infty} \frac{b}{n} \sum_{i=0}^n f\left(i \frac{b}{n}\right) \quad (17)$$

after the following transformations:

$$\lim_{n \rightarrow \infty} -\beta_n \sum_{k=1}^n \left[\frac{m\omega_n^2}{2} (x_k - x_{k+1})^2 + V(x_k) \right] \quad (18)$$

$$= \lim_{n \rightarrow \infty} -\frac{\beta}{n} \sum_{k=1}^n \left[\frac{mn^2}{2\hbar^2\beta^2} (x_k - x_{k+1})^2 + V(x_k) \right] \quad (19)$$

$$= \frac{-1}{\hbar^2\beta} \lim_{n \rightarrow \infty} d\tau \sum_{k=1}^n \left[\frac{m}{2} \left(\frac{x_k - x_{k+1}}{d\tau} \right)^2 + V(x_k) \right] \quad \text{where } d\tau = 1/n \text{ is small} \quad (20)$$

Lie-Trotter splitting introduced n and so it has no *a-priori* physical meaning. We previously thought about the indexing variable of x as a imaginary time step, so it follows to interpret $d\tau = 1/n$ as an infinitesimal in this imaginary time τ . Using the definition of a Riemann sum in Equation 20, we get the integral

$$\frac{-1}{\hbar^2\beta} \int_0^\infty d\tau \left[\frac{m}{2} \left(\frac{dx}{d\tau} \right)^2 + V(x) \right] \quad (21)$$

If we interpret $1/n$ in this way, then the quantity in brackets is the classical Lagrangian L in imaginary time τ

$$L = \frac{1}{2}m \left(\frac{dx}{d\tau} \right)^2 - V(x) \quad (22)$$

The classical Lagrangian summed over a path gives the classical action S along that path.

$$S = \int_{\text{path}} d\tau L \quad (23)$$

In classical mechanics, the path through time and space which produces a minimum in the action S is the path that occurs.

We’ve now shown that the quantity in the exponent of Equation 16 is the classical action over a path through imaginary time multiplied by $-1/(\hbar^2\beta)$. This gives an interpretation of the equation for the quantum expectation value of an observable:

$$\langle A \rangle = \frac{1}{Z} \underbrace{\lim_{n \rightarrow \infty}}_{\text{look at all}} \left(\frac{m}{2\pi\beta_n\hbar^2} \right)^{n/2} \underbrace{\int dx_1 \cdots \int dx_n}_{\text{paths through space}} \underbrace{\exp\left(\frac{-1}{\hbar^2\beta} S[x_1 \dots x_n]\right)}_{\text{and weight by imaginary time action}} \underbrace{A(x_1)}_{\text{the observable}}$$

2.1.3. *Path Integral Molecular Dynamics.* The formalism of path integral quantum mechanics and the classical isomorphism are useful because they provide a route to simulating quantum systems using the tools of classical simulation.

Looking at Equation 16 again (repeated here):

$$\begin{aligned} \langle A \rangle = & \frac{1}{Z} \lim_{n \rightarrow \infty} \left(\frac{m}{2\pi\beta_n\hbar^2} \right)^{n/2} \\ & \times \int dx_1 \cdots \int dx_n \exp \left(-\beta_n \sum_{k=1}^n \left[\frac{m\omega_n^2}{2} (x_k - x_{k+1})^2 + V(x_k) \right] \right) A(x_1) \end{aligned}$$

We can think of a single quantum particle as a set of classical particles, each called a *bead*, that are each evolving in their own potential (notice V is only a function of one x_i) at a temperature n times that of the physical system. There is also one interaction between these system copies that is a harmonic spring between “adjacent” beads (given by cyclic index k

and force constant ω_n^2). See Figure 1 for an illustration. Since these beads are connected in a cycle, they are called *ring polymers*, and this cycle forms an imaginary time path.

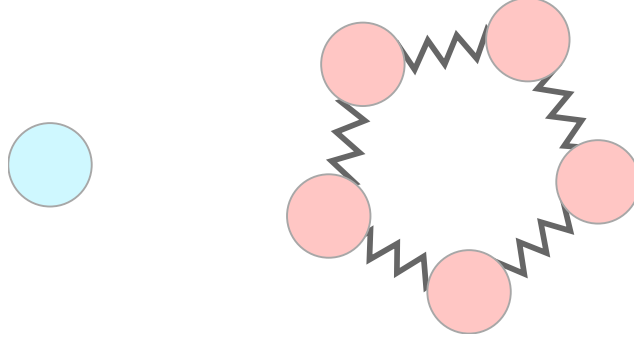


FIGURE 1. The classical isomorphism of a quantum particle turns it into a ring polymer with beads that are independent copies of the original particle evolving at a higher temperature connected by harmonic springs. The strength of the springs and the effective temperature depend on the original mass of the particle and physical system temperature.

The higher temperature that the beads evolve with and the harmonic connection between beads allows the ring polymers to explore parts of phase space that would be classically energetically forbidden. This is the basis of how quantum effects are introduced.

Since each of the n systems evolves in their own potential, an n bead PIMD simulation requires about n times more computational resources than a classical one. This method becomes formally exact as $n \rightarrow \infty$, however one only needs to obtain values to a given numerical precision in practice and hence a finite number of beads can be used. Not every system has large quantum effects; systems with low temperature and high frequencies will need larger numbers of beads to capture their quantum effects accurately. As a guideline

[12] the quantum effects in a system can be accurately captured when the number of beads

$$n > \beta \hbar \omega_{\max} \quad (24)$$

where $\beta = 1/T$ and ω_{\max} is the highest frequency in the system. Techniques for reducing the number of beads to obtain accurate measurements using thermostating will be discussed in subsection 2.2.

PIMD uses the dynamical evolution of the classical Hamiltonian in Equation 16 to extract equilibrium quantum expectation values [15]. This is done by the introduction of a set of momenta and masses for the beads using the identity

$$1 = \left(\frac{\beta_n}{2\pi} \right)^{n/2} \left(\frac{1}{\det \mathbf{M}} \right)^{1/2} \int d^n \mathbf{p} \exp \left(-\frac{\beta_n}{2} \mathbf{p}^T \mathbf{M} \mathbf{p} \right) \quad (25)$$

where “det” is the determinant, \mathbf{M} is the mass matrix, and \mathbf{p} contains the bead momenta.

This can be incorporated into the path integral expression for an expectation value of Equation 16 to yield [12]

$$\langle A \rangle = \lim_{n \rightarrow \infty} \frac{1}{Z_n} \left(\frac{m^n}{(2\pi\hbar)^n \det \mathbf{M}} \right)^{1/2} \int d^n \mathbf{p} \int d^n \mathbf{x} \exp(-\beta_n H_n(\mathbf{p}, \mathbf{x})) A_n(\mathbf{x}) \quad (26)$$

where H_n and Z_n are given by

$$H_n = \frac{1}{2} \mathbf{p}^T \mathbf{M}^{-1} \mathbf{p} + \sum_{k=1}^n \left[\frac{m\omega_n^2}{2} (\mathbf{x}_k - \mathbf{x}_{k-1})^2 + V(\mathbf{x}_k) \right] \quad (27)$$

$$Z_n = \left(\frac{m^n}{(2\pi\hbar)^n \det \mathbf{M}} \right)^{1/2} \int d^n \mathbf{p} \int d^n \mathbf{x} \exp(-\beta_n H_n(\mathbf{p}, \mathbf{x})) \quad \text{so } \lim_{n \rightarrow \infty} Z_n = Z \quad (28)$$

Regardless of the masses chosen for the PIMD simulation, static classical equilibrium properties will be calculated properly; classical *dynamics*, though, will only occur if the masses are the actual atomic masses.

We can now use the classical dynamics described by the ring polymer Hamiltonian in Equation 27 to calculate equilibrium quantum observables by time averaging over generated trajectories [15].

2.1.4. *Calculating Quantum Kinetic Energy.* In this thesis, we are only interested in the quantum kinetic energy equilibrium observable. The simplest way to calculate it [19] starts with the statistical mechanics route from the partition function Z to total energy E :

$$E = -\frac{\partial}{\partial \beta} \ln Z = \frac{1}{Z} \frac{\partial Z}{\partial \beta} = \lim_{n \rightarrow \infty} \frac{1}{Z_n} \frac{\partial Z_n}{\partial \beta} \quad (29)$$

Using the definition of the ring polymer partition function Z_n from Equation 28

$$Z_n = \left(\frac{m^n}{(2\pi\hbar)^n \det \mathbf{M}} \right)^{1/2} \int d^n \mathbf{p} \int d^n \mathbf{x} \exp \left(\frac{-\beta_n}{2} \mathbf{p}^T \mathbf{M}^{-1} \mathbf{p} \right. \\ \left. - \beta_n \sum_{k=1}^n \left[\frac{m\omega_n^2}{2} (\mathbf{x}_k - \mathbf{x}_{k-1})^2 + V(\mathbf{x}_k) \right] \right) \quad (30)$$

and partial differentiating, we derive [19]

$$E = \lim_{n \rightarrow \infty} \frac{n}{2\beta} - \sum_{k=1}^n \left[\frac{mn}{2\beta^2 \hbar^2} (\mathbf{x}_{k+1} - \mathbf{x}_k)^2 \right] + \frac{1}{n} \sum_{k=1}^n [V(\mathbf{x}_k)] \quad (31)$$

Notice that the kinetic energy term¹ has an n in the sum; this means that the individual terms that are being averaged get more and “noisy” as the number of beads is increased, even if the average stays the same. This makes it numerically more challenging to calculate QKE using this method. The centroid virial kinetic energy estimator [1] does not suffer from this problem.

In classical mechanics, the virial kinetic energy estimator says we can know the expectation value for kinetic energy given the expectation value for the potential. The centroid virial kinetic energy estimator uses this intuition and the idea of kinetic energy in imaginary time to give

$$T_n = \frac{9}{2\beta} + \frac{3}{2n} \sum_{k=1}^n \left[(\mathbf{x}_k - \bar{\mathbf{x}}) \frac{\partial V(\mathbf{x}_k)}{\partial \mathbf{x}_k} \right] \quad (32)$$

where the centroid of the particle $\bar{\mathbf{x}}$ is defined as

$$\bar{\mathbf{x}} = \frac{1}{n} \sum_{k=1}^n \mathbf{x}_k \quad (33)$$

¹It is interesting that the quantum kinetic energy has a very similar form to the classical kinetic energy, just with a “velocity” through imaginary time, or from one bead to the next.

The $9/(2\beta)$ term is a classical contribution due to the thermal kinetic energy of the particle, while the rest is a quantum kinetic energy contribution. Since fractionation is calculated through a difference in free energies (see subsection 2.3), only the quantum terms will not cancel, and thus it practically suffices to only calculate the quantum contribution.

2.2. PIGLET. In order to accurately sample equilibrium thermodynamic quantum properties in PIMD simulations, there needs to be enough beads on the ring polymers to capture the degree to which the system is behaving quantum mechanically. Because increasing the number of beads increases computational cost, there is a need for techniques to enable convergence of quantum expectation values with fewer beads. One method of achieving this through careful thermostating of the beads is called *PIGLET*.

2.2.1. Generalized Langevin Equation. To understand how thermostating can be used to control the convergence of equilibrium properties, an overview of the formalism necessary is presented here based on the review of Ceriotti [3].

Brownian motion is the motion of a given larger tagged particle due to the constant random buffeting of a sea of smaller particles, or *heat bath*, at constant temperature. Since the tagged particle can exchange energy with the bath, within the constraints of constant temperature, the tagged particle will sample the *canonical ensemble* of the potential surface it is experiencing. Many physical processes happen in an environment like a heat bath. Naively simulating this requires the computationally intensive approach of explicitly taking into account all of the particles in the bath.

Langevin dynamics attempts to ignore the specifics of the bath but correctly manifest the effects of the bath on the system. The specific actions of bath particles on the tagged ones are integrated out into a random stochastic buffeting term and a frictional force. The simplest expression for these dynamics is

$$\dot{q} = p \tag{34}$$

$$\dot{p} = -V(q) - \underbrace{a \cdot p}_{\text{friction}} + \underbrace{b \cdot \xi(t)}_{\text{buffeting}} \tag{35}$$

where q is position, p is momentum, a dot indicates the time derivative, and V is potential. These resemble Hamilton’s equations of motion, but with an extra friction term a that acts to oppose velocity, and a random buffeting term, where $\xi(t)$ is a Gaussian random number and b is a force scaling constant. The character of this stochastic force is crucial; many proofs using this formalism require that the random noise has certain distributions or variances. In the form of Equation 35, this noise is “white noise” because it couples to all frequencies in the system uniformly.

To have a thermostat that can selectively affect different frequencies in the system, one with a more powerful form called a generalized Langevin equation (GLE) is needed. “White noise” Langevin dynamics has noise as a random function of only the current state of the system; it has no *memory*, or the random noise can’t depend on previous system configurations. To have the random noise couple to different system frequencies differently, called “colored noise,” the thermostat needs some way to keep track of frequencies in the system.

By introducing a vector of fictitious momenta \mathbf{s} , the thermostat can remember the past of the system; these momenta do not correspond to the momenta of any real particles. We then generalize the friction and buffeting scalars from the standard Langevin equation to the matrices \mathbf{A} and \mathbf{B} . The result is the GLE:

$$\dot{q} = p \tag{36}$$

$$\begin{pmatrix} \dot{p} \\ \dot{\mathbf{s}} \end{pmatrix} = \begin{pmatrix} -V(q) \\ 0 \end{pmatrix} - \mathbf{A} \begin{pmatrix} p \\ \mathbf{s} \end{pmatrix} + \mathbf{B}(\boldsymbol{\xi}) \tag{37}$$

These equations of motion are written here in a general form. To sample the canonical ensemble, the *fluctuation-dissipation theorem* needs to be imposed, and that involves a specific relationship between \mathbf{A} and \mathbf{B} . *Other* ensembles or distributions can be enforced by choosing the matrices carefully.

2.2.2. PI+GLE. PIMD simulations capture tunneling and zero-point energy but can be computationally expensive with many beads. A GLE thermostat can enforce the correct quantum distribution of an observable in a path integral simulation using fewer beads, and this technique is referred to as PI+GLE.

There are other methods for reducing the cost of PIMD simulations. One is ring polymer contraction [13], but it requires the ability to break up forces acting upon a particle into long and short-range ones, which is not possible when using *ab-initio* techniques to calculate them. PI+GLE is a general technique that can use any method of calculating forces.

To go about enforcing the correct quantum distribution, one can show [4] that the equilibrium variance in position of a single bead in the ring polymer $\langle q^2 \rangle$ can be expressed by averaging the variance in position of each of the n beads $\langle q_k^2 \rangle$.

$$\langle q^2 \rangle = \frac{1}{n} \sum_{k=0}^{n-1} \langle q_k^2 \rangle \quad (38)$$

If the ring polymer is in a harmonic potential well, then the normal mode frequencies ω_f present in that ring polymer are analytically given by [4]

$$\omega_f = \sqrt{\omega^2 + 4\omega_n^2 \sin^2(f\pi/n)} \quad (39)$$

where f is the normal mode index, n is the number of ring polymer beads, ω is the characteristic frequency of the harmonic potential well, and ω_n is the harmonic spring frequency of the ring polymer itself (remember that ω_n depends on temperature, \hbar , and n). Since the ring polymer normal modes are analytically known in the limit of a harmonic environment, rewriting the expression for the variance in a single bead's position from Equation 38 in terms separate harmonic oscillators at the normal mode frequencies, each coupled to their own GLE thermostat, yields

$$\langle q^2 \rangle = \frac{1}{n} \sum_{f=0}^{n-1} c_{qq}(\omega_f) \quad (40)$$

where each GLE thermostat is enforcing the normal mode variances c_{qq} for a single quantum harmonic oscillator with the normal mode frequency

$$c_{qq}(\omega) = \frac{\hbar}{2\omega} \coth \frac{\hbar\omega}{2k_B T} \quad (41)$$

The variance is called c_{qq} because it is a part of a \mathbf{C} matrix that is part of a constraint on \mathbf{A} and \mathbf{B} arising from enforcing the fluctuation–dissipation theorem.

The \mathbf{A} and \mathbf{B} matrices have to be fit to ensure this behavior during the simulation. This fitting is “general” in that, as long as you are close to the harmonic limit, the same matrices will apply to any system simulated. The specifics of fitting these matrices is discussed by Ceriotti [3].

2.2.3. *PIGLET*. The PI+GLE method only thermostats individual beads and does not take into account correlations between beads. The virial kinetic energy estimator in Equation 32 of subsection 2.1.4 depends on the positions of the beads relative to the centroid, so it includes correlations between bead positions. Because bead correlations are involved, the kinetic energy equilibrium expectation value will not converge as quickly using standard PI+GLE as it would using a method that explicitly thermostats these correlations to enforce the quantum distribution directly.

One can manipulate the equation for the expectation value of kinetic energy $\langle T \rangle$ from the virial kinetic energy estimator in order to reveal where the contributions to the value are

coming from. Start with Equation 32 in one dimension

$$\langle T \rangle = \frac{1}{2\beta} + \frac{1}{2n} \sum_{k=1}^n \left[\left\langle (x_k - \bar{x}) \frac{\partial V(x_k)}{\partial x_k} \right\rangle \right] \quad (42)$$

where $\beta = 1/(k_B T)$, n is the number of beads, and x_k is the position of bead k . If the ring polymer is in a harmonic potential, then $V(x) = \omega^2 x^2/2$ and that yields

$$\langle T \rangle = \frac{1}{2\beta} + \frac{\omega^2}{2n} \sum_{k=1}^n [\langle (x_k - \bar{x}) x_k \rangle] \quad (43)$$

Rearranging the equation and breaking up the sum shows that $\langle T \rangle$ can be constructed in terms of the expectation values for $\langle V \rangle$ and $\langle \bar{x} \rangle$:

$$= \frac{1}{2\beta} + \frac{\omega^2}{2n} \sum_{k=1}^n [\langle x_k^2 - \bar{x} x_k \rangle] \quad (44)$$

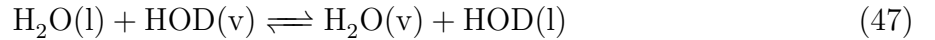
$$= \frac{1}{2\beta} + \underbrace{\frac{1}{n} \sum_{k=1}^n \left[\left\langle \frac{\omega^2}{2} x_k^2 \right\rangle \right]}_{\langle V \rangle} - \frac{\omega^2}{2} \langle \bar{x} \rangle \underbrace{\frac{1}{n} \sum_{k=1}^n [\langle x_k \rangle]}_{\langle \bar{x} \rangle} \quad (45)$$

$$= \underbrace{\frac{1}{2\beta} - \frac{\omega^2}{2} \langle \bar{x} \rangle^2}_{\text{centroid}} + \underbrace{\langle V \rangle}_{\text{ring polymer}} \quad (46)$$

Equation 46 says that the total kinetic energy expectation value can be broken into a component due to the centroid kinetic energy in the harmonic potential field, and a component due to the quantum expectation of potential felt by all the beads.

A method described by Ceriotti [5] called PIGLET classically thermostats the centroid, while using a GLE thermostat on ring polymer normal modes to enforce the correct distribution in the quantum term in Equation 46 with a smaller bead requirement. Since the correct quantum bead correlations are now explicitly imposed by the thermostat, $\langle T \rangle$ converges with fewer beads.

2.3. Fractionation Ratios. We are interested in an experimental manifestation of nuclear quantum effects, namely the equilibrium fractionation ratio between hydrogen and deuterium between liquid and vapor phase water. The equilibrium being studied is



i.e. exchanging one deuterium in the vapor phase with one hydrogen in the liquid phase.

The equilibrium coefficient is a static property, so the Helmholtz free energy change associated with Equation 47 will determine whether deuterium prefers one phase or another. We can quantify this preference using the liquid–vapor fractionation ratio α_{lv} which is a function of the equilibrium mole fractions and thus the equilibrium’s ΔA .

$$\alpha_{lv} = \frac{\chi_{D,l}/\chi_{H,l}}{\chi_{D,v}/\chi_{H,v}} = \exp(-\Delta A/(k_B T)) \quad (48)$$

Many interesting practical uses of hydrogen–deuterium fractionation occur in the dilute deuterium limit. Examining the origin of a atmospheric water sample (where deuterium is

6000 times less common than hydrogen) or using very dilute HOD as a spectroscopic probe both are instances of this limit. Because of this, we will consider this limit in our simulations.

To calculate ΔA from PIMD simulations, we use the centroid virial estimator described in subsection 2.1.4, along with the thermodynamic integration [20]

$$\Delta A = \int_{m_H}^{m_D} dm \frac{\langle T_v(m) \rangle - \langle T_l(m) \rangle}{m} \quad (49)$$

In thermodynamics, the change in a state variable is independent of path so if we make up a path that connects the correct starting and ending points, we are guaranteed the correct value. The path invoked here is non-physically changing the mass of one hydrogen to that of deuterium.

The straightforward way of evaluating the integral in Equation 49 would be to run a set of simulations varying the mass of one of the hydrogens and finding a numerical approximation. Looking at Figure 2, QKE versus inverse mass squared forms a linear relationship and thus only two points are required to uniquely define the integral. This practically means that only two, instead of around ten, simulations are needed to calculate a fractionation ratio.

ΔA is a function of a difference of two kinetic energy expectation values, and since the thermal contribution to each of those is the same, the difference can be further reduced to just the difference in QKE. To obtain ΔA we perform simulations to calculate:

- the QKE of an hydrogen surrounded by $\text{H}_2\text{O}(\text{l})$
- the QKE of a deuterium surrounded by $\text{H}_2\text{O}(\text{l})$

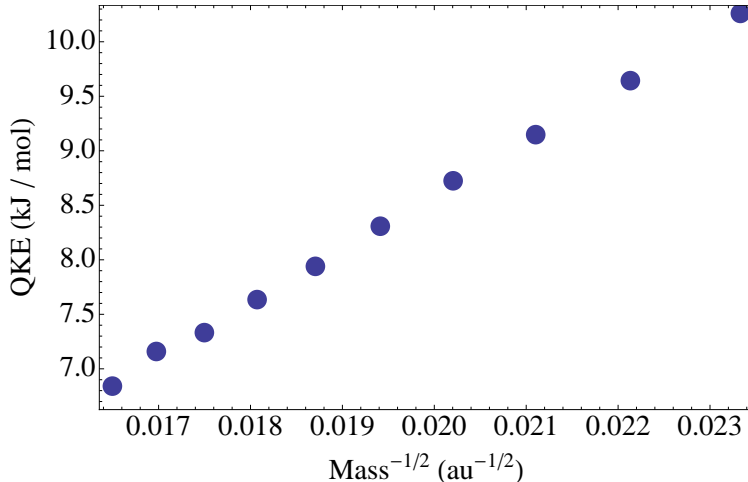


FIGURE 2. QKE of X in a simulation of bulk HOX using the parameters of subsection 3.2.2. X is a hydrogen-like atom with a mass that varies between that of hydrogen and deuterium. Error bars are smaller than the points. The relationship presented here is linear, so its integral can be calculated knowing only two data points. This greatly simplifies the calculations needed for a fractionation ratio.

- the QKE of hydrogen in $\text{H}_2\text{O}(\text{v})$
- the QKE of deuterium in $\text{HOD}(\text{v})$

ΔA can then be used to calculate the fractionation ratio using Equation 48. Since the fractionation ratio is a function of QKE, patterns in relative QKE can be looked at as a proxy for the behavior of the fractionation ratio.

3. SIMULATION RESULTS

3.1. Hydrogen Bonding Effects on Quantum Kinetic Energy. Because atomic-scale water structure is dominated by hydrogen bonding, it is useful to use hydrogen bonding donor and acceptor count to describe a water's environment. We calculated hydrogen bonding donor and acceptor count using the method of Hynes [10] and binned the QKE of each water's hydrogens by this donor and acceptor count.

3.1.1. *System Setup.* We ran PIGLET simulations using 4 beads at 300 K. The system was $6 \times 6 \times 6 = 216$ waters using periodic boundary conditions (PBC). A 0.5 fs timestep was used, with 5 ps of equilibration and 500 ps of production. Both q-TIP4P/F [7] and q-SPC/Fw [14] water models were simulated. Pure H_2O , pure HOD, and pure D_2O systems were investigated.

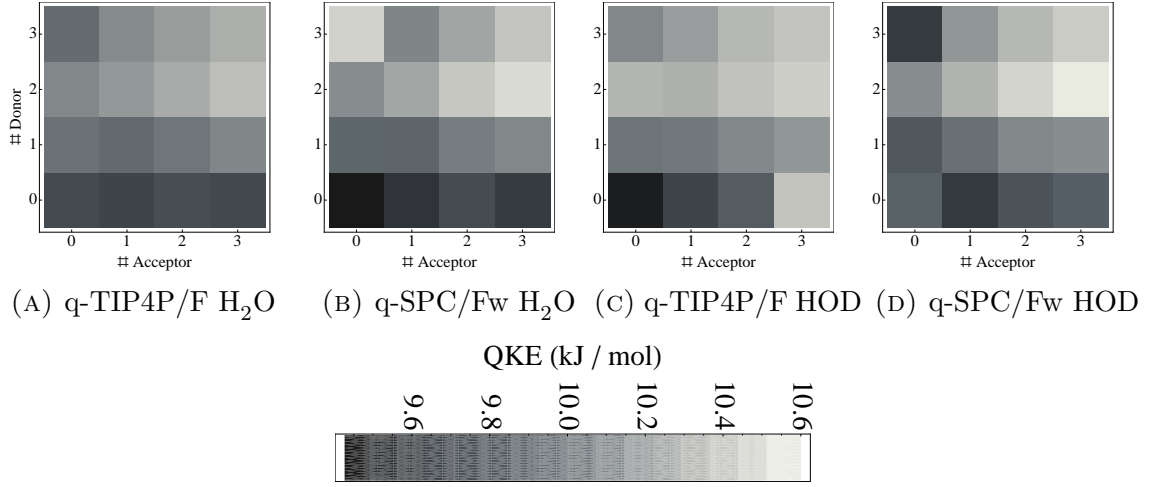


FIGURE 3. Average hydrogen QKE as a function of hydrogen bonding environment in PIGLET simulations in bulk H_2O and HOD (deuterium QKE is not shown). x -axis is the donor and y -axis is the acceptor count of the water whose hydrogen are binned. The bins labeled 3 contain all waters with ≥ 3 counts.

3.1.2. *Breakdown by Hydrogen Bonding.* Our results from the bulk H_2O and HOD simulations are visualized in Figure 3. A qualitative result from general quantum mechanics is that QKE increases as confinement increases. When thinking about confinement in a liquid, it helps to think about the problem in geometric and structural terms. There is a trend of increasing hydrogen QKE with the total number of hydrogen bonds correlating with increased hydrogen QKE because more hydrogen bonding partners mean more coordinating

neighbors to confine the given water. This is visualized in the lightening of the squares from the lower-left (0-donor 0-acceptor) to the upper-right of the plots.

As might be expected, there is a stronger correlation between donation count and positive hydrogen QKE; this can be seen in the larger lightness steps as you go up the y -axis as opposed to across the x -axis. Hydrogen bond donation directly involves the hydrogens on a given water; hydrogen bond acceptance is due to other waters coordinating with the oxygen, and so does not as directly affect the confinement of the hydrogens on a given water.

Waters that are donating to or accepting from three or more other waters are given a count of 3 in these plots. This makes any structural argument about the last column and row less robust, but it is interesting to consider the reversal in the trend of increasing hydrogen QKE with increasing donor count. One possible argument is as follows: When a given water accepts hydrogen bonds, more waters impinge on the given water's oxygen and can squeeze the hydrogens on the backside into the remaining space, increasing confinement and QKE. When a given water donates three hydrogen bonds, one of the given water's hydrogens will have to be pointing at two other water's oxygen atoms, providing a larger stabilizing negative charge than a single accepting oxygen could provide, and providing a larger stabilizing volume for that hydrogen, thus reducing confinement and QKE.

Further work is needed to determine the structures and environments that contribute to the decrease in QKE upon donating to three or more surrounding waters and to study how this behavior relates to the hydrogen bonding definition used.

In comparing the bulk H_2O simulations in Figure 3a and 3b to the bulk HOD simulations in Figure 3c and 3d, there isn't a qualitative difference in how hydrogen bonding affects hydrogen QKE, since a similar general pattern is still observed. The overall magnitude of QKE also is the same across these simulations. The effect of HOD in H_2O will be further explored in subsection 3.2, but these results point to there being little difference on the energy scale involved in hydrogen bonding.

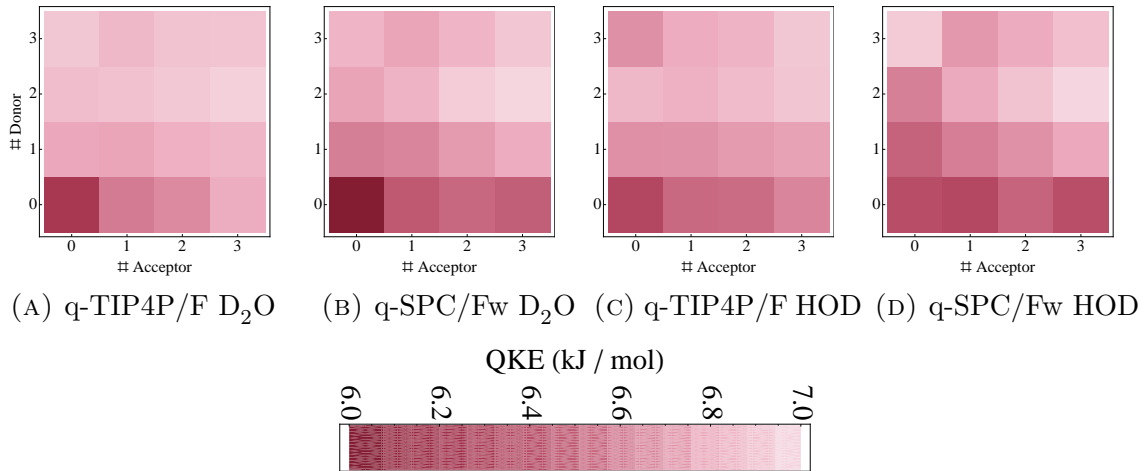


FIGURE 4. Average deuterium QKE as a function of hydrogen bonding environment in PIGLET simulations in bulk D_2O and HOD (hydrogen QKE is not shown). x -axis is the donor and y -axis is the acceptor count of the water whose deuterium are binned. The bins labeled 3 contain all waters with ≥ 3 counts.

3.1.3. Deuterium Results. Figure 4 shows the same qualitative behavior for deuterium QKE in D_2O and HOD as hydrogen displayed in Figure 3. All of the above discussion for hydrogen similarly applies in the analysis of the deuterium results.

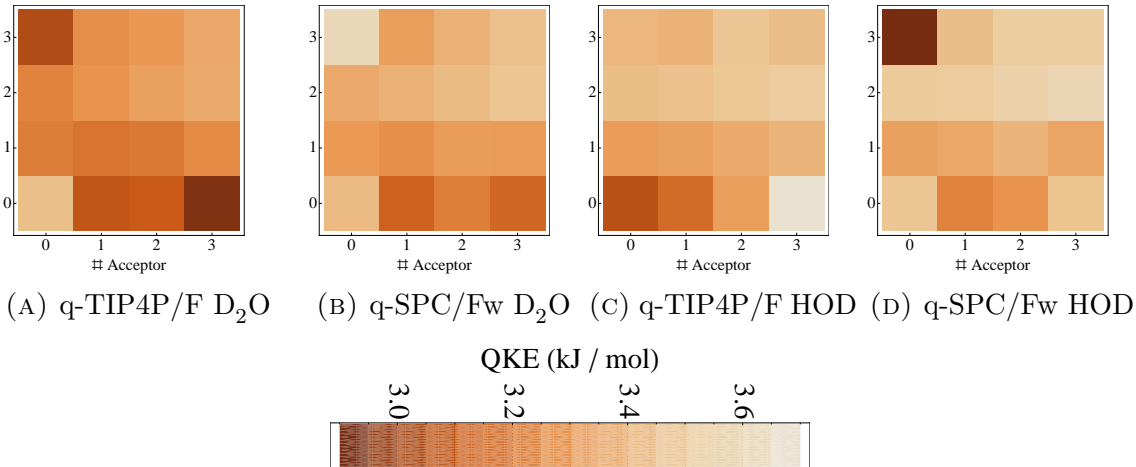


FIGURE 5. Difference in average QKE between hydrogen and deuterium as a function of hydrogen bonding environment in PIGLET simulations in bulk H_2O or D_2O and HOD. x -axis is the donor and y -axis is the acceptor count of the water whose deuterium are binned. The bins labeled 3 contain all waters with ≥ 3 counts.

3.1.4. *Fractionation Results.* Figure 5 shows the difference in QKE between hydrogen and deuterium as a function of hydrogen bonding environment. As was discussed in subsection 2.3, QKE difference is a proxy for fractionation ratio. It is very interesting that there is still the same general visual pattern in Figure 5, as this means that fractionation will not be uniform across all hydrogen bonding environments. Since higher donor and acceptor counts correspond to larger QKE differences, there is a slight fractionation between the hydrogen bonding environments. Situations that shift the balance of hydrogen bond structures in water, like interfaces or solutes, will change the fractionation ratio.

3.2. **Sampling of the Dilute Deuterium Limit.** As described in subsection 2.3, we are interested in investigating the dilute deuterium limit. In order to rigorously do this, all of our simulations have to sample the properties of deuterium in a single HOD surrounded by only H_2O . This exact simulation setup makes it computationally more intensive to fully sample

the equilibrium distribution for that deuterium. Here we investigate how the concentration of HOD in the simulation affects the observed equilibrium value of hydrogen and deuterium QKE. If there is little dependence on concentration, this means that conclusions from experiments and simulations that are not in the dilute deuterium limit will still be applicable there. This is specifically useful since we can ease sampling in our simulations by averaging over the environment of multiple deuteriums in the simulation box.

3.2.1. Quantum Kinetic Energy Decorrelation Timescale. We need to have a method for estimating error in equilibrium QKE values. One basic way to do this is to ensure that the measurements from simulations that contribute to calculating those values are completely uncorrelated, then basic statistical equations can return the variance and standard deviation of the expectation values [16].

For an observable A , the expectation value $\langle A \rangle$ is given by

$$\langle A \rangle = \frac{1}{M} \sum_i^M A_i \quad (50)$$

where M is the total number of measurements and A_i is the value of A during measurement i . The variance in A , $\sigma^2(A)$ is given by

$$\sigma^2(A) = \sum_i^M (A_i - \langle A \rangle)^2 = \langle A^2 \rangle - \langle A \rangle^2 \quad (51)$$

while the variance in $\langle A \rangle$, $\sigma^2(\langle A \rangle)$ is given by

$$\sigma^2(\langle A \rangle) = \frac{\sigma^2(A)}{M} \quad (52)$$

but *only* if all of the measurements of A are uncorrelated. To use these simple statistical equations, we need to ensure that our measurements are uncorrelated. In PIMD and PIGLET simulations, as in classical MD, there is a strong correlation between neighboring frames, as the trajectory is evolved dynamically. To discover the simulation timescale on which QKE decorrelates, we can plot the autocorrelation of QKE as a function of simulation time, and see when it decays to almost zero and call this the decorrelation time. If measurements are taken from the simulation only every decorrelation time, we can ensure uncorrelated data and use the above statistical analysis.

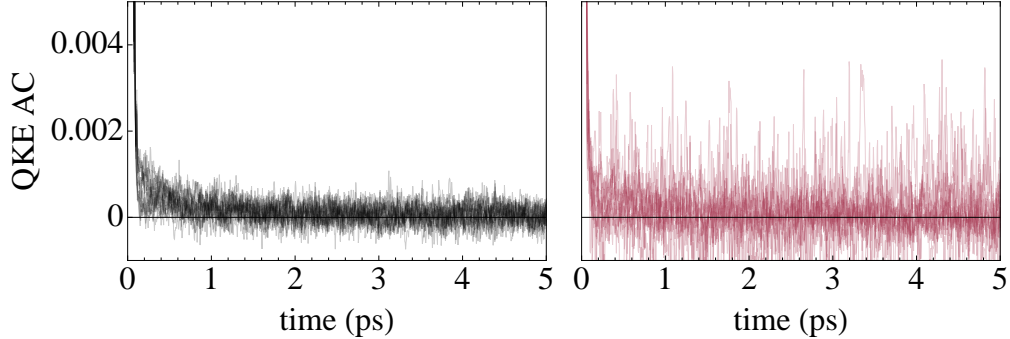


FIGURE 6. Decay in the autocorrelation of quantum kinetic energy for hydrogen (black) and deuterium (red) of all PIGLET simulations in this thesis superimposed.

The decay times for all of the bulk simulations used in this thesis are presented in Figure 6. The QKE autocorrelation curves show two decay timescales: the first is a precipitous drop off from 1 within a few femtoseconds, which matches the timescale of the fluctuations of the

ring polymer in its local chemical environment; the second is a slower decay ending around 2 – 3 ps, which is about the rotational decorrelation time of bulk water at 300 K.

Because of the slower timescale, we'll only be sampling every 3 ps of trajectory in order to get independent snapshots for this concentration HOD QKE error analysis.

3.2.2. System Setup. We ran PIGLET simulations using 4 beads at 300 K. The system was $6 \times 6 \times 6 = 216$ waters using PBC. A 0.5 fs timestep was used, with 5 ps of equilibration and 500 ps of production. Both q-TIP4P/F [7] and q-SPC/Fw [14] water models were simulated. Pure H₂O, pure HOD, pure D₂O, and concentrations of 0% (a single HOD molecule), 5%, 25%, 50%, 75%, and 100% HOD in H₂O systems were investigated.

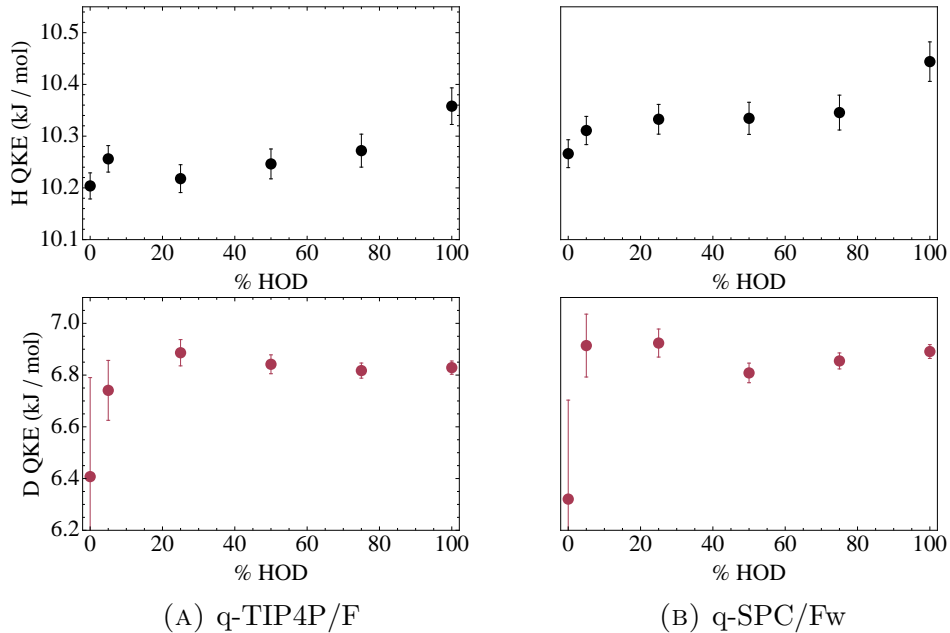


FIGURE 7. Dependence of hydrogen (black) or deuterium (red) QKE on percentage of HOD in H₂O. One standard deviation error bars are shown.

3.2.3. HOD Concentration Dependence. The results showing the dependence of hydrogen and deuterium QKE on concentration of HOD in H₂O are visualized in Figure 7. There is

a weak positive correlation between HOD concentration and hydrogen QKE. There is not a clear trend in deuterium QKE with HOD concentration.

What is clear from this data, is that the change in QKE in the bulk due to varying amounts of deuterium is small ($\approx 0.2 \text{ kJ/mol}$) relative to the QKE difference between hydrogen and deuterium ($\approx 3 \text{ kJ/mol}$) or to their isolated gaseous values ($\gtrsim 1 \text{ kJ/mol}$; see subsection 3.3.1). This means that, depending on how subtle of an isotopic effect is being examined, it might be acceptable to run simulations with higher percentages of HOD for better deuterium statistics, but still be able to apply those results to the dilute deuterium limit.

3.2.4. Intramolecular versus Intermolecular Deuterium Effects. In subsection 3.2.3, the effect of increasing the ratio of HOD to H_2O in the simulation box was investigated. The deuterium introduced was permanently bound to a given HOD since these simulation hydrogens are bound to their oxygens with an unbreakable empirical potential. This simulation setup ignores any intramolecular QKE effects between deuteriums on the same water molecule.

Simulations with identical parameters as the single component bulk simulations, but with 50% H_2O and 50% D_2O were run. The results in Figure 8 show that, given the same total concentration of deuterium in the simulation, having it in D_2O raises the deuterium QKE but lowers the hydrogen QKE relative to having all of the deuterium in HOD for both water models. 0% deuterium results come from a single HOD in H_2O simulations, while 100%

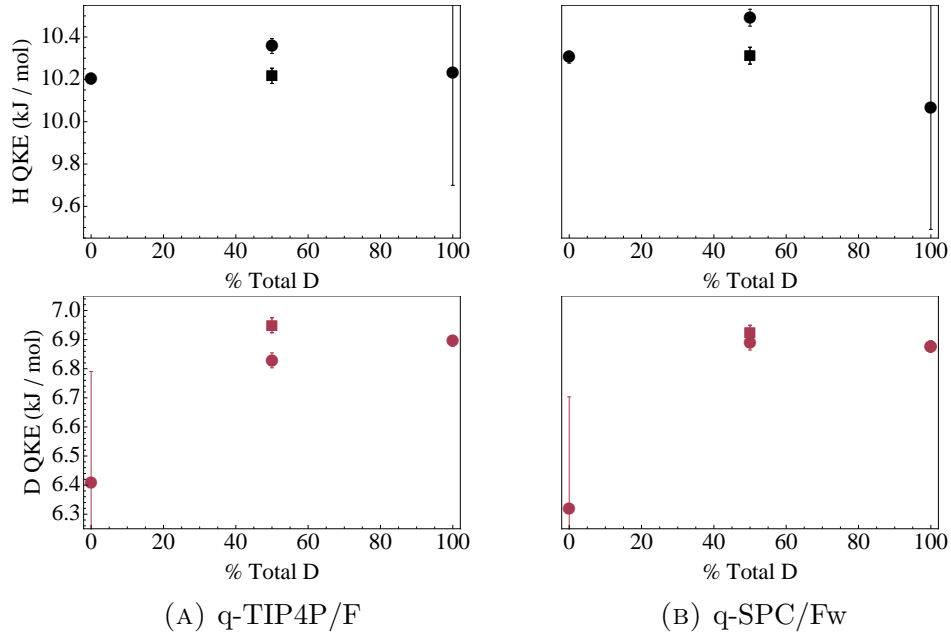


FIGURE 8. Dependence of hydrogen (black) or deuterium (red) QKE on total deuterium concentration in the simulation. 50% total deuterium concentration is investigated via one pure HOD simulation and one fifty–fifty mixture of H_2O and D_2O simulation (square marker).

deuterium results come from a single HOD in D_2O simulations, which are meant to explore the dilute deuterium and dilute hydrogen limits respectively.

The specific differences between the nature of hydrogen bonding and covalent oxygen–hydrogen bonding that are giving rise to this slight difference in QKE is currently unknown and further work is needed.

3.3. Quantum Kinetic Energy Distribution at Liquid–Vapor Interfaces. As was described in subsection 2.3 and expected from thermodynamics, equilibrium liquid–vapor fractionation can be calculated from the change in free energy between the bulk and the vapor. The interface itself is not considered explicitly, and any deviations from the bulk behavior in that environment are not included. To start to look at the properties of the

interface explicitly, whether there be fractionation between the bulk and the interface, or necessary corrections to small system liquid–vapor fractionation because of surface area, we have run and analyzed a set of water interface simulations.

3.3.1. Vapor Limit. In order to understand how the interface deviates from expected behavior, a full understanding of the two phases that are present is required. We’ve discussed the bulk liquid phase in the previous sections, but have not presented any information on hydrogen or deuterium QKE in the vapor phase.

To analyze vapor properties, we ran PIGLET simulations with 4 beads at 300 K. A 0.5 fs timestep was used with 5 ps of equilibration and 500 ps of production. Both q-TIP4P/F [7] and q-SPC/Fw [14] water models were simulated. A single H_2O , a single D_2O , and a single HOD molecule in a 100 Å on a side PBC box were used to simulate the vapor phase. A method identical to that in subsection 3.2.1 was employed to determine that QKE decorrelated over 500 fs of trajectory to get independent samples for error analysis. Table 1 has a summary of the results.

<i>water model</i>	<i>molecule</i>	<i>hydrogen QKE</i>	<i>deuterium QKE</i>
q-TIP4P/F	H_2O	6.3(1)	—
q-TIP4P/F	D_2O	—	5.1(1)
q-TIP4P/F	HOD	6.5(2)	4.9(2)
q-SPC/Fw	H_2O	9.2(2)	—
q-SPC/Fw	D_2O	—	6.3(1)
q-SPC/Fw	HOD	9.5(2)	6.1(1)

TABLE 1. quantum kinetic energy (QKE) of hydrogen and deuterium in isolated water isotopes for the q-TIP4P/F and q-SPC/Fw water models.

The two water models produce significantly different vapor QKE values but similar bulk liquid values. Previous analysis of the two empirical potentials [11] shows that q-SPC/Fw does not yield correct fractionation ratios; an anharmonic component must be added to q-SPC/Fw’s OH stretch.

3.3.2. *System Setup.* To analyze interfacial properties, we ran PIGLET simulations with 4 beads at 300 K. A 0.5 fs timestep was used, with 5 ps of equilibration and 500 ps of production. Both q-TIP4P/F [7] and q-SPC/Fw [14] water models were simulated. A $37.4 \times 37.4 \times 81.1$ Å PBC box, with 1000 HOD molecules placed in an ≈ 30 Å thick slab in the xy plane in the center of the box, was used to simulate a bulk liquid–vapor interface in the z direction. Since the results in subsection 3.2.3 showed that there was not a strong dependence of QKE on HOD concentration relative to the liquid–vapor QKE difference, we hope that this simulation will provide a acceptable first approximation of the dilute deuterium limit.

3.3.3. *Interfacial Distance Dependence.* We analyzed these simulations by binning hydrogen and deuterium’s QKE by z position, which should roughly correspond to distance from the interface (after subtracting off the slab’s middle z position). The difference in QKE is plotted in Figure 9. Based on the current simulation trajectories, we can not tell if there is a systematic variance in difference in QKE as a function of distance into the slab. Although it is difficult to definitively say from this data, it is interesting that the vapor phase QKE difference (plotted as the green horizontal line) appears to not be gradually approached in

q-TIP4P/Fw, where there is a larger gap between the vapor phase energy difference and the liquid phase energy difference; this points to the conclusion that bulk liquid like fractionation occurs out until the edge of the interface.

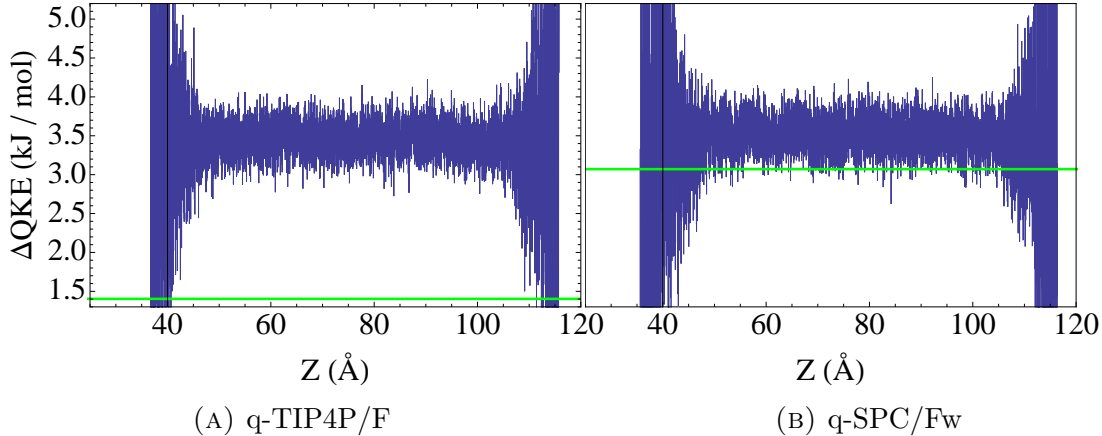


FIGURE 9. Difference in QKE between hydrogen and deuterium as a function of position through a slice of the simulation that contains the liquid–vapor interface. The horizontal green line is the vapor phase QKE difference.

There are subtleties in the definition of the liquid–vapor interface that have not been explored in this analysis that might bring about better resolution of energy differences. We are currently binning QKE by the z coordinate in the simulation box, and center of mass momentum removal ensures that the interface will stay in the middle of the box. We are ignoring capillary waves and the clearer picture that would result from using the distance from a fitted surface to the current atomic positions, similar to the Gaussian isosurface fitting that Chandler proposed [21].

3.4. Quantum Kinetic Energy Distribution in Isolated Water Clusters. Small water clusters are interesting due to their atmospheric significance. Interfacial QKE effects

that manifest themselves in droplets with tens to hundreds of water molecules could impact the interpretation of measured fractionation ratios in these experiments.

We are also interested in small water clusters in order to lay the groundwork for more complex density functional theory treatments of water systems. More computationally complex interaction potentials require smaller systems, and understanding how system size affects calculated QKE ensures that results can be interpreted properly.

3.4.1. System Setup. We ran PIMD 32 bead PRO thermostated simulations at 300 K. A 0.75 fs timestep was used, with 7.5 ps of equilibration and 750 ps of production, and using ring polymer contraction [13] on electrostatic interactions over 5 Å. Only the q-TIP4P/F [7] water model was simulated. Systems of 32, 48, 64, 96, 128, 256, and 512 H₂O molecules in a spherical droplet in vacuum with no PBC were investigated. A bulk PBC H₂O simulation and a vapor phase isolated single molecule simulation were also investigated with identical parameters. The droplet was constrained with a harmonic potential to the origin, applied to the centroid of each atom, with a force constant of 2×10^{-6} Hartree/Bohr.

This value of harmonic potential force constant was arrived at by looking at 32 water cluster trajectories with varying force constants and choosing the largest one that did not appreciably change the radial density profile of the droplet. Even though this was not observed here, it is possible that larger droplets will need to have lower constraining harmonic force constants since their edge water molecules inherently must sit farther up in energy in the harmonic potential.

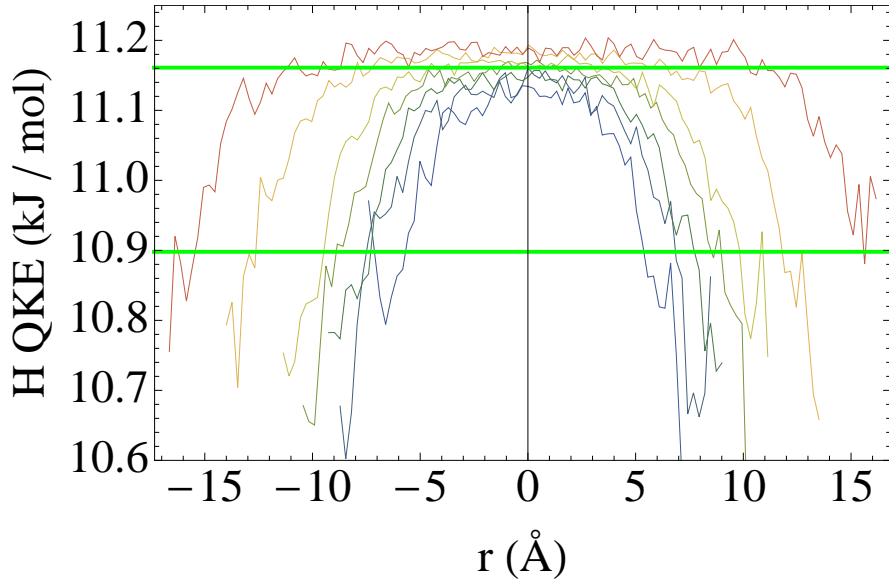


FIGURE 10. Hydrogen QKE as a function of radius from the center of water clusters with varying number of molecules. Lower green horizontal line is the vapor hydrogen QKE value; upper green horizontal line is the bulk PBC water simulation value. Simulations with more waters in the cluster show a monotonic increase in center hydrogen QKE values, but this does not approach the bulk value.

3.4.2. *Radial Dependence of QKE.* The results of all droplet sizes are presented in Figure 10.

It does appear that 32 water molecules are enough to see a plateau and conclude that the cluster reaches some sort of “bulk value” but it is interesting that there is a systematic increase in the QKE of center hydrogens as a function of cluster size. This is not a function of the harmonic constraining force perturbing the system, since doubling and halving the restraining force with 96 waters was found to not appreciably affect the center QKE values; turning off ring polymer contraction also had no effect. The fact that the cluster center QKE values are not approaching the bulk liquid value could be caused by the bulk simulations being run with PBC and long range Ewald forces being used.

It is interesting that the cluster simulations show a much clearer transition between the vapor hydrogen QKE value and the bulk liquid, unlike the slab interface simulations.

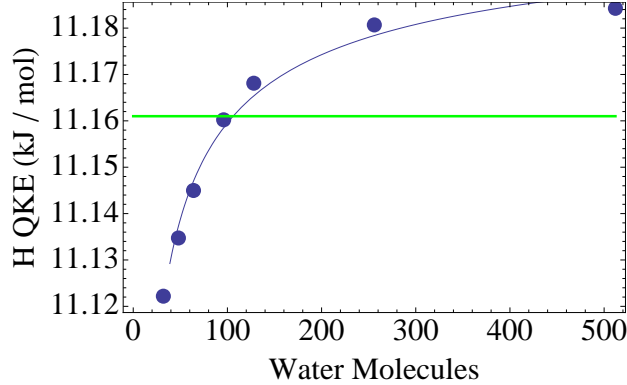


FIGURE 11. Cluster center hydrogen QKE as a function of number of waters in the cluster. A best fit line $11.21 - 0.502N^{-1/2}$, where N is the number of water molecules, is superimposed. The green horizontal line is the bulk PBC simulation hydrogen QKE value.

Plotting the cluster center hydrogen QKE values from Figure 10 as a function of cluster size, shows an approximate $N^{-1/2}$ dependence in Figure 11, where N is the number of molecules in the cluster. Further work is required to understand the origin of the center QKE dependence on system size. An extrapolation of the best fit line to $N \rightarrow \infty$ does not approach the bulk hydrogen QKE value of 11.15 kJ/mol. It is possible that this energy shift is due to long distance Coulombic forces being calculated by reciprocal space Ewald summation in the bulk simulation, but not in the isolated cluster simulations.

4. FUTURE DIRECTIONS

The work presented here highlights a number of immediate threads of future inquiry: Elucidating an explicit spatial and geometric relationship between hydrogen bonding configurations and QKE is necessary to confirm the spatial and electrostatic hypothesis presented

previously. Analyzing the influence of deuterium in intermolecular versus intramolecular interactions is necessary to begin understanding specifically how its concentration affects bulk liquid properties. It will also be interesting to include refinements to the definition of the liquid–vapor interface, calculate fractionation ratios due to interfacial QKE, and examine the specifics of cluster size effects in order to better understand the role of interfaces in fractionation.

Further work includes investigating fractionation around ions [9], the validation of density functional theory forces in PIMD simulations through fractionation ratios, and comparing to experimental fractionation work at various types of water interfaces.

REFERENCES

- [1] Jianshu Cao and Bruce J Berne. On energy estimators in path integral Monte Carlo simulations: Dependence of accuracy on algorithm. *The Journal of Chemical Physics*, 91(10):6359–6366, 1989.
- [2] R Car and Michele Parrinello. Unified Approach for Molecular-Dynamics and Density-Functional Theory. *Physical Review Letters*, 55(22):2471–2474, 1985.
- [3] Michele Ceriotti, Giovanni Bussi, and Michele Parrinello. Colored-Noise Thermostats à la Carte. *Journal of Chemical Theory and Computation*, 6(4):1170–1180, April 2010.
- [4] Michele Ceriotti, David E Manolopoulos, and Michele Parrinello. Accelerating the convergence of path integral dynamics with a generalized Langevin equation. *The Journal of Chemical Physics*, 134(8):084104, 2011.
- [5] Manolopoulos, David Ceriotti, Michele. Efficient first-principles calculation of the quantum kinetic energy and momentum distribution of nuclei. *Pre-print*, 2012.

- [6] D Chandler and PG Wolynes. Exploiting the Isomorphism Between Quantum-Theory and Classical Statistical-Mechanics of Polyatomic Fluids. *Journal of Chemical Physics*, 74(7):4078–4095, 1981.
- [7] Scott Habershon, Thomas E Markland, and David E Manolopoulos. Competing quantum effects in the dynamics of a flexible water model. *The Journal of Chemical Physics*, 131(2):024501, 2009.
- [8] Jochen Hoefs. *Stable Isotope Geochemistry*. Springer Verlag, 2009.
- [9] J HORITA, D COLE, and D WESOLOWSKI. The activity-composition relationship of oxygen and hydrogen isotopes in aqueous salt solutions: II. Vapor-liquid water equilibration of mixed salt solutions from 50 to 100 °C and geochemical implications. *Geochimica et Cosmochimica Acta*, 57(19):4703–4711, October 1993.
- [10] Damien Laage and James T Hynes. On the Molecular Mechanism of Water Reorientation. *The Journal of Physical Chemistry B*, 112(45):14230–14242, November 2008.
- [11] T E Markland and B J Berne. Unraveling quantum mechanical effects in water using isotopic fractionation. *Proceedings of the National Academy of Sciences*, 109(21):7988–7991, May 2012.
- [12] Thomas E Markland. Development and Applications of Path Integral Methods. pages 1–206, August 2009.
- [13] Thomas E Markland and David E Manolopoulos. A refined ring polymer contraction scheme for systems with electrostatic interactions. *Chemical Physics Letters*, 464(4-6):256–261, October 2008.
- [14] Francesco Paesani, Wei Zhang, David A. Case, Thomas E. Cheatham, and Gregory A. Voth. An accurate and simple quantum model for liquid water. *The Journal of Chemical Physics*, 125(18):184507–184507–11, 2006.
- [15] Michele Parrinello and A Rahman. Study of an F center in molten KCl. *The Journal of Chemical Physics*, 80:860, 1984.
- [16] D C Rapaport. *The Art of Molecular Dynamics Simulation*. Cambridge University Press, Cambridge, 2011.

- [17] L S Schulman. *Techniques And Applications of Path Integration*. Dover Pubns, December 2005.
- [18] H F Trotter. On the product of semi-groups of operators. *Proceedings of the American Mathematical Society*, 10(4):545–545, April 1959.
- [19] M Tuckerman. Statistical Mechanics: Theory and Molecular Simulation - Mark Tuckerman - Google Books, 2010.
- [20] Jiri Vanicek and William H. Miller. Efficient estimators for quantum instanton evaluation of the kinetic isotope effects: Application to the intramolecular hydrogen transfer in pentadiene. *Journal of Chemical Physics*, 127(11):–, 2007.
- [21] Adam P Willard and David Chandler. Instantaneous Liquid Interfaces. *The Journal of Physical Chemistry B*, 114(5):1954–1958, February 2010.