Combining Research in Theoretical Chemistry and Chemistry Education

Part A: The Development and Use of Guided Diffusion Monte Carlo to Study Molecules that Undergo Large Amplitude Vibrational Motions

Part B: Using Marzano's Taxonomy to Classify Multiple-Choice General Chemistry Exam Questions

Jacob Finney

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2023

Reading Committee:

Anne McCoy, Chair

Sotiris Xantheas

Stefan Stoll

Program Authorized to Offer Degree:

Chemistry

University of Washington

Abstract

Combining Research in Theoretical Chemistry and Chemistry Education

Part A: The Development and Use of Guided Diffusion Monte Carlo to Study Molecules that Undergo Large Amplitude Vibrational Motions

Part B: Using Marzano's Taxonomy to Classify Multiple-Choice General Chemistry Exam Questions

Jacob Finney

Chair of the Supervisory Committee:
Anne McCoy
Department of Chemistry

This dissertation is a combination of work in theoretical vibrational spectroscopy and chemical education. The research in the former focuses on studies of molecules that exhibit large amplitude motions. One such system is water clusters, which serve as intermediates between a single water molecule and bulk water. The addition of water molecules introduces hydrogen bonding, a strong intermolecular interaction that is an ongoing topic of study. An effective method for studying systems that undergo large amplitude motions is diffusion Monte Carlo (DMC), a method that can be used to obtain the ground state vibrational wave function and zero-point energy for an arbitrary system. The research in the latter part of this work includes developing and utilizing a rubric to classify the cognitive effort required by a student to solve multiple-choice general chemistry problems.

To begin, the diffusion Monte Carlo algorithm will be reviewed as well as extensions to guided DMC and the evaluation of excited states and molecular properties. The approach has been shown to require increasingly large ensembles when intra- and intermolecular vibrations

are weakly coupled. Our group recently proposed a guided variant of diffusion Monte Carlo to address these challenges for water clusters, $(H_2O)_n$. An extension of this approach is then applied to more strongly bound molecular ions, specifically CH_5^+ and $H^+(H_2O)_{n=1-4}$. For the protonated water systems, this work shows that the guided DMC approach that was developed for studies of $(H_2O)_n$ can be used to describe the OH stretches and HOH bends in the solvating water molecules, as well as the free OH stretches in the hydronium core. For the hydrogen bonded OH stretches in the H_3O^+ core of $H^+(H_2O)_n$ and the CH stretches in CH_5^+, adaptive guiding functions based on the instantaneous structure of the ion of interest were developed. Using these guiding functions, converged zero-point energies and ground state wave functions are obtained using ensemble sizes that are as small as 10% the size that is needed to obtain similar accuracy from unguided calculations.

The isomerization pathway between the energetically low-lying Zundel and Eigen isomers of the protonated water hexamer is then investigated using high-level *ab initio* calculations including a proper treatment of zero point corrections. On the basis of these calculations, the Zundel-Eigen isomerization was found to proceed through a stable intermediate isomer, which consists of a four-membered ring with two single acceptor water molecules. The inclusion of vibrational zero-point energy is shown to be important for accurately establishing the relative energies of the three relevant isomers involved in the Zundel-Eigen isomerization. Diffusion Monte Carlo calculations including anharmonic vibrational effects show that all three isomers of $H^+(H_2O)_6$ and $D^+(D_2O)_6$ have well-defined structures. The energetic ordering of the three isomers changes upon deuteration. Implications of these results for the vibrational spectra of $H^+(H_2O)_6$ and $D^+(D_2O)_6$ are also discussed.

Further exploring the spectra of $H^+(H_2O)_6$, a spectral map is created using relationships between vibrationally averaged hydrogen-bonded OH and OD distances and the frequencies and intensities associated with those vibrations. This allows spectral features to be predicted using vibrationally averaged distances obtained from DMC calculations. This spectral map-

ping method is compared against scaled harmonic and vibrational perturbation theory, as well as experiment. It is calibrated using smaller protonated water clusters which shows better agreement to experiment for vibrational modes with large anharmonicities. The spectral map is then used with four isomers of $H^+(H_2O)_6$ and $D^+(D_2O)_6$ to compare the spectral features to experiment.

Finally, Marzano's taxonomy was used as a basis for developing a rubric that specifically assessed the cognitive level of general chemistry multiple choice questions. The rubric was then applied to a bank of multiple-choice questions, which came from four University of Washington chemistry instructors who, though teaching a common syllabus from a common textbook, had idiosyncratic approaches to the material and how they presented it in class. An Inter-Rater Reliability of 92% for at least two out of three coders was obtained. This work shows that the cognitive level of most questions was quite low and that a solid majority of all questions were classified as "execution of a learned algorithm," which is consistent with trends noted in the literature on chemistry assessments.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

A Ph.D is no small undertaking, and as someone who is the first in my family to pursue a graduate degree, I made sure to get a lot of advice and help in order to finish it. I would like to thank all of the people who have helped me push through and help me continue to find the motivation to complete this degree.

First off, I would like to thank Dr. Anne McCoy. Thank you for always being on my side and working with me to finish this degree. You have helped me through countless times, both in science and in my personal life, where I hit a brick wall and did not know how to fix a problem I was having. From taking quantum from you in the first quarter here, to teaching classes with you throughout my time, you have continued to inspire me to become a teacher that not only teaches chemistry, but also one who teaches more fundamental skills such as problem solving. Even though I can't remember the definition of kinetic energy, thank you for always taking the time to help me slow down and actually understand the underlying science.

I would also like to thank my other advisor, Dr. Colleen Craig. I am very grateful to have joined your research group after my general exam. Learning that there is research that focuses on the part of chemistry that I have always enjoyed, teaching it to others, has really expanded my view and inspired me to learn as much as I can about the frontier of this field. Thank you for helping me on my path to be the best instructor that I can and to critically think about ways to pursue that. I greatly appreciate the in depth discussions we have had about the minutiae of general chemistry exam problems.

I would like to thank all of the members of the McCoy group, both past and present, who have been very welcoming and supported my habits of procrastination. Dr. Victor Lee,

Dr. Mark Boyer, and Dr. Ryan DiRisio, you all taught me so much and I loved being able to work closely with each of you as I grew as a scientist. I would also like to thank Chloe Chiu, Coire Gavin-Hanner, Mickey Moonkaen, Yarra Hassan, and Greta Jacobson. It was a pleasure working with you all and I appreciate all of your support during my degree. To Dr. Rachel Huchmala, I could not have done this degree without you. I am extremely grateful to have joined this group with you and I deeply appreciate your friendship. I would also like to thank the members of the Xantheas group, Dr. Joseph Heindel, Kristina Herman, Garrett Santis, and Max Hoffman. With all of you, I never had a reason to leave the fourth floor of Bagley.

I would also like to thank all the people I met through my fellowship with the Clean Energy Institute. I would like to give special thanks to Dr. Danica Hendrickson, Catherine Matson, Madelynne Zornes, and Ricardo Rivera. The outreach that I was able to do with all of you brought back my spark for working with younger students like I used to do in high school. The energy that you all brought made all of the work that we put in worth it.

I would like to thank all of my friends who helped keep me sane during these long years. Bailey Dugan, Jacob Gross, Joel Thomas, and Kyle Bendebel, I am extremely grateful to all of you. I hope that we can spend more time playing table-top games in the future.

I want to thank my mom, Erin Cleary, and my dad, Kenn Cook, who both raised me to have faith in myself and to set high goals. I would like to thank Ethan and Dylan Cook, even though we did not see each other much, you both always expressed support. I would like to thank Charlie Finney, I am glad we got to live together for some of my time here in Seattle. It was amazing to see you grow and I am glad that we were there for each other during COVID lockdown.

I would like to express my gratitude to my wife's family, Kimmberly and James Krevitz, Jason Biehner, Sadie Williams, and all of the kids: Raven and Tannon Biehner, MiaElena, Elias, Elijah, and SophiaLynn Krevitz. You all may be much louder than my family, but

you have all made sure that I was supported and loved the entire time I have known you. I appreciate you all greatly and I couldn't ask for better in-laws.

Finally, I would like to thank my wonderful wife, Justine Wilson. I cannot even begin to express my gratitude to you for journeying through this with me. It has been a tough five years, especially with that whole COVID thing, and I am very glad that you, and our precious feline boy Edward, have been with me. I am very excited for the next chapter in our life now that I am finally done with this whole being a student thing.

DEDICATION

To my loving and supportive wife, Justine

I hope to have many adventures with you in the future.

Chapter 1

INTRODUCTION

1.1 Diffusion Monte Carlo: A Method for Studying Large Amplitude Motions

Developing and implementing general approaches for solving the Schrödinger equation for vibrational systems is a long-standing challenge in theoretical spectroscopy. For molecules and ions where the vibrations are weakly coupled and the wave functions for the states of interest are localized in a single minimum in the potential energy surface, methods based on a harmonic treatment can provide a good starting point for more sophisticated methods. These include vibrational perturbation theory,[16,17] or vibrational self-consistent field (VSCF)-based approaches,[18,19] and extensions built off of these approaches.[20] Such approaches have been extended to studies of molecular clusters, for example $(H_2O)_n$, in which the Hamiltonian based on the intramolecular degrees of freedom is solved for a fixed cluster geometry.[21] The situation becomes more challenging when the molecular system of interest displays large amplitude vibrations, which allow it to sample multiple minima on the potential surface even in the vibrational ground state. Additional complications emerge when these large amplitude vibrations are strongly coupled to other vibrational degrees of freedom.

A method that circumvents many of these challenges is diffusion Monte Carlo (DMC).[22–27] In this approach, a large ensemble of localized functions, called walkers, is used to provide a Monte Carlo sampling of the ground state wave function. The coordinates and the relative weights of the walkers are propagated based on the imaginary-time time-dependent Schrödinger equation. The long time, equlilibrated ensemble obtained from this propagation provides a Monte Carlo sampling of the ground state wave function of the system of interest. The associated zero-point energy is obtained by imposing the requirement that the amplitude of the wave function does not change during the propagation.

The DMC approach has been used with great success to study a broad range of molecular systems that would have been very challenging to study by approaches that rely on a basis set to expand the wave functions.[8,28–30] An example of such a molecule is CH_5^+. The ground state wave function of this ion has amplitude at the 120 symmetry equivalent minima on the potential surface as well as near the 180 saddle points that connect these minima.[31,32] At the same time, the harmonic CH stretch frequencies span the range of roughly 2500 to 3300 cm^{-1} at each of the three low-energy stationary point structures.[5] This indicates that the frequency of the CH stretch is sensitive to its environment within the ion. Similarly, the frequencies of the water-bound OH bonds in protonated water clusters, $H^+(H_2O)_n$, are sensitive to their environments. This is reflected in the breadth of the features in the vibrational spectrum that are assigned to this motion in various size-selected protonated water clusters.[33] The structures and frequencies of water molecules that make up a water cluster, $(H_2O)_n$, also show sensitivities to the structure of the cluster, although the couplings between the intra- and intermolecular vibrations are weaker in this case compared to the protonated water clusters.[4]

1.2 Guiding Functions

In the case of the neutral water clusters, Mallory and Mandelshtam showed that very large ensembles of walkers were required to obtain reliable results using standard DMC approaches. For example, in their study of $(H_2O)_6$, they used an ensemble in excess of 10^6 walkers, run for 8.8×10^4 time steps.[34] In DMC simulations the potential energy of each walker needs to be evaluated at each time step. As potential functions become more accurate, they often also become more expensive to evaluate. Thus, these requirements of large ensembles to obtain accurate results make the DMC approach prohibitive, and limits the size and types of systems that can be studied using DMC.

As we explored the origins of this behavior,[27,35] we concluded that these challenges of applying DMC to studies of neutral water clusters reflected the fact that the shifts in the frequencies of the OH stretching vibrations with the environment were comparable in size

to the frequencies of the low-frequency vibrations in these clusters. Additionally, when all degrees of freedom are included in the calculation, the instantaneous equilibrium structure of the cluster as a function of the intermolecular degrees of freedom is sensitive to the precise OH bond lengths and HOH angles of the individual water molecules that make up the cluster. The need for a very large ensemble to converge this calculation reflects difficulties in sampling both the intermolecular and intramolecular degrees of freedom effectively. An additional complication arises from the fact that the time step used in the simulation needs to be chosen based on the shortest vibrational period in the system. At the same time the simulation needs to be run long enough to sample the motions in the degrees of freedom that are associated with the low-frequency vibrations. These considerations contributed to the very large ensemble sizes Mallory and Mandelshtam found they needed to fully sample the ground state of water clusters.

On the other hand, while the above considerations make these clusters particularly challenging for DMC calculations, they represent a situation for which an adiabatic separation of the low- and high-frequency vibrations provide a good zero-order description of the vibrational dynamics.[21,36] In this approach, rather than focusing on the instantaneous values of the intramolecular degrees of freedom that are associated with the high frequency vibrations, the potential that is used to study the intermolecular vibrations is the average of the full potential over the probability amplitude associated with the wave function that describes the high frequency vibrations. More specifically, the vibrational energy and wave function associated with the intramolecular degrees of freedom are obtained using the potential evaluated at an instantaneous cluster geometry. Taken together, these energies are used to construct an adiabatic potential where the energy dependence on the intermolecular degrees of freedom reflects the vibrationally averaged behavior of the high frequency modes rather than their instantaneous values. Such an approach expressed in this form would be challenging for studies of such clusters using DMC, as each potential energy evaluation for the intermolecular degrees of freedom would require averaging the full potential over the ground state wave function in the intramolecular vibrations. Additionally, such a treatment would render the

results of the calculation at best approximate, removing one of the attractive features of DMC, e.g. its use for obtaining an exact description of the ground state wave function, and the associated zero-point energy of the system of interest.

In work by a previous student in the group,[27,35] they demonstrated that we could achieve a similar separation of time scales using DMC through the introduction of carefully developed guiding functions, Φ_T. We applied the approach to studies of water clusters and demonstrated an order of magnitude saving in ensemble sizes used for the simulations without a loss of accuracy. Specifically, for water clusters, these guiding functions were expressed as direct products of the solutions to one-dimensional cuts through the potential surface of an isolated water molecule along each of the two OH stretch coordinates along with a harmonic description of the bend. The introduction of the guiding function in the DMC calculation has the effect of replacing the potential function with the local energy,

$$E_{\mathrm{L}} = \frac{H\Phi_T}{\Phi_T} \tag{1.1}$$

Our choice for Φ_T has the effect of reducing the dependence of E_{L} on the intramolecular degrees of freedom compared to that of the full potential function. For example, if the OH stretch component exactly matched that of one of the OH oscillators of the water molecule in the cluster then for that cut through the potential, the local energy is constant, and the average value of the local energy for that coordinate is independent of how the walkers sample that coordinate. By removing the dependence of the local energy on the high frequency vibrations, E_{L} in Eq. 1.1 resembles an adiabatic potential that only depends on the intermolecular coordinates. Further by expressing the trial wave functions as direct products of one-dimensional wave functions, the evaluation of the E_{L} does not have a significant impact on the overall expense of the DMC calculation. On the other hand, the introduction of E_{L} does not introduce any approximations to the treatment of the ground state wave function. By removing the dependence of E_{L} on the OH bond lengths and HOH bend angles we found that we could realize substantial savings due to the ability to use much smaller ensemble

sizes.

This approach differs from earlier studies on hydrogen clusters[37,38] and H_5^+,[39] which used normal mode descriptions for all of the vibrational degrees of freedom. Since the normal modes are evaluated for a reference structure, such an approach makes it difficult to describe systems that have multiple low-energy minima on the potential, as a Φ_T that is based on one of minima may not provide a good description of the ground state wave function near other minima. By focusing the trial wave function on only the high frequency vibrations, we remove this bias from our sampling at the expense that we are only describing a subset of the degrees of freedom. Such an approach was shown to work quite well for the $(H_2O)_6$ cluster in which there are several low-energy minima on the potential.[35]

1.3 Comparisons to Vibrational Spectroscopy

While Diffusion Monte Carlo studies offer many insights into systems that are difficult to study by other approaches, DMC is a ground state method. This introduces serious limitations when considering vibrational spectra due to the lack of information about the excited state that is required to identify the frequency and intensity of transitions and assign the spectrum. There have been a few approaches that utilize the high quality wave functions that can be obtained from diffusion Monte Carlo in order to obtain excited state information. One approach is the fixed-node approximation which has been applied successfully in many systems such as $H_5O_2^+$,[40] $H_3O_2^-$,[41,42] H_5^+,[43] complexes of two neon atoms with SH or OH,[44] and recently $C_2H_5^+$,[45] all of which contain large amplitude motions. This approximation is based on the nodal surface being described in a normal mode model. When describing vibrations as normal modes, the node is only along the coordinate, or linear combination of coordinates, that the excitation is in. This method has shown good agreement with both experiment and variational calculations. However, in these calculations the node has to be placed in the correct position. Symmetry can often be used to place the node, for example in the antisymmetric stretch in water. If the node can not be determined by symmetry an adiabatic DMC calculation can be used to find the optimal node location for a given coordi-

nate.[46–49] In all of these cases, multiple DMC calculations need to be run, one for the ground state, and at least one calculation for each vibrational mode of interest. For large systems this can quickly become unfeasible. A larger limitation is that this method can calculate the frequencies in specific transitions but the intensities are much more difficult to obtain. There has been an approach to obtain matrix elements involving excited states, but the method is extremely computational expensive.[42] In Chapter 2 we also discuss a similar approach with guiding functions where an excited state guiding function is used in the vibrational mode of interest to obtain not only frequencies, but also overlaps between the ground and excited state wave functions.

Another approach for obtaining excited state information using DMC uses the ground state probability amplitude to approximate integrals involving the ground state and excited state wave functions. These integrals are then used to calculate the transition frequencies and transition moments required to predict spectra. This approach is outlined in two studies done on $H_3O_2^-$ and $H_5O_2^+$.[41,50] The method was also used on H_5^+ as well as $H^+(H_2O)_{3-4}$ and $D^+(D_2O)_{3-4}$,[43,51] and it showed good agreement to both experiment and other theoretical methods. However, the method requires a careful selection of internal coordinates for the system, which can get complicated as system size increases, or with the introduction of ring-like structures.

A related, and somewhat simpler approach, involves mapping structural features to spectral signatures. For example, in a study of changes in the spectrum of the protonated water tetramer upon complexation with H_2, N_2, CO and one or two water molecules Wolke *et al.* showed that the frequencies of the hydrogen-bonded OH stretching vibrations could be correlated to the distance between the oxygen atoms in the donor and acceptor water molecules.[7] While such a correlation is not surprising, it provides an approach for mapping the spectrum onto structural parameters. Likewise, through a series of studies, Xantheas and co-workers illustrated a linear relationship between the shift in the OH bond length and the shift in the frequency of the OH oscillator,[52,53] demonstrating a similar correlation to the one made by Badger that related shifts in diatomic bond lengths and the force constant.[54] The slope

(-19.1 cm^{-1} / 0.001 Å) was found to be independent of the level of electronic structure theory used for the calculation and to whether the harmonic frequency is related to the equilibrium bond length or the anharmonic frequency is related to the zero-point averaged bond length.[4] In a more recent study, Xantheas and co-workers identified a similar relationship for the hydronium ion.[55] Interestingly, the proportionality constant of approximately -19 cm^{-1} / 0.001 Å applies to both water and hydronium. In analyzing these relationships for water clusters, we showed that we could relate the linear relationships between the OH bond lengths and frequencies to the response of these quantities to the local electric field,[4] thereby connecting the above treatment to the spectral mapping approaches that Skinner, Geissler and others have used to interpret the spectrum of water, both in bulk and for smaller water clusters.[56–61]

The above relationship only provides information about the expected positions of peaks in the spectra. To model the spectrum also requires information about the intensities of these transitions. Work on hydrogen-bonded systesm, particularly water clusters has shown that there is also a linear relationship between the intensities and the frequencies of the OH stretching transitions.[62,63] Additionally Kananenka and Skinner showed that they could extend their spectral maps to include intensity information.[58] They used these maps to reproduce the spectrum of the book form of water hexamer that had been reported by Johnson and co-workers.[64] These results indicate that we should be able to obtain a good approximation with the anharmonic vibrational frequencies of OH oscillators in either hydronium or water from the zero-point averaged OH bond lengths obtained using the accurate ground state probability amplitudes, for example those obtained from Diffusion Monte Carlo calculations.

1.4 Chemistry Education

The type of assessment that an instructor employs sends a strong message to students about what kind of reasoning the instructor values,[65–68] as well as the kinds of students that can succeed in the course.[69] Therefore, if an instructor values higher order thinking, they should design assessments that explicitly require students to exhibit higher order thinking. This

sounds very obvious, but instructors can easily and unintentionally undermine a course that they have carefully constructed to elicit higher order thinking by composing test questions that only elicit lower level thinking.[69–71] Indeed, in higher education chemistry courses, many assessments stress mathematical skills rather than chemical reasoning,[68,69] and focus on relatively low level cognitive skills.[69] In addition, recent work suggests that chemistry assessments that focus on algorithmic math skills may exacerbate existing patterns of inequity, due to the inequitable distribution of pre-college mathematics preparation.[69,72–74] To encourage students to start thinking like practicing chemists,[68] and to level the playing field for students from societal groups with less access to pre-college mathematics preparation,[69] assessments should require higher order cognition skills.

One way to approach this misalignment of instructor intention and the quality of exam items is to classify the questions on exams according to the cognitive level that they require, then using the questions that target the desired cognitive level of the instructor. A popular classification scheme for this process is Bloom's Taxonomy of Educational Objectives shown in Figure 1.1.[1] Although it was originally developed to help educators design learning objectives, it has also been used to assess cognitive level of exam questions in many higher education STEM contexts: biology,[66,70,71,75] chemistry,[76,77] physics,[66] and mathematics.[78] The original version of Bloom's taxonomy used six levels to categorize types of thinking: Knowledge, Comprehension, Application, Analysis, Synthesis, and Evaluation. A notable revision to this taxonomy was put forward by Anderson et al.,[2] in which the language of the taxonomy was modified from nouns to verbs to focus on the actions required at each cognitive level shown in Figure 1.1: Remember, Understand, Apply, Analyze, Create, and Evaluate. To futher delinieate the levels of cognition, this revision included sublevels for each of the levels above. In addition, a second, "knowledge" dimension was added to the taxonomy to allow a distinction between Factual, Conceptual, Procedural, and Metacognitive knowledge that the cognitive dimension acts on.

While Bloom's taxonomy has been used with much success in many contexts,,[66,70,71,75,76,78] there are some problems with this taxonomy, especially in its application to questions that

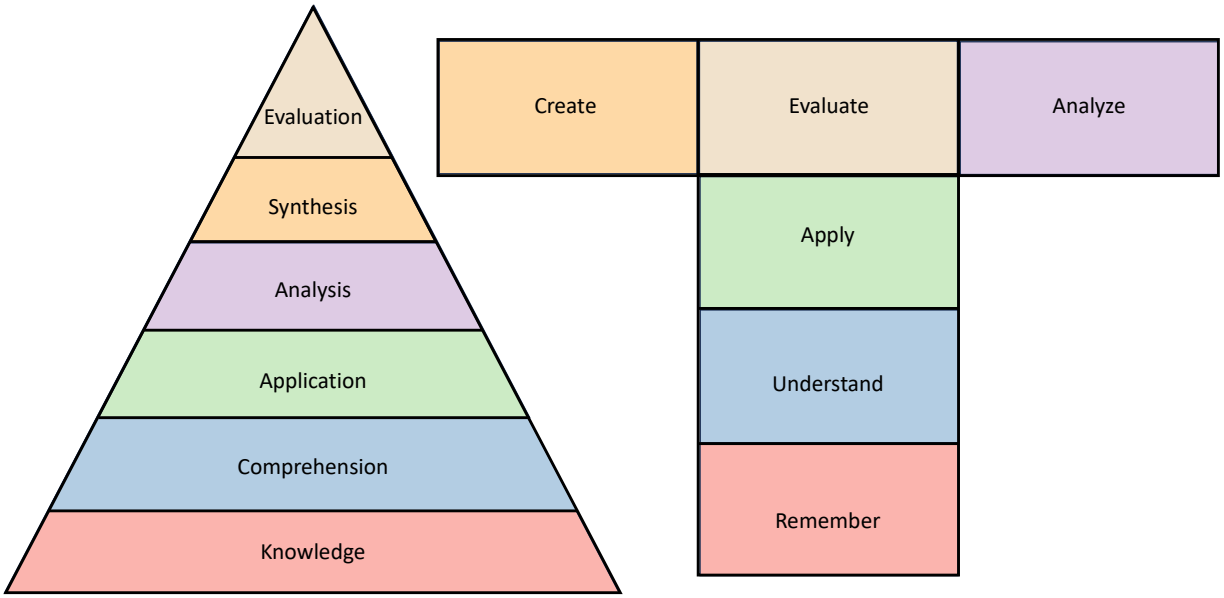


Figure 1.1: A common representation of Bloom's taxonomy (left).[1] A representation of the revision of Bloom's taxonomy by Anderson et al. (right).[2,3]

are common in general chemistry. One problem arises when assigning the cognitive level to an algorithmic type question, such as reaction stoichiometry, or Lewis dot structures. In Bloom's taxonomy such questions would fall under the sublevel Executing which is in the third cognitive level, Apply.[2] There has been recent work separating Bloom's taxonomy into "higher" and "lower" levels reffered to as higher-order cognitive skills (HOCS), which includes analysis, synthesis, and evaluation, and lower-order cogntive skills (LOCS), which includes knowledge, comprehension, and application.[70,77] However, the LOCS are typically treated as hierarchical for the first three levels, as shown on the right of Figure 1.1.[3] This means that an algorithmic type question—in which a student follows a set of pre-established steps—would be classified at a higher cognitive level than the second level, Understand. The essence of this level is sense-making: taking in new information, making sense of it, and then communicating one's new understanding in some form (verbal explanation, a figure or diagram, etc.).

Questions that only require the execution of an algorithm do not require the student to demonstrate any conceptual understanding, only that they remember and correctly apply the steps necessary to "run" the algorithm. These types of questions are often colloquially called "plug-and-chug" type questions, because students can take the values provided in the question statement and "plug" them into a known equation or relationship and then "chug" through the calculation to obtain an answer, all without using any problem solving skills or content-specific information. In other words, such ostensibly "chemistry" questions can often be solved using mathematical recipes alone.[69,79] In addition, general chemistry students typically receive considerable practice on these types of problems through homework or in-class examples.[68] All of this practice lowers the cognitive load for students to employ the related algorithms to obtain the answers, especially for those who are comfortable with the math involved in answering such questions. By placing more weight on algorithmic problems, students with greater pre-college mathematics preparation are at an advantage.[68,69,73] Ralph et al.[69] showed that assessments that emphasize mechanistic reasoning skills as opposed to purely mathematical assessments had a positive impact on predicted pass rates based on

students' performance on all tasks for Black and Latinx STEM students. Therefore, using a taxonomy that places more weight on conceptual learning is helpful in analyzing and creating assessments that focus on eliciting higher order thinking.

There have been other notable developments of taxonomies in addition to Bloom's Taxonomy. One example is the Structure of the Observed Learning Outcomes (SOLO) taxonomy.[80–82] This taxonomy was designed with the intent of aligning curricula with learning objectives. It has five levels that describe a way of assessing how much understanding the student has demonstrated but not necessarily the process that the student used to obtain the answer. Another example that is specific to chemistry is the Expanded Framework for Analyzing General Chemistry Exams by Smith el al.[83] This was based on the work of Zoller in identifying higher-order cognitive skills and lower-order cognitive skills in chemistry, as well as other research in chemistry assessment.[2,77,84–88] This framework is broken up into three primary levels; Definition, Algorithmic, and Conceptual. These levels are further delineated into secondary levels that describe specific types of chemistry problems that the framework was built around. Due to the framework being constructed based on chemistry problems, the levels can have overlap and some questions can have aspects of multiple levels. In this work we employed another taxonomy, Marzano's Taxonomy of Educational Objectives to assess the cognitive levels of general chemistry problems.

1.5 Overview of Work

Diffusion Monte Carlo is a powerful method when investigating systems that display large amplitude motions. In Chapter 2, we will first describe the details of this method and the extension to guided diffusion Monte Carlo. We will use a one-dimensional model to illustrate the changes that happen with the introduction of a guiding function as well as extensions to calculating molecular properties.

Guided diffusion Monte Carlo relies on the use of a guiding function that describes some or all of the vibrational modes within the system. A general guiding function for neutral water clusters was developed in a previous pair of studies that showed a large reduction in the

needed simulation size without a reduction in the quality of the ground state wave function and ground state energy.[27,35] This guiding function is made of a product of one-dimensional wave functions that describe the high frequency, intramolecular modes water clusters. Chapter 3 describes an extension of this work to systems that exhibit large amplitude motions; $H^+(H_3O)_{1-4}$ and CH_5^+. Due to the large anharmonicity in OH stretches in protonated water clusters and CH stretches in CH_5^+, the guiding function for these modes was modified based on their local environment.

Using the guiding function developed in Chapter 3 for protonated water clusters, larger water clusters which were intractable for DMC simulations are able to be studied due to the smaller simulation sizes that can be used. One such system is the protonated water hexamer where the spectrum of the deuterated isotopologue shows a different band structure, suggesting a different population of isomers upon deuteration.[7,12] Previous double resonance experiments showed that the spectrum of $H^+(H_2O)_6$ was primarily a mixture of two isomers.[12] In Chapter 4, we use a combination of electronic structure and diffusion Monte Carlo to investigate the differences in zero-point energy of three isomers of $H^+(H_2O)_6$ and $D^+(D_2O)_6$. We also use vibrational perturbation theory to determine which of three isomers have characteristic peaks to determine the isomer that could have significant population once deuterated.

Chapter 5 focuses on using the spectral mapping to further explore the isomers that have vibrational frequencies that could be identified in the $D^+(D_2O)_6$ spectrum. This approach uses the formerly observed relationship between hydrogen-bonded OH distances and frequencies to identify the peak positions based on vibrationally averaged OH distances using the DMC wave functions.[4,52,53] This work also establishes this relationship for hydrogen-bonded OD stretches to compare to the deuterated spectrum. Correlations between hydrogen-bonded OH/D distances and intensities are also created and the spectral map is validated with smaller protonated water clusters where the spectra are simpler.

Finally, in Chapter 6, we use Marzano's taxonomy to develop a rubric to categorize the cognitive level of general chemistry multiple choice questions. Marzano's taxonomy is

explained in the context of multiple choice questions. The rubric is then used to rate questions within a large test set of multiple choice questions pulled from tests from the first quarter of a three quarter general chemistry class. The analysis of the percentage of questions in each level highlights the need for analyzing the cognitive effort that a student requires to answer questions correctly.

Chapter 2

DIFFUSION MONTE CARLO APPROACHES FOR STUDYING LARGE AMPLITUDE VIBRATIONAL MOTIONS IN MOLECULES AND CLUSTERS

Preprint of an chapter published in [Jacob M. Finney, Ryan J. DiRisio, and Anne B. McCoy. Diffusion Monte Carlo Approaches for Studying Large Amplitude Vibrational Motions in Molecules and Clusters. In *Vibrational Dynamics of Molecules* Bowman, J. M.; World Scientific Publishing, New Jersey, **2022**.; pp 145-173]

2.1 Introduction

As we consider vibrational dynamics and how it is expressed through spectroscopy, we often rely on zero-order models that provide semi-quantitative descriptions of the system and processes of interest. When successful, analysis of these models provides a set of quantum numbers that can be used to assign the spectra. Such models are also used to motivate the choice of basis sets that are used for more sophisticated treatments of the vibrational dynamics. For example, for molecules that remain localized near a single minimum in the potential, a zero-order description based on uncoupled oscillators can be extremely effective. When larger amplitude motions correspond to hindered rotations or internal rotations, such motions can be incorporated into the model. While such approaches are often very effective, when molecules sample multiple low-energy minima on the potential surface or contain strongly coupled, large amplitude vibrational motions, identifying an appropriate zero-order description can become challenging. Additionally, the approaches that are taken to calculate the spectra can become system specific.

On the other hand, understanding properties of the ground state, and possibly low-energy

excited states, of such molecular systems can often provide significant insights into how it samples its potential surface. Such insights can be used both for the interpretation of spectra and to motivate other approaches. A method that is well-suited for such studies is Diffusion Monte Carlo (DMC).[22,24] This approach, which was credited to Fermi by Metropolis,[89] exploits the fact that when the time-dependent Schrödinger equation is expressed in imaginary time ($\tau = it/\hbar$)

$$\frac{d}{d\tau}\Psi(\mathbf{x},\tau) = \left[\frac{\hbar^2}{2}\sum_{k=1}^{n_{\text{atom}}}\frac{1}{m_k}\nabla_k^2 - V(\mathbf{x})\right]\Psi(\mathbf{x},\tau) \tag{2.1}$$

has the same structure as the diffusion equation with a coordinate-dependent source-sink term

$$\frac{dC(\mathbf{x}(t))}{dt} = [D + k(\mathbf{x})]\, C(\mathbf{x}(t)) \tag{2.2}$$

Just as the diffusion equation can be solved by Monte Carlo approaches based on the evolution of coordinates of the particles that are diffusing, if we express the solution to the imaginary-time, time-dependent Schrödinger equation, $\Psi(\mathbf{x},\tau)$, as an ensemble of localized objects, called walkers, the computational machinery that has been used to obtain the equilibrium solution to the diffusion equation can also be applied to the obtaining the equilibrium solution to the imaginary-time, time-dependent Schrödinger equation. While this may seem like a strange thing to calculate, $\lim_{\tau\to\infty}\Psi(\mathbf{x},\tau) = \Psi_0(\mathbf{x})$, which is the ground state solution to the corresponding time-independent Schrödinger equation. DMC has been used to study a broad range of chemical problems.[27,34,35,39,90–100] While much of the work has focused on solving the electronic Schrödinger equation, a significant challenge introduced in this application is the treatment of the nodes brought about by the requirement that the wave function must be antisymmetric with respect to permutation of the electrons. In some senses, though, DMC is better suited to vibrational problems, where the ground state is nodeless. Suhm and Watts,[24] and Buch,[46,91,101] pioneered work in this area. Their studies demonstrated the variety of insights that can be obtained from DMC calculations. Much of their work focused

on hydrogen-bonded clusters, where the interacting molecules were constrained to be rigid.

Over the past fifteen years, the increased interest in exotic molecules, like H_5^+[39,102–104] and CH_5^+[31,94,105–107] have led to the need for methods that allow us to study molecules that have multiple strongly-coupled large amplitude vibrational motions. Critical to increased interest in these systems has been the availability of potential surfaces for a broad range of molecules that undergo large amplitude motions[5,32,39,108] and the development of flexible general potentials,[8,109] which can be used to explore nuclear quantum effects in assemblies of water molecules and protonated water clusters. Unlike the electronic structure problems, where the potential can be described by pairwise Coulombic interactions, once the Born-Oppenheimer approximation is invoked, the form for the vibrational potential loses its simplicity.

As we move from molecules with ten or fewer atoms to larger molecular clusters, we have found that the size of the ensemble that is needed to effectively sample the potential grows rapidly, and modifications to the DMC approach are needed. Over the past several years, we have explored a modified version of Diffusion Monte Carlo, which employs guiding functions, to improve the sampling. In contrast to much of the previous work in using guiding functions in studies of vibrational problems, which focused on developing general functions of the intermolecular degrees of freedom,[39,100,110] this work focuses on the development of guiding functions, which are functions of individual bond length and angles. We have applied this approach to water clusters with up to six water molecules,[27,35] protonated water clusters, and molecules like CH_5^+ that undergo large amplitude motions.[97] In the next section of this chapter, we will describe this extension and explore some of the considerations that need to be be accounted for in applying the approach. We explore the method in the context of calculations on water, water hexamer and CH_5^+.

2.2 Diffusion Monte Carlo (DMC)

Diffusion Monte Carlo provides a general approach for obtaining the ground state wave function and zero-point energy for an arbitrary molecular system.[22–24,26] To develop the

working equations for DMC, the time-dependent Schrödinger equation is expressed as a propagator in imaginary time $\tau = it/\hbar$

$$|\Psi(\tau)\rangle = \exp[-\hat{H}\tau]|\Psi(0)\rangle \tag{2.3}$$

At long times, $|\Psi(\tau)\rangle$ converges to the ground state solution to the corresponding time-independent Schrödinger equation, $|\Psi_0\rangle$, and the amplitude of the wave function decays as $\exp[-E_0\tau]$, where E_0 is the ground state energy. The simplest DMC algorithm can be developed by approximating the propagator in Eq. 2.3 by a split operator, and shift the potential energy by E_{ref}

$$|\Psi(\tau)\rangle = \exp[-(\hat{V} - E_{\text{ref}})\tau]\exp[-\hat{T}\tau]|\Psi(0)\rangle \tag{2.4}$$

This shift is introduced as a parameter that is determined by the requirement that the amplitude of $|\Psi(\tau)\rangle$ remains constant. At the end of the simulation, the time-averaged value of E_{ref} provides the ground state energy for the system of interest. In contrast to other approaches discussed in this volume, where the wave function is expanded in a carefully chosen basis set, in DMC we represent the wave function by a weighted sum of localized functions, $\{g(\mathbf{x} - \mathbf{x}_i)\}$, commonly referred to as walkers

$$\Psi(\tau) = \sum_{i=1}^{N_{\text{w}}} w_i(\tau) g(\mathbf{x} - \mathbf{x}_i(\tau)) \tag{2.5}$$

where $\mathbf{x}$ represents the Cartesian coordinates of the atoms in the molecular system of interest, $\mathbf{x}_i$ is the geometry at which g is localized, and w_i provides the associated weight in the context of an ensemble of N_{w} walkers. As noted in Eq. 2.5, both the positions and weights of the walkers evolve according to Eq. 2.4.

Operationally, the simulation is divided into a series of short time steps, $\Delta\tau$. During each time step, the Cartesian coordinates of each of the atoms in each of the walkers are

displaced based on a Gauss-random distribution with a width $\hbar(\Delta\tau/m_j)^{1/2}$, where m_j is the mass of the jth atom.[22] This provides an approximation to the action of the kinetic energy contribution to the propagator in Eq. 2.4 on one of the localized functions. Next, the potential energy is evaluated at the coordinates of each of the displaced walkers, and

$$P_i(\tau + \Delta\tau) = \exp[-(V(\mathbf{x}_i(\tau + \Delta\tau)) - E_{\text{ref}}(\tau))\Delta\tau] \tag{2.6}$$

is used to obtain $w_i(\tau + \Delta\tau)$.

Two approaches are used to update the weights. In the first, termed continuous weighting,

$$w_i(\tau + \Delta\tau) = P_i(\tau + \Delta\tau)w_i(\tau) \tag{2.7}$$

This approach keeps the ensemble size constant, but can lead to numerical challenges as a small number of walkers can carry most of the weight. As a result, even for a large ensemble, as the simulation progresses the sampling can deteriorate. To circumvent this problem, we impose a range that the weights must remain within. If the weight of a walker becomes smaller than the minimum allowed weight, that walker is removed from the simulation and the walker with the largest weight is replicated. The weight of the replicated walker and its replica are each half the weight of the high-weight walker prior to replication. This ensures that the sum of the weights of the walkers and the number of walkers remain constant. A similar procedure is applied when the weight of a walker exceeds the maximum value. We have found that a minimum weight of 0.01 and a maximum weight of 20 provide stable simulations.[97] The second approach, referred to as discrete weighting, requires that all of the walkers have $w_i = 1$. In this case, after displacement $\text{int}(P_i(\tau + \Delta\tau))$ is used to determine the number of walkers that should have the coordinates of the ith walker at the start of the next iteration of the propagation. The fractional part of $P_i(\tau + \Delta\tau)$ is compared to a random number between 0 and 1. If $P_i(\tau + \Delta\tau)$ is larger than the random number, an additional walker is placed at these coordinates. This approach has the advantage that all

walkers contribute equally to the wave function, often leading to more efficient sampling, at the expense that the size of the ensemble fluctuates throughout the simulation.

The above procedure requires the evaluation of $E_{\mathrm{ref}}(\tau)$. Following Anderson,[23]

$$E_{\mathrm{ref}}(\tau) = \overline{V}(\tau) - \alpha \frac{W(\tau) - W(0)}{W(0)} \tag{2.8}$$

Here $\overline{V}$ provides the ensemble average of the potential energy, while $W(\tau)$ provides the sum of the weights of the walkers. The second term in Eq. 2.8 introduces a penalty when the $W(\tau)$ deviates from $W(\tau = 0)$, where α is a simulation parameter, which controls the size of the fluctuations in E_{ref}. The effect of the size of α on the time evolution of E_{ref} has been discussed in a recent review (see Fig. 6).[111] We have found that a value of $\alpha = 0.5\hbar/\Delta\tau$ often provides a good choice for this parameter.[98,112] As noted above, the time-averaged value of E_{ref} provides the energy of the state of interest.

2.3 Introducing Guiding Functions

A common modification to the standard DMC algorithm is to introduce a guiding function, Φ_{T}.[25,113] The introduction of well-chosen guiding functions can improve the efficiency of DMC simulations, but Φ_{T} needs to be chosen with care.[97,100] As we will illustrate, one of the roles of the guiding function is to shift the walkers toward regions of configuration space where Φ_{T} has amplitude. Consequently, a poor choice of Φ_{T} can introduce bias and lead to poor results. At the same time, the potential energy is replaced by the local energy,

$$E_{\mathrm{L}} = \frac{\hat{H}\Phi_{\mathrm{T}}}{\Phi_{\mathrm{T}}} \tag{2.9}$$

If $\Phi_{\mathrm{T}} = \Psi_0$, the local energy will be the ground state energy, and E_{ref} will be constant and equal to the ground state energy. This is not an interesting situation, as if we already know Ψ_0 there is no need to use DMC to evaluate it. More typically $\Phi_{\mathrm{T}} \approx \Psi_0$. In this case E_{L} is not constant, but has a weaker functional dependence on the coordinates compared to the

underlying potential.

To generate the working equations for this formulation of DMC,[25,113] we start with the imaginary-time time-dependent Schrödinger equation,

$$\frac{d\Psi}{d\tau} = \sum_{j=1}^{n_{\text{atoms}}} \left[\frac{\hbar^2}{2m_j} \nabla_j^2 - (V(\mathbf{x}) - E_{\text{ref}}) \right] \Psi(\mathbf{x}) \tag{2.10}$$

and substitute f/Φ_{T} for Ψ, where f represents the product of the trial wave function and the wave function for the state of interest. If we define a drift term,

$$\vec{D}_j(\mathbf{x}) = \frac{\hbar^2}{m_j} \frac{1}{\Phi_{\text{T}}(\mathbf{x})} \vec{\nabla}_j \Phi_{\text{T}}(\mathbf{x}) \tag{2.11}$$

upon substitution, Eq. 2.10 becomes

$$\begin{aligned} \frac{df}{d\tau} &= \sum_{j=1}^{n_{\text{atoms}}} \left[\frac{\hbar^2}{2m_j} \nabla_j^2 - (E_{\text{L}}(\mathbf{x}) - E_{\text{ref}}) \right] f(\mathbf{x}) \\ &\quad - \sum_{j=1}^{n_{\text{atom}}} \vec{\nabla}_j \cdot \left(\vec{D}_j(\mathbf{x}) f(\mathbf{x}) \right) \end{aligned} \tag{2.12}$$

The first line of Eq. 2.12 is treated in the same manner as standard DMC simulations, described above. Two additional steps need to be introduced to account for the drift term. Following the diffusion step, each walker is displaced by $\mathbf{D}(\mathbf{x})\Delta\tau$, and the new coordinates of the ith walker are denoted as $\mathbf{x}_i'$. This move is accepted or rejected based on the the ratio of the probability of displacing from $\mathbf{x}_i' \rightarrow \mathbf{x}_i$ in a time step of $\Delta\tau$ to the probability of displacing from $\mathbf{x}_i \rightarrow \mathbf{x}_i'$ in the same time.[25,42,114] If the ratio is larger than 1, then the move is accepted. If not, it is only accepted if the ratio is larger than a random number selected in the range from 0 to 1. If the move is rejected, the $\mathbf{x}_i(\tau + \Delta\tau) = \mathbf{x}_i(\tau)$.[25,114]

To illustrate how this works, in Fig. 2.1, we plot the quantities that comprise Eqs. 2.9 to 2.12 for a one-dimensional model system based on a Morse oscillator that models the hydrogen-bonded OH bond in water dimer, with a frequency of 3704.5 cm^{-1}, an anhar-

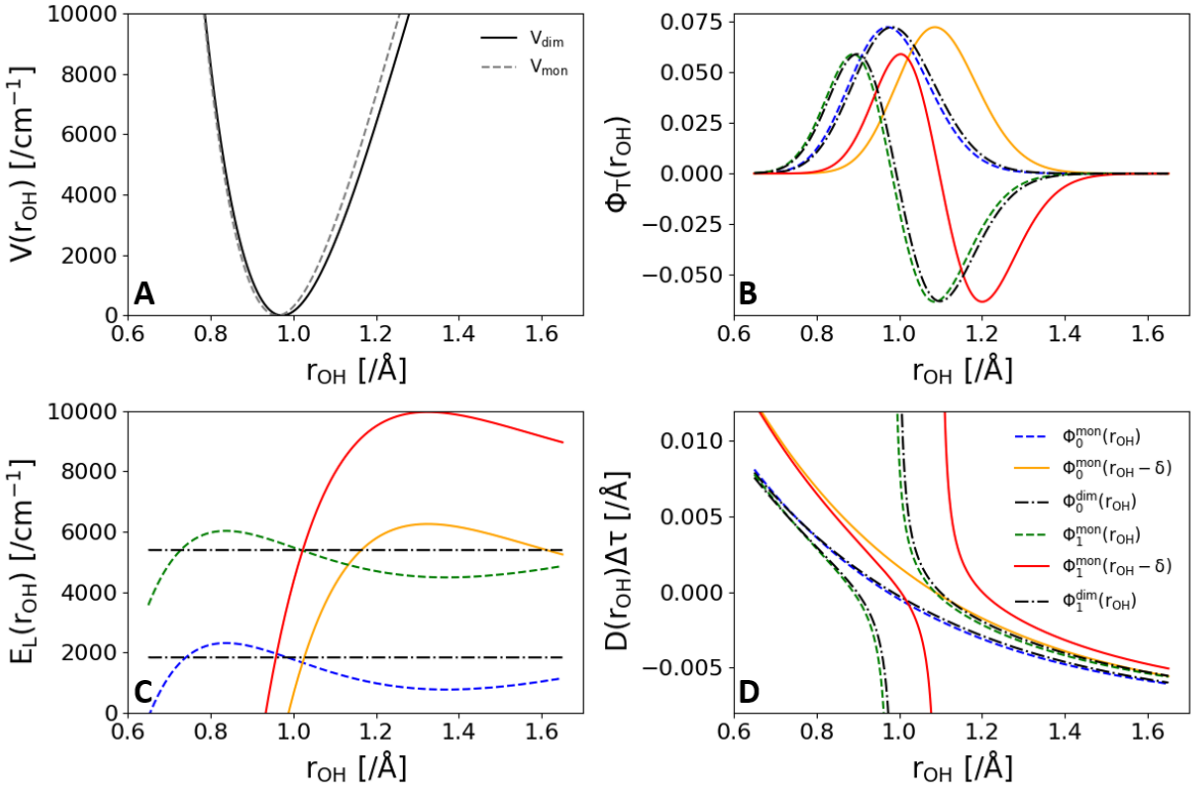


Figure 2.1: Illustration of the relationship between the quantities described in Eqs. 2.9-2.12. A. One-dimensional Morse Oscillator potentials model the hydrogen-bonded OH bond in water dimer, with a frequency of 3704.5 cm^{-1}, an anharmonicity of 75.3 cm^{-1}, and an equilibrium OH bond length of 0.973 Å ($V_{\rm dim}$ black solid line), and a Morse oscillator potential that describes an OH bond in an isolated water molecule, with a frequency of 3884.8 cm^{-1}, an anharmonicity of 86.9 cm^{-1}, and an equilibrium OH bond length of 0.961 Å ($V_{\rm mon}$, grey dashed line),[4] plotted as a function of the OH bond length, $r_{\rm OH}$. B. Ground and first excited state wave functions for the Morse oscillators plotted in A. C. Local energy evaluated using Eq. 2.9 using $V_{\rm dim}$, and wave functions plotted in panel B. D. The value of the product of the drift term and $\Delta\tau$ with $\Delta\tau = 1$, and the drift term evaluated using Eq. 2.11 with the wave functions plotted in panel B. In panels B-D the blue and green dashed curves provide results when we use $\Phi_n^{\rm mon}(r_{\rm OH})$, the gold and red solid lines provide the wave functions we use $\Phi_n^{\rm mon}(r_{\rm OH}-\delta)$, where $\delta = 0.115$ Å, which is 10 times the difference between $r_{{\rm OH},e}$ for the monomer and the dimer, and the black dotted lines provide the results when we use $\Phi_n^{\rm dim}$.

monicity of 75.3 cm^{-1}, and an equilibrium OH bond length of 0.973 Å. The potential, $V_{\rm dim}$ is plotted with a solid black line in Fig. 2.1A. For $\Phi_{\rm T}$, we use the ground state solution to a Morse oscillator that describes an OH bond in an isolated water molecule, with a frequency of 3884.8 cm^{-1}, an anharmonicity of 86.9 cm^{-1}, and an equilibrium OH bond length of 0.961 Å.[4] This potential is plotted with a grey dashed line in Fig. 2.1A. The blue and green dashed curves in Fig. 2.1B show the $n = 0$ and 1 wave functions obtained using the Morse oscillator potential for water monomer, denoted $\Phi_n^{\rm mon}$. The gold and red solid lines show $\Phi_n^{\rm mon}$ when the $\Delta r_{\rm OH}$ axis is shifted by 0.115 Å, which is 10 times the difference between $r_{{\rm OH},e}$ for the dimer and the monomer. This shift was chosen to explore what happens when a poor choice for $\Phi_{\rm T}$ is made. Finally, the black dotted lines provide the results when we use $\Phi_n^{\rm dim}$.

Before discussing the results in panels C and D it is useful to recall that compared to the wave functions that describe the OH bond in water dimer, shown with dotted lines in Fig. 2.1B, the wave functions for the monomer have more amplitude at smaller values of $r_{\rm OH}$ and less amplitude at larger values of $r_{\rm OH}$. Near $r_{\rm OH} = r_{{\rm OH},e}^{\rm dim}$, the local energy for the ground and excited state are roughly constant when the unshifted wave functions are used. Likewise, in this region, the drift term for the ground state vanishes, while it diverges for the $n = 1$ state. When the monomer wave function is shifted, the local energy becomes constant at larger values of $r_{\rm OH}$. The shift in the wave function simply shifts in the drift term, as expected from Eq. 2.11.

If we only considered the local energy, we find that it decreases as $r_{\rm OH}$ is decreased, but remains roughly constant out to 1.6 Å. Based on this, we expect the ensemble of walkers to shift to smaller distances the longer the propagation is run. This will be more dramatic for the shifted wave functions, plotted in gold and red. This is exactly the behavior that is needed to obtain Ψ_n from the shifted version of $\Phi_n^{\rm mon}$. Considering only $E_{\rm L}$ one might expect the build up of amplitude at small values of $r_{\rm OH}$ to be problematic. The drift term, plotted in Fig. 2.1D, prevents this from happening by shifting the location of walkers away from regions where the amplitude of the trial wave function is small to regions where it is larger. This can be seen by the divergence of the drift term where there is a node in the

Table 2.1: Energies (in cm^{-1}) obtained from DMC calculations using $V_{\rm dim}$, and the wave functions shown in Fig. 2.1.

$\Phi_{\rm T}$	ΔE_0^a	$\Delta E_1^{{\rm left},b}$	$\Delta E_1^{{\rm right},b}$
$\Phi_n^{\rm dim}(r_{\rm OH})$	< 0.001	< 0.001	< 0.001
$\Phi_n^{\rm mon}(r_{\rm OH})$	$-0.43(1.39)$	$533.20(0.08)$	$-362.57(0.39)$
$\Phi_n^{\rm mon}(r_{\rm OH} - 0.011$ Å$)$	$0.10(0.12)$	$-25.20(0.38)$	$18.34(0.07)$
$\Phi_n^{\rm mon}(r_{\rm OH} - 0.115$ Å$)$	$-44.62(20.42)$	$-3246.76(34.19)$	$4092.42(0.39)$

[a] $\Delta E_0 = E_{\rm DMC} - E_0^{\rm dim}$, where $E_0^{\rm dim} = 1833.42$ cm^{-1}.

[b] $\Delta E_1^{\rm left/right} = E_{\rm DMC} - E_1^{\rm dim}$, where $E_1^{\rm dim} = 5387.37$ cm^{-1}. These values were obtained from DMC calculations considering only one side of the node.

wave function. As noted above, care needs to be taken in selecting $\Phi_{\rm T}$ as a poor choice can lead to poor sampling. Characteristics of a poor choice for $\Phi_{\rm T}$ include trial wave functions that do not have amplitude in regions of configuration space where the wave function that is being modeled has amplitude, or if the node is not correctly placed for an excited state.

To illustrate the sensitivity of the quality of the results to the choice of $\Phi_{\rm T}$, we ran DMC simulations using the potential and wave functions shown in Fig. 2.1 as well as $\Phi_n^{\rm mon}(r_{\rm OH} - 0.011$ Å$)$. The resulting energies are reported in Tab. 2.1. When we use $\Phi_n^{\rm dim}$, the resulting energies are in exact agreement with the expected energies. Ground state calculations using $\Phi_0^{\rm mon}$ produce an accurate ground state energy, with an error of 0.4 cm^{-1}, and an uncertainty that is more than three times this error. Shifting $\Phi_0^{\rm mon}$ by 0.011 Å, which is the difference between $r_{{\rm OH},e}$ for the dimer and monomer, improves the agreement, and the error is reduced to less than 0.1 cm^{-1}. When we shift $\Phi_0^{\rm mon}$ by a much larger amount, the accuracy of the results deteriorates. In the case of excited states, the presence of the node introduces some complications, and separate calculations are preformed for the regions where $\Phi_{\rm T}$ has positive and negative amplitude, denoted as left and right, respectively, in Tab. 2.1. Overall, none of the results are as accurate as the ground state calculations, but the same general trends hold. The most accurate results are obtained when the trial wave function is shifted by 0.011 Å.

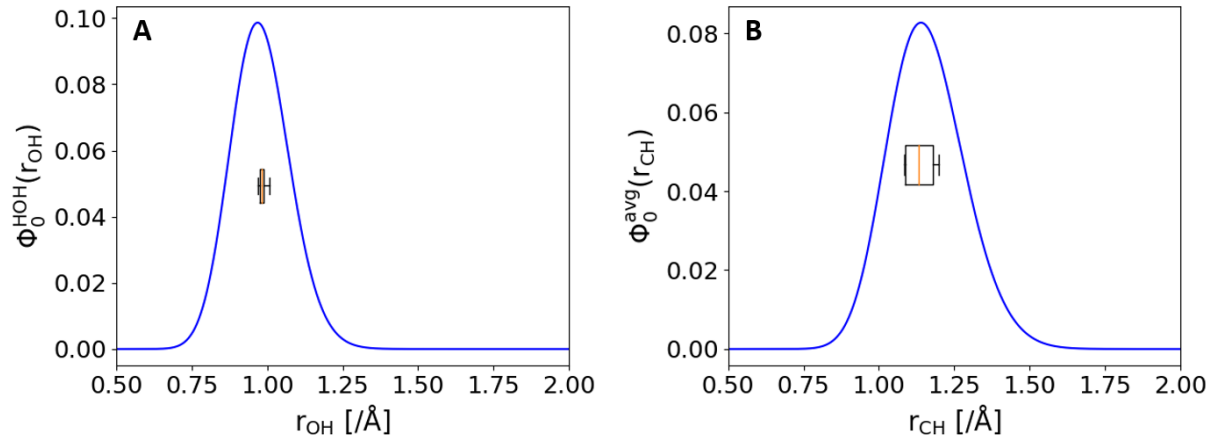


Figure 2.2: Comparison of the range of bond lengths in stationary point structures (box and whiskers plots) to the trial wave functions (plotted with blue solid lines) used in the development of Φ_T for DMC simulations. In panel A we focus on hydrogen-bonded OH bonds in water clusters with six or fewer water molecules[4] and in panel B we explore the stationary point structures of CH_5^+.[5]

As we move from one-dimensional model problems to real molecules and molecular clusters, we need to be careful in our choice of guiding functions. In DMC the high frequency vibrations often cause the greater challenges than the low-frequency ones. They are also easier ones to develop approximate wave functions for. Specifically, our approach to guided DMC is based on using trial wave functions that are products of one-dimensional functions of the high frequency XH stretching, and HXH bending vibrations.[27,35,97] For example, for water clusters, we use a guiding function that is a product of one-dimensional wave functions that are functions of OH bond lengths. These functions are obtained by solving for the ground state solution using a one-dimensional cut through the potential for an isolated water molecule[115] along the OH stretch coordinate. For the bend, we use a one-dimensional harmonic oscillator.[27] A similar approach was previously suggested by Suhm and Watts,[24] but it does not appear to have been applied to many studies. This is likely due to the lack of availability of potentials that accurately describe intra- and intermolecular motions in molecular clusters.

This approach is expected to work well for water clusters, as OH and HOH wave functions show only small changes when the isolated water molecule is placed in the context of a cluster. For example, in Fig. 2.2A we compare the range of OH bond lengths in hydrogen-bonded OH bonds in water clusters with six or fewer water molecules to the ground state OH stretch wave function used in the development of Φ_T. As can be seen, the OH bond lengths generally deviate by less than 0.036 Å,[4] which is comparable to the shift in the wave functions shown in Fig. 2.1. This contrasts the situation for the CH bonds in CH_5^+, shown in Fig. 2.2B, which shows a much larger spread of CH bond lengths. In this case, if we used a single CH trial wave function, which is the average of the ground state wave functions for the five CH oscillators in the equilibrium geometry of the ion, to describe the five CH wave functions in this ion in our guided calculation, we obtain a zero-point energy of 10 926(4) cm^{-1},[97] which is roughly 10 cm^{-1} larger than the zero-point energy of this ion, 10917.[94] A more sophisticated approach, which adjusts the form for the trial wave function depending on the local environment of the CH bond will be needed to obtain an accurate zero-point energy for CH_5^+. The details of such an approach is discussed below.[97]

2.4 Evaluation of Excited States and Molecular Properties

The ability to evaluate ground state wave functions and energies allow us to obtain insights into systems that undergo large amplitude vibrational motions. For example, they allow us to explore the relative stability of various isomers of a molecular species of interest, and how this is affected by deuteration.[27,34,35,99] On the other hand, much more can be learned from the ability to evaluate properties of the wave function and to calculate the energy and wave functions for excited states. In the discussion that follows, we will focus on DMC with guiding functions, in which the ensemble of walkers samples $f = \Psi \times \Phi_T$. One can consider the unguided situation as the special case when Φ_T is constant.

2.4.1 Obtaining excited state wave functions

Obtaining excited state wave functions with guiding functions requires little adjustment to the algorithm described above beyond using a trial wave function that describes the excited state of interest. If after the displacement of the walker, the phase of Φ_{T} has changed, the displacement is rejected, and the coordinates of the walkers are restored to their value prior to the displacement step. While this change appears straightforward, the sensitivity of the sampling to deficiencies in the trial wave function are more significant for excited states than for ground states. The source of the problem is illustrated by the divergence of the drift term at the position of the node, shown in 2.1D and Tab. 2.1. An improperly placed node can lead to poor sampling and errors in the energy obtained using DMC. This is why $\Delta E_1^{\mathrm{left}} \neq \Delta E_1^{\mathrm{right}}$ in Tab. 2.1.

To address this problem, prior to performing the DMC simulations, we identify the optimal position of the node or nodal surface. To accomplish this, we draw from approaches we developed for identifying nodes in unguided DMC calculations.[46,47,116] Specifically, we obtain an excited state trial wave function by replacing the one-dimensional contribution to the trial wave function in the coordinate that is to be excited by the appropriate excited state based on the potential used to evaluate the ground state wave function. This procedure is likely to produce a trial wave function where the node is slightly shifted from the optimal position. To adjust the position of the node we take advantage of the fact that displacements of the walkers that change the phase of Φ_{T} are rejected. This allows us to run independent simulations for regions of configuration of space where Φ_{T} is positive and negative. In one-dimension, if the node is correctly placed, the two simulations must provide identical values for the energy. In general, this becomes more complicated in higher dimensions. Since we are assuming a product form for Φ_{T} it is straightforward to obtain the nodal surface under this approximation of separability by requiring the energies obtained for the regions of configuration space where Φ_{T} is positive and negative be equal. Using this constraint, we run two DMC simulations in which at each time step the coordinate on which

the wave function depends is shifted by a small amount. By making the displacement of the wave function a linear function of propagation time the approximate ground state energy will also change with τ. Plots of E_{ref} as a function of this displacement are made for both simulations, as illustrated in Fig. 2.3. For this calculation, we are using Φ_1^{mon} in Fig. 2.1A for the trial wave function to describe Ψ_1^{dim}. The point at which these two curves cross (0.993 Å and 5384 cm^{-1}) provides the optimal shift of the trial wave function, and the corresponding energy is the energy of this excited state. To help illustrate the accuracy of the approach, the expected position of the node and energy are shown with a black circle, which is at 0.993 Å and 5387.37 cm^{-1}.

2.4.2 Obtaining molecular properties

Guided diffusion Monte Carlo provides a Monte Carlo sampling of f. As such, an integral over this distribution can be replaced by a sum over the walkers as

$$\int A(\mathbf{x}) f(\mathbf{x}) d\mathbf{x} = \sum_{i=1}^{N_W} w_i A(\mathbf{x_i}) \tag{2.13}$$

In this way, the expectation value of a multiplicative operator $\hat{O} = O(\mathbf{x})$ can be evaluated using

$$\langle O \rangle = \int \Psi^2(\mathbf{x}) O(\mathbf{x}) d\mathbf{x} = \sum_{i=1}^{N_W} w_i O(\mathbf{x}_i) \left(\frac{\Psi(\mathbf{x}_i)}{\Phi_{\text{T}}(\mathbf{x}_i)} \right) \tag{2.14}$$

while the projection of the probability amplitude onto an arbitrary coordinate, r is evaluated using

$$P(r') = \int \Psi^2(\mathbf{x}) \delta(r(\mathbf{x}) - r') d\mathbf{x} = \sum_{i=1}^{N_W} w_i \delta(r(\mathbf{x}_i) - r') \left(\frac{\Psi(\mathbf{x}_i)}{\Phi_{\text{T}}(\mathbf{x}_i)} \right) \tag{2.15}$$

Note operationally, $\delta(r(\mathbf{x}) - r')$ is replaced by a bin of width Δr.

Evaluation of the expressions in Eqs. 2.14 and 2.15 requires the ability to calculate Ψ/Φ_{T} at the coordinates of the walkers. Barnett et al. and Suhm and Watts[24,114] have shown that $(\Psi(\mathbf{x}_i)/\Phi_{\text{T}}(\mathbf{x}_i))$ is proportional to the ratio of the weight of the ith walker at a propagation

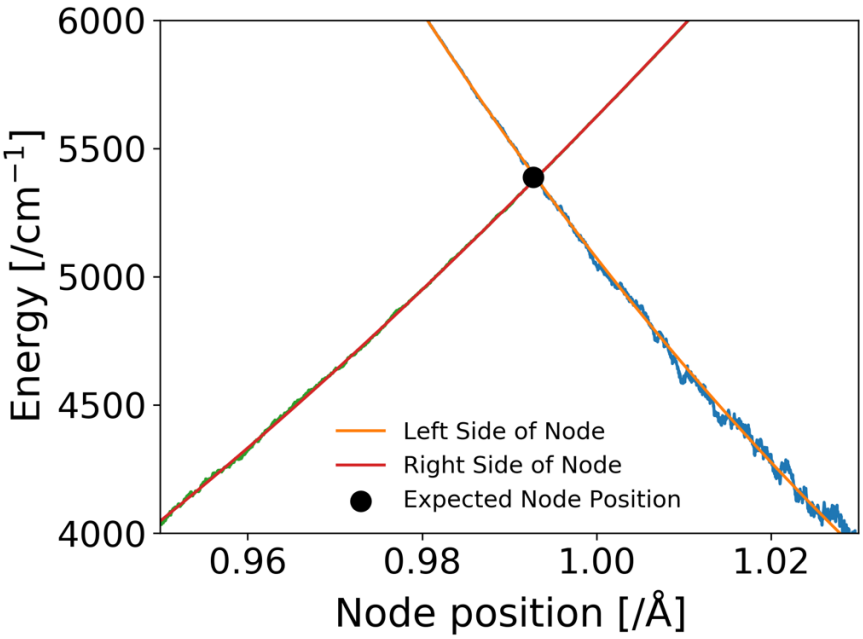


Figure 2.3: Two example simulations where Φ_1^{mon} is used as the guiding function to describe the first excited state of the Morse oscillator that was constructed to approximate the hydrogen-bonded OH bond in water dimer, see Fig. 2.1. In one simulation (blue) the walkers are placed on the left side of the node and Φ_1^{mon} is shifted to the right throughout the simulation. In the other simulation (green) the walkers are placed on the right side of the node and Φ_1^{mon} is shifted to the left throughout the simulation. The E_{ref} is plotted with respect to node position and fit to a cubic polynomial (red and gold lines) to find the correct node position for this system. Finally, the black circle shows the expected location of the node and the corresponding energy.

time of $\tau' = \tau + \delta\tau$ to the weight at propagation time τ.

$$\frac{\Psi(\mathbf{x}_i(\tau))}{\Phi_{\mathrm{T}}(\mathbf{x}_i(\tau))} \propto \frac{w_i(\tau')}{w_i(\tau)} \tag{2.16}$$

and, for example

$$\langle O \rangle = \sum_{i=1}^{N_W} w_i(\tau) O(\mathbf{x}_i(\tau)) \left(\frac{\Psi(\mathbf{x}_i(\tau))}{\Phi_{\mathrm{T}}(\mathbf{x}_i(\tau))} \right) = \sum_{i=1}^{N_W} w_i(\tau') O(\mathbf{x}_i(\tau)) \tag{2.17}$$

Typically, $\delta\tau$ corresponds to roughly 200 - 300 a.u. of propagation time, and this approach is referred to as descendant weighting. For a discrete weighting simulation, this ratio is replaced by the number of walkers at time τ' that can be traced to the ith walker in the ensemble at time τ. An advantage of this approach is that it is straightforward to incorporate into the DMC simulation. The disadvantage is that the quality of the results can be sensitive to the value of $\delta\tau$, and the optimal value can depend on the system that is being studied. In addition, while the approach is effective for evaluating expectation values, its extension to the evaluation of matrix elements that involves the overlap of two different wave functions, while feasible, is challenging due to the fact that the ensembles of walkers used to sample the two states are are not the same.[42]

An alternative approach takes advantage of the fact that Φ_{T} provides a good approximation to Ψ, at least for the coordinates that are included in the development of Φ_{T}. As such, $\Psi - \Phi_{\mathrm{T}} = \delta$ can be assumed to be small. Following Singer and co-workers,[93] based on this approximation for Ψ,

$$\Psi^2 = 2f - \Phi_{\mathrm{T}}^2 + \delta^2 \approx 2f - \Phi_{\mathrm{T}}^2 \tag{2.18}$$

With this relationship,

$$P(r') = 2 \sum_{i=1}^{N_W} w_i \delta(r(\mathbf{x})_i - r') - \phi_{\mathrm{T}}^2(r') \tag{2.19}$$

where $\phi_{\mathrm{T}}(r')$ is the one-dimensional wave function that contributes to Φ_{T}, which is a function

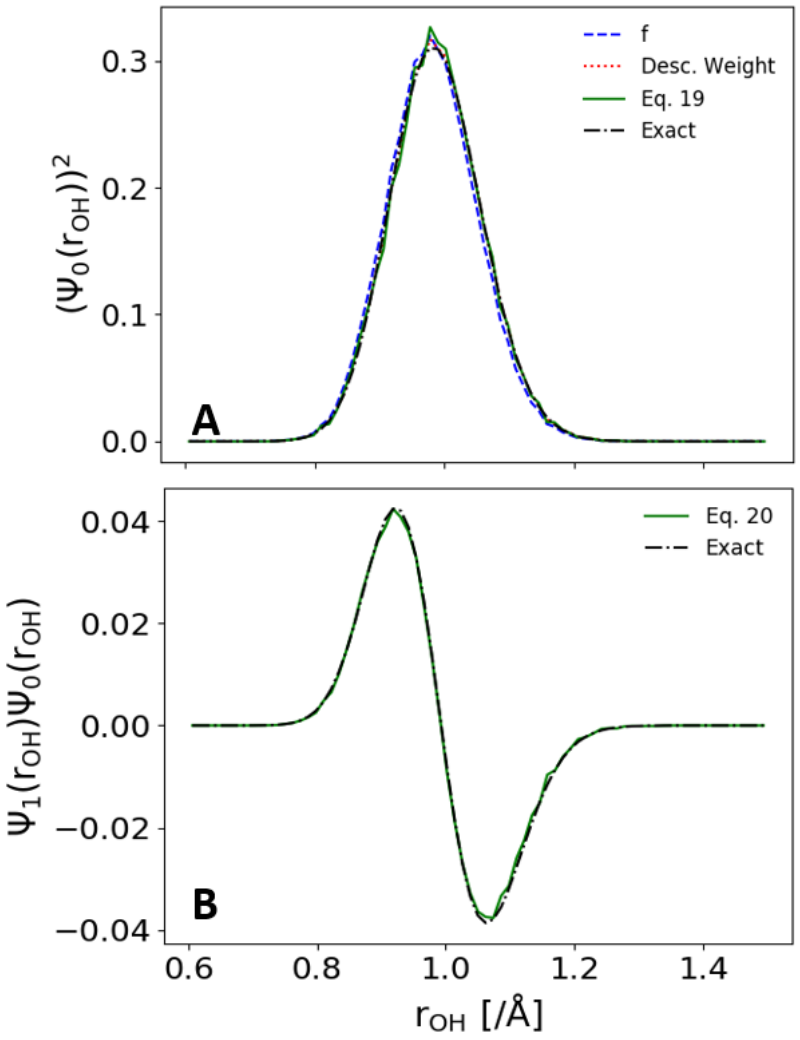


Figure 2.4: Exploration of several approximations for evaluating $(\Psi_0^{\text{dim}})^2$ (panel A), and $\Psi_0^{\text{dim}}\Psi_1^{\text{dim}}$ (panel B) shown in black dotted curves in Fig. 2.1B.

of r. Expectation values of operators can be obtained by a similar substitution into Eq. 2.13. By extension, if we are interested in evaluating $\langle\Psi'|O|\Psi\rangle$, we can also assume that $\Psi' = \Phi'_{\mathrm{T}} + \delta'$, where δ' is also small. This leads to the approximation

$$\begin{aligned}\langle\Psi'|O|\Psi\rangle &= \sum_{i=1}^{N_W} w_i O(\mathbf{x}_i)\left(\frac{\Phi'_{\mathrm{T}}(\mathbf{x}_i)}{\Phi_{\mathrm{T}}(\mathbf{x}_i)}\right) + \sum_{i'=1}^{N'_W} w'_{i'} O(\mathbf{x}_{i'})\left(\frac{\Phi_{\mathrm{T}}(\mathbf{x}_{i'})}{\Phi'_{\mathrm{T}}(\mathbf{x}_{i'})}\right) \\ &\quad - \int O(\mathbf{x})\Phi_{\mathrm{T}}(\mathbf{x})\Phi'_{\mathrm{T}}(\mathbf{x})d\mathbf{x} \end{aligned} \tag{2.20}$$

where the summations are over the distribution of walkers that represent f and f'.

To illustrate the relative accuracy of these treatments, we apply them to the Morse oscillators shown in Fig. 2.1. In Fig. 2.4A, we plot the ground state probability amplitude. The results based on Φ_0^{dim} are shown in red, and are compared to the value of f and Ψ^2 obtained using descendant weighting and Eq. 2.19. As can be seen all three approaches yield accurate results. Likewise, the approximation to the overlap of Ψ_0^{dim} and Ψ_1^{dim} provided by Eq. 2.20 shows good agreement with the expected results. The distributions plotted in Fig. 2.4 can also be used to evaluate expectation values or matrix elements of operators. For this discussion, we will focus on $x = r_{\mathrm{OH}} - r_{\mathrm{OH},e}^{\mathrm{dim}}$. The results of the various approximations are provided in Tab. 2.2. As can be seen, all of the approaches yield distributions that closely resemble the expected distribution (plotted with a black dash/dot line). The agreement is poorest when we approximate Ψ^2 by f. This is more easily seen in the comparison of the values of $\langle x\rangle$, reported in Tab. 2.2, which have been evaluated either using descendant weights or the expressions provided in Eqs. 2.19 and 2.20. In the case of the overlaps, using Eq. 2.20 to obtain these quantities is much less computationally expensive than using descendant weighting.[42] These results are provided for one-dimensional systems. When moving to larger systems the approximations provided in Eqs. 2.19 and 2.20 are only expected to provide accurate results when the guiding function is a function of the coordinate of interest.

Table 2.2: Accuracy of approximate approaches for evaluating Ψ^2 and $\Psi_n\Psi_{n'}$ based on the Morse oscillator that describes the hydrogen-bonded OH oscillator in water dimer, see Fig. 2.1

Approach	$\langle\Psi_0\|x\|\Psi_0\rangle^a$	$\langle\Psi_1\|x\|\Psi_0\rangle^a$
Exact	0.0151	0.0700
Descendant Weighting	0.0150 (0.0007)	-
f	0.0095 (0.0006)	-
Eqs. 2.19 or 2.20	0.0151 (0.0012)	0.0706 (0.0015)

[a] $x = r_{\mathrm{OH}} - r_{\mathrm{OH},e}^{\mathrm{dim}}$

2.5 Conclusions

In this chapter, we have described approaches for studying molecules that undergo large-amplitude vibrational motions, which are based on Diffusion Monte Carlo, through the introduction of carefully chosen trial wave functions. The introduction of guiding functions allow us to extend the size and types of molecular systems that can be studied using DMC. We have described some of the challenges introduced by using guiding functions as well as the opportunities they provide. Most importantly, these functions need to be selected with care as they can bias the results of DMC simulations. This is why our use of guiding functions has focused on intramolecular vibrations. Since the equilibrium bond lengths and vibrational frequencies of the molecular units that make up molecular clusters show relatively small variability with cluster size or the environment that the oscillator is in, developing trial wave functions from products of one-dimensional functions of the high frequency intramolecular motions allows us to realize substantial savings in the size of the DMC calculations needed to achieve converged results without sacrificing their accuracy.[27,35] In some cases, like CH_5^+ or protonated water clusters, this approach requires minor modifications to account for changes in the equilibrium bond length or vibrational frequency of certain oscillators with low-frequency, large amplitude motions.[97]

We have also illustrated extensions to these ground state methods, by illustrating for one-dimensional systems, how guided DMC can be used to obtain excited state wave functions

and energies. Work is currently underway to apply this approach to larger systems. With the ability to obtain ground and excited states, the next step is to obtain properties, like transition moments (e.g. matrix elements of the dipole moment between two of these wave functions). This is complicated by the fact that in a typical implementation, each vibrational state is evaluated independently. As a result, these states are sampled at distinct sets of geometries, making the evaluation of matrix elements computationally demanding. We have extended the approach for obtaining probability amplitudes from $f = \Psi_0 \Phi_T$, described by Cho and Singer,[93] to allow us to obtain matrix elements using Eq. 2.20.

Finally, arguably the most time consuming part of the DMC simulation comes in the evaluation of the potential function. We have recently developed an extension of DMC to employ deep neural networks to fit the potential in the region that is relevant for a ground state DMC simulation.[117,118] Along with those described in this Chapter, this advance is expected to greatly extend our capabilities for studying molecular systems that undergo large amplitude vibrational motions using DMC.

Chapter 3

GUIDED DIFFUSION MONTE CARLO: A METHOD FOR STUDYING MOLECULES AND IONS THAT DISPLAY LARGE AMPLITUDE VIBRATIONAL MOTIONS

Reproduced in part with permission from [Jacob M. Finney, Ryan J. DiRisio, and Anne B. McCoy. Guided Diffusion Monte Carlo: A Method for Studying Molecules and Ions That Display Large Amplitude Vibrational Motions. *J. Phys. Chem. A* **2020**, 124 (46), 9567–9577]. Copyright [2020] American Chemical Society.

3.1 Introduction

In this chapter, we explore whether previous strategies of guided diffusion Monte Carlo,[27,35] could be used to study protonated water clusters and CH_5^+ for which the equilibrium OH or CH bond lengths depend strongly on the values of the coordinates associated with the low-frequency vibrations. For example, although the five CH bonds in CH_5^+ are equivalent once zero-point energy is considered, the equilibrium CH bond lengths in the coordinates of the three lowest energy stationary points on the potential energy surface range from 1.09 to 1.2 Å. This range is comparable to the width associated with the ground state probability amplitude for a CH oscillator with a frequency of 3000 cm^{-1}, which is 0.17 Å. Likewise the frequency of the shared proton OH stretch in protonated water clusters ($H^+(H_2O)_n$) can vary between 1000 and 2700 cm^{-1} when $n = 2$ and 4, respectively. As such, the approach for describing Φ_T that proved to be effective for the neutral water clusters will not be as effective for these motions. Instead, we propose an approach that allows the maximum in the wave function that describes the CH or OH vibrations to depend on the environment of the CH or OH oscillator. In the case of the OH stretches in the protonated water systems,

the width of the trial wave function is also allowed to adjust based on the environment. This approach will be applied to CH_5^+ and protonated water clusters to explore the savings that can be achieved as well as insights into couplings of the intermolecular and intramolecular vibrations in these ions.

3.2 Theory

Diffusion Monte Carlo and our implementation have been described in the previous chapter. As mentioned above, Φ_T is the product of one-dimensional wave functions that are each functions of one of the high frequency intramolecular coordinates, e.g. an XH bond length or an HOH angle. These guiding functions will be described in the following section.

3.2.1 DVR Calculation of the Wave Functions for the CH and OH Stretches

To obtain the trial wave function for the CH stretch used in this work, we performed a one-dimensional scan of the potential energy along this coordinate, constraining all of the other bond lengths and HCH angles to their values in the minimum energy structure. For these calculations, we used the Potential of Jin *et al.*[5] For this scan, we evaluated the potential at 5000 CH bond lengths ranging from 0.21 to 3.18 Å. This provided the potential function for a one-dimensional discrete variable representation (DVR) calculation of the ground state wave function.[119] Similar calculations were performed to obtain the other CH wave functions and vibrational frequencies reported in Figures 3.6 and 3.1 and Table 3.1.

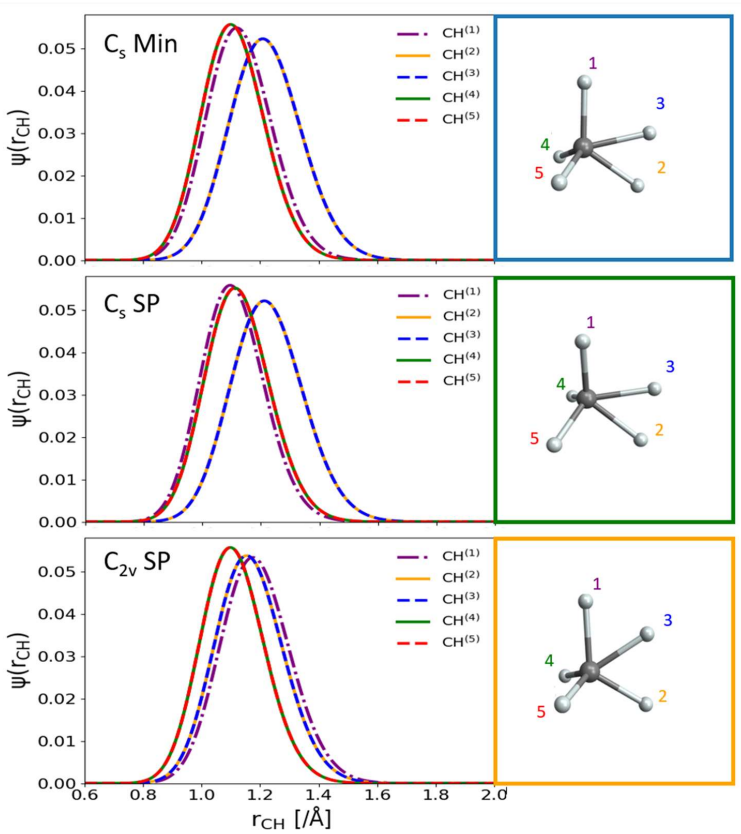


Figure 3.1: Ground state wave functions for CH stretches evaluated at the C_s minimum energy geometry, the C_s saddle point geometry, and the C_{2v} saddle point geometry of CH_5^+. The associated frequencies of these stretches are provided in Table 3.1.

Table 3.1: Frequencies and Equilibrium Bond Lengths in CH_5^+

C_s Min		
CH	$\nu_{CH}^{(i)}$ $(cm^{-1})^a$	$r_{CH,e}^{(i)}$ $(Å)^b$
1	2896.28	1.1076
2	2383.95	1.1966
3	2393.78	1.1971
4	3069.90	1.0881
5	3069.90	1.0881
C_s SP		
CH	$\nu_{CH}^{(i)}$ $(cm^{-1})^a$	$r_{CH,e}^{(i)}$ $(Å)^b$
1	3110.05	1.0842
2	2373.51	1.1996
3	2373.51	1.1996
4	2968.06	1.0991
5	2968.06	1.0991
C_{2v} SP		
CH	$\nu_{CH}^{(i)}$ $(cm^{-1})^a$	$r_{CH,e}^{(i)}$ $(Å)^b$
1	2587.22	1.1626
2	2671.41	1.1414
3	2671.41	1.1414
4	3073.04	1.0870
5	3073.04	1.0870

[a] Frequency of the CH stretch based on a DVR calculation.[119]

[b] Equilibrium bond length.

A similar approach was taken to obtain the OH stretch wave function for the OH stretch in water. For these calculations, we use the water potential developed by Partridge and Schwenke.[115] The scan included 500 OH bond lengths ranging from 0.26 to 2.11 Å. These potential evaluations were then used in a one-dimensional DVR calculation to generate the ground state wave function for the OH stretch used in this study.[119]

3.2.2 Calculation of $r_{\mathrm{OH}}^{\mathrm{max}}$ and σ_{OH} for $H^+(H_2O)_3$ and $H^+(H_2O)_4$

Starting from the equilibrium structure, we performed two-dimensional potential energy scans in the hydrogen-bonded OH stretch coordinate (r_{OH}) and the corresponding OO stretch (R_{OO}) using the potential energy surface that was used in this study. The hydrogen atom was constrained to lie on the OO axis.

The R_{OO} scans ranged from 2.17 to 3.55 Å and 2.24 to 3.62 Å for $H^+(H_2O)_3$ and $H^+(H_2O)_4$ respectively. At each of the 50 R_{OO} values, we evaluated the potential at 2000 values of r_{OH}, ranging from 0.56 to 1.67 Å and 0.53 to 1.64 Å for $H^+(H_2O)_3$ and $H^+(H_2O)_4$ respectively. The potential evaluations at the r_{OH} scan points were then used in a one-dimensional DVR calculation to generate the ground state wave function as a function of the OH bond length for each value of R_{OO}.[119] Using these ground state wave functions, we then calculated the standard deviation σ_{OH} and the value of r_{OH} at which the maximum amplitude of the wave function occurred, which is referred to as $r_{\mathrm{OH}}^{\mathrm{max}}$ throughout the discussion.

3.2.3 Numerical Details

For the simulations of the protonated water clusters, $H^+(H_2O)_{n=2,3,4}$ the potential used to propagate the walkers was developed by Yu and Bowman.[8] The version of this potential used in this study incorporates an updated three-body term involving the hydronium core and two water molecules. We needed to use this modified surface because preliminary studies using the previously reported surface has a low-energy region that corresponds to dissociation of H^+ from the water cluster, which was sampled by our continuous weight simulations. The potential energy surface for H_3O^+ is the one developed by Huang, Carter, and Bowman.[120] The simulations for the protonated water clusters were first equilibrated by running for 5000 time steps with a $\Delta\tau$ of 10 a.u. without the use of a guiding function, then the simulations were run for an additional 20 000 time steps with a $\Delta\tau$ of 1 a.u. either with or without a guiding function. This added equilibration step was introduced to ensure that the potential energy surface was fully sampled at the beginning of the simulation. Simulations with a $\Delta\tau$

of 10 a.u. were run for 20 000 time steps without this added equilibration step. The zero-point energy was then obtained by time averaging E_{ref} over the final 15 000 time steps of the simulation. Although the simulations that are run using the 10 a.u. time steps are allowed to propagate for roughly three times as long as the simulations that use the 1 a.u. time step, we find that the average zero-point energies obtained when we perform the averages over the same propagation times as are used to analyze the 1 a.u. results agree with those obtained using longer propagation times, albeit with larger statistical uncertainties.

For the calculations involving CH_5^+ and its deuterated analogues, the initial ensemble of walkers is distributed randomly among the 120 equivalent minima on the potential energy surface. For these calculations we use the potential of Jin, Braams, and Bowman, which was fit to electronic energies of CH_5^+ calculated at the CCSD(T)/aug-cc-pVTZ level of theory/basis.[5] Each simulation was run for 20 000 time steps with a $\Delta\tau$ of 1 a.u. The zero-point energy was obtained by time averaging E_{ref} from 5000 time steps to the end of the simulation. This allows us to ensure that the ensemble is equilibrated before the energies are collected.

All of the reported energies are based on five independent simulations, and the uncertainties represent one standard deviation based on these values, while descendant weights were collected 20 times for 250 time steps over the last 15 000 time steps in order to collect wave functions for each simulation.

For the guided simulations, the HOH bends were described by harmonic oscillators based on a harmonic frequency of 1668 cm^{-1} and a $G-$matrix element[121] of 2.338 amu^{-1} $\mathrm{Å}^{-2}$.[35] The wave functions that describe the CH bonds in CH_5^+ and the OH bonds in the protonated water clusters were obtained using a discrete variable representation (DVR)[119] based on one-dimensional slices through the potentials of interest. The resulting wave functions were interpolated using a cubic spline. Calculations of the drift term and local energy require the first and second derivatives of Φ_T with respect to the Cartesian coordinates. These derivatives were evaluated numerically using a three-point finite difference.

3.3 Results

3.3.1 $H^+(H_2O)_n$

Protonated water clusters, like neutral water clusters, present challenges for DMC approaches. The smallest of these ions, H_3O^+, is covalently bound, and the OH stretch frequencies are 3445 and 3536 cm^{-1},[120] which is about 200 cm^{-1} lower than the OH stretch frequencies in an isolated water molecule. The next smallest system, $H^+(H_2O)_2$, shows large amplitude vibrations and large couplings between high and low frequency vibrations. These stronger couplings result from the very low frequency of the shared proton stretch, $\approx$ 1000 cm^{-1}, which is strongly coupled to the stretching and bending vibrations of the terminal water molecules.[50,122] As we add more water molecules, the amplitude of the water-bound OH stretching vibrations in the hydronium core are intermediate between H_3O^+ and $H^+(H_2O)_2$, and the frequencies of these vibrations in $H^+(H_2O)_3$ and $H^+(H_2O)_4$ are approximately 2100 and 2653 cm^{-1}, respectively.[9,10]

As we consider which vibrational degrees of freedom to include in the guiding functions, several options emerge. Given recent successes in applying a similar approach to studies of neutral water clusters,[27,35] one option is to use a direct product of wave functions that describe the OH stretch and HOH bends in the terminal water molecules as well as the vibrations of the OH stretches in the hydronium core that are not bound to water molecules. For these vibrations, we use the ground state wave functions obtained for an isolated water molecule.[115] Since the water bound OH stretches in the hydronium core of $H^+(H_2O)_n$ display a broad range of frequencies, it is not likely that the same distribution can be used to describe the ground state wave function of this oscillator in all of the possible structures of these systems. As we consider the development of the guiding functions to use for these systems, we focus on functional forms that are transferable to larger protonated water clusters, and which do not presume a specific bonding configuration.

We start by considering the convergence behavior of guided and unguided DMC calculations for H_3O^+. The results of this analysis are shown in panel A of Figure 3.2. In all of

the panels in this figure, the red symbols and lines provide the results of unguided DMC calculations with a time step of 1 a.u., while the gold lines and symbols provide the results when the guiding functions based on the OH stretches in an isolated water molecule are used to describe all of the unbound OH stretches in the hydronium core, and the OH stretches in the solvating water molecules. In addition, the guiding functions include a harmonic description of the HOH bend in the solvating water molecules. We do not try to describe the HOH bends in the hydronium core with the guiding functions due to redundancies in these coordinates in planar geometries. To aid in comparisons, the results based on the guided DMC calculation using the largest ensemble are extended using a dotted line, and the uncertainties are shown with shading of the same color. For hydronium, both the guided and unguided approaches yield results that are in good agreement with the previously reported DMC ground state energy of 7453 cm^{-1}, which was obtained using an unguided simulation with 20 000 walkers and a 1 a.u. time step.[98] For the unguided calculations, ensembles larger than 10 000 walkers yield energies that are well-converged, while for the guided calculations, 1000 walkers are needed to achieve similar accuracy. These results are consistent with previous studies of water clusters in our group. Unlike the study of water clusters, in which the trial wave function was chosen to provide a good description of the OH stretch in water, here we use the water-based OH wave function to describe the OH stretches in H_3O^+. We have also performed these calculations using guiding functions based on the ground state wave function for an OH stretch in H_3O^+ and obtain results that are nearly identical to those reported in Figure 3.2A (see Table 3.2).

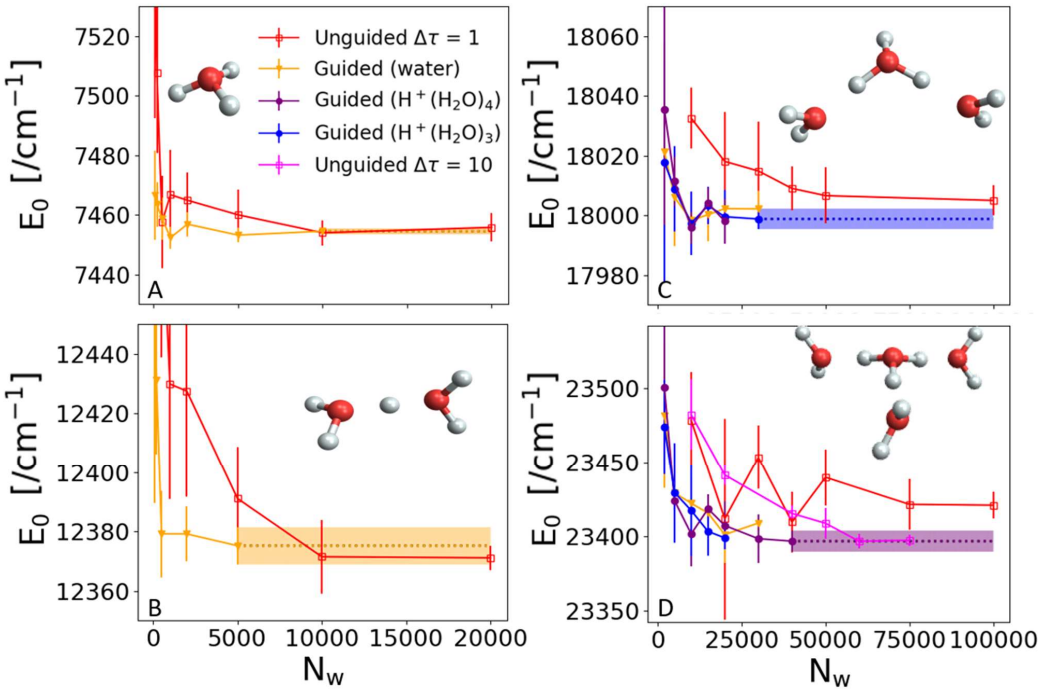


Figure 3.2: Calculated zero-point energies for (A) H_3O^+, (B) $H^+(H_2O)_2$, (C) $H^+(H_2O)_3$, (D) $H^+(H_2O)_4$ plotted as functions of the number of walkers used in the simulation. The red or pink symbols and lines provide results of unguided simulations using a 1 a.u. and 10 a.u. time step, respectively. The gold, blue and purple symbols and lines provide results of three types of guided simulations, which are described in the text. The dotted line and shading extends the results of the largest guided simulation in each panel to facilitate comparisons with other calculations. The energies used to generate these plots are also provided in Tables 3.2-3.5.

Table 3.2: Convergence of Zero-Point Energy for H_3O^+ in cm^{-1} from Figure 3.2

N_W	Guided(water)[a]	Guided(H_3O^+)[b]	Unguided
100	7467 (17)	7459 (26)	7535 (47)
200	7459 (8)	7475 (9)	7508 (30)
500	7457 (8)	7454 (11)	7458 (17)
1000	7455 (4)	7455 (8)	7467 (17)
2000	7457 (5)	7454 (5)	7465 (10)
5000	7453 (3)	7453 (2)	7460 (10)
10000	7455 (1)	7454 (1)	7454 (5)
20 000[c,d,e]	-	-	7456 (5)

[a] Guiding function using OH stretch wave functions from an isolated water molecule.
[b] Guiding function using OH stretch wave functions from an isolated hydronium molecule.
[c] Previous literature value based on 20 000 walkers and $\Delta\tau = 1$ a.u. is 7453 (2) cm^{-1}.[98]
[d] Unguided energy based on 20 000 walkers and $\Delta\tau = 10$ a.u., averaged over 15 000 time steps (150 000 a.u.) is 7448 (1) cm^{-1}.
[e] Unguided energy based on 20 000 walkers and $\Delta\tau = 10$ a.u., averaged over 15 000 a.u. is 7449 (3) cm^{-1}.

A similar approach was applied to calculations of the zero-point energies of $H^+(H_2O)_n$ for $n = 2$, 3 and 4, and the results are shown with the red and gold symbols and lines in panels B, C and D of Figure 3.2. For $n = 2$, shown in panel B, the energies obtained using the guided DMC approach do not change significantly when 500 or 5000 walkers are used, although the uncertainties of the results decrease as larger ensembles are used. For the unguided simulations, 10 000 walkers are needed to achieve similar results. These results are about 20 cm^{-1} lower than previously reported zero-point energies for this ion, 12 393 (5).[123] The difference reflects small differences in the potential surface used here, which is based on the potential for larger protonated water clusters, developed by Yu and Bowman.[8] Similar improvements in performance are found for the clusters with three and four water molecules, shown with red and gold symbols and lines in panels C and D of Figure 3.2.

Table 3.3: Convergence of Zero-Point Energy for $H^+(H_2O)_2$ in cm^{-1} from Figure 3.2

N_W	Guided[a]	Unguided
100	12 463 (82)	12 872 (306)
200	12 431 (28)	12 767 (141)
500	12 379 (16)	12 489 (56)
1000	-	12 430 (43)
2000	12 379 (10)	12 427 (40)
5000	12 375 (7)	12 391 (20)
10 000	-	12 371 (14)
20 000[b,c]	-	12 371 (5)

[a] Guiding function using OH stretch wave functions from an isolated water molecule and harmonic HOH bend wave functions.

[b] Unguided energy based on 20 000 walkers and $\Delta\tau = 10$ a.u., averaged over 15 000 time steps (150 000 a.u.) is 12 362 (4) cm^{-1}.

[c] Unguided energy based on 20 000 walkers and $\Delta\tau = 10$ a.u., averaged over 15 000 a.u. is 12 358 (6) cm^{-1}.

For $H^+(H_2O)_3$, both the guided and unguided DMC calculations give a zero-point energy of roughly 18 000 cm^{-1} while for the guided simulations this result is achieved with as few as 2000 walkers, at least 20 000 walkers are required to obtain this energy using the unguided approach. In both cases the uncertainties are uncomfortably large. The energies are within 5 cm^{-1} of the converged zero-point energy when 10 000 walkers for the guided calculation and 100 000 are required to obtain similar accuracy from the unguided calculations.

In the case of $H^+(H_2O)_4$, the guided DMC simulations give a zero-point energy of 23 409 (6) cm^{-1} when 30 000 walkers are used. The unguided simulations give an energy of 23 421 (10) cm^{-1} when 100 000 walkers were propagated using a 1 a.u. time step. This energy is 11 cm^{-1} above the energy obtained using the guided approach. For comparison, when we perform large unguided simulations with a larger 10 a.u. time step we obtain a zero-point energy of 23 397 (4) cm^{-1}, which is in very good agreement with the guided results. The convergence behavior for $H^+(H_2O)_4$ is similar to the behavior we noted for

Table 3.4: Convergence of Zero-Point Energy for $H^+(H_2O)_3$ in cm^{-1} from Figure 3.2

N_W	Guided ($H^+(H_2O)_3$)[a]	Guided (water)[b]	Guided ($H^+(H_2O)_4$)[c]	Unguided
2000	18018 (45)	18021 (33)	18035 (43)	-
5000	18009 (16)	18006 (18)	18011 (9)	-
10000	17997 (12)	17998 (8)	17996 (6)	18033 (11)
15000	18003 (7)	18000 (10)	18004 (6)	-
20000[d,e]	18000 (10)	18002 (4)	17998 (8)	18018 (18)
30000	17999 (3)	18004 (6)	-	18015 (19)
40000	-	-	-	18009 (8)
50000	-	-	-	18007 (11)
100000	-	-	-	18005 (6)

[a] Guiding function using OH stretch wave functions from an isolated water molecule and harmonic HOH bend wave functions with the shared proton stretch described using the $H^+(H_2O)_3$ parameters from Figure 3.5.

[b] Guiding function using OH stretch wave functions from an isolated water molecule and harmonic HOH bend wave functions.

[c] Guiding function using OH stretch wave functions from an isolated water molecule and harmonic HOH bend wave functions with the shared proton stretch described using the $H^+(H_2O)_4$ parameters from Figure 3.5.

[d] Unguided energy based on 20 000 walkers and $\Delta\tau = 10$ a.u., averaged over 15 000 time steps (150 000 a.u.) is 18 004 (6) cm^{-1}.

[e] Unguided energy based on 20 000 walkers and $\Delta\tau = 10$ a.u., averaged over 15 000 a.u. is 18 006 (18) cm^{-1}.

Table 3.5: Convergence of Zero-Point Energy for $H^+(H_2O)_4$ in cm^{-1} from Figure 3.2

N_W	Guided ($H^+(H_2O)_4$)[a]	Guided (water)[b]	Guided ($H^+(H_2O)_3$)[c]	Unguided[d]	Unguided[e]
2000	23 501 (63)	23 481 (55)	23 474 (36)	-	-
5000	23 424 (8)	23 429 (20)	23 429 (38)	-	-
10 000	23 401 (24)	23 422 (18)	23 417 (34)	23 478 (36)	23 482 (27)
15 000	23 418 (11)	23 416 (7)	23 403 (18)	-	-
20 000	23 407 (18)	23 401 (21)	23 399 (9)	23 412 (76)	23 441 (8)
30 000	23 398 (18)	23 409 (6)	-	23 453 (24)	-
40 000	23 397 (8)	-	-	23 409 (23)	23 415 (6)
50 000	-	-	-	23 440 (21)	23 409 (12)
60 000	-	-	-	-	23 397 (5)
75 000	-	-	-	23 421 (19)	23 397 (4)[f]
100 000	-	-	-	23 421 (10)	-

[a] Guiding function using OH stretch wave functions from an isolated water molecule and harmonic HOH bend wave functions with the shared proton stretch described using the $H^+(H_2O)_4$ parameters from Figure 3.5.

[b] Guiding function using OH stretch wave functions from an isolated water molecule and harmonic HOH bend wave functions.

[c] Guiding function using OH stretch wave functions from an isolated water molecule and harmonic HOH bend wave functions with the shared proton stretch described using the $H^+(H_2O)_3$ parameters from Figure 3.5.

[d] Based on $\Delta\tau = 1$ a.u., averaged over 15 000 a.u.

[e] Based on $\Delta\tau = 10$ a.u., and averaging over 15 000 time steps (150 000 a.u.).

[f] Unguided energy based on 20 000 walkers and $\Delta\tau = 10$ a.u., averaged over 15 000 a.u. is 23 414 (15).

$(H_2O)_6$ in a recent paper by a previous group member.[35] As in that work, significantly larger ensembles are required for the unguided simulations than for the guided ones. In the case of $(H_2O)_6$, 1 000 000 walkers were needed to obtain accurate zero-point energies using unguided DMC approaches.[35,99] In that work, we also found that the large unguided calculations that employed the 1 a.u. time step appeared to be converging to an energy that was larger than the value obtained from the guided DMC calculations.

We can also compare the zero-point energies obtained by the guided DMC approach to the value obtained when 20 000 walkers are used, in an unguided simulation with a 10 a.u. time step. For $H^+(H_2O)_3$, the calculation with this larger time step appears to be converged at this ensemble size, and nearly identical zero-point energies are obtained when 25 000 or 15 000 walkers are used. In these cases, the calculated energy agrees with the values obtained from unguided simulations with ensembles of 40 000 or more walkers when the 1 a.u. time step was used. In the case of $H^+(H_2O)_4$, the DMC simulation with 20 000 walkers and a 10 a.u. time step gives a zero-point energy of 23 441 (8) cm^{-1}, which is 32 cm^{-1} higher than the value obtained using the guided approach. When we increase the ensemble size used for the unguided simulation to 75 000 walkers, the agreement between the energy calculated by this approach is in excellent agreement with the values obtained using the guided approach. As with $(H_2O)_6$, the larger time step compensates for sampling issues. While the total propagation times for the 1 a.u. and 10 a.u. simulations differ, if we evaluate the results of the 10 a.u. simulations using the same propagation times used for the 1 a.u. ones, the energies do not change significantly, but the uncertainties are larger when the smaller amount of propagation time is used (see Tables 3.3 to 3.5).

One concern with using smaller ensemble sizes in the DMC simulations is that the potential may not be as well-sampled. To explore the effect of ensemble size on the description of the wave function, we focus on the OOO bend angle in $H^+(H_2O)_3$. We choose this angle because earlier work exploring projections of the probability amplitude for this cluster ion onto various internal coordinates indicated that this was among the most problematic.[122] This can be seen in the plots of projections of the probability amplitude onto several internal

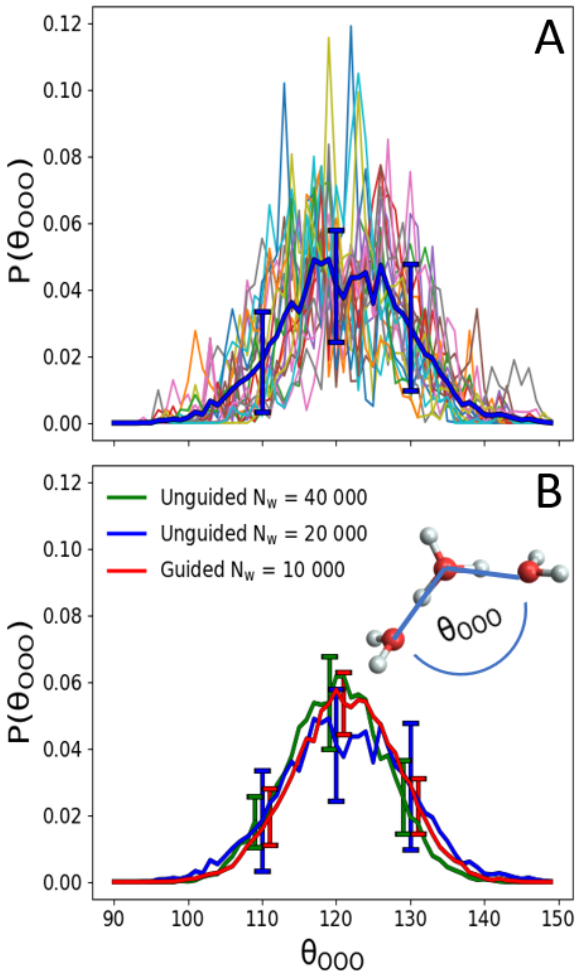


Figure 3.3: (A) Projections of the ground state probability amplitude for $H^+(H_2O)_3$ onto the OOO angle based on the probabilty amplitudes obtained from 20 DMC wave functions (thin lines). The average of these distributions is plotted with the thick blue line, and the standard deviations at 110°, 120° and 130° are represented by error bars. These results are based on an unguided simulation with 20 000 walkers. (B) Comparison of the average of the projected probability amplitudes based on twenty DMC wave functions obtained from unguided DMC simulations with 20 000 walkers (blue) and 40 000 walkers (green). In addition the projected probability amplitude obtained from a guided DMC simulation with 10 000 walkers is shown in red. For the guided calculation, the guiding function is based on the OH stretches of the outer water molecules and unbound OH bonds in the hydronium core as well as the HOH bends in the outer water molecules. The individual probability distributions used to obtain the green and red curves in panel B are shown in Figure 3.4.

coordinates based on unguided simulations with 20 000 walkers, in the left panels of Figure 3.4, with the projection onto the OOO angle also shown in the results reported in the upper panel of Figure 3.3. For these plots, we show the probability distributions obtained by projecting the probability amplitude for twenty wave functions onto the angle of interest. As is seen, while the average distribution looks reasonable (thick blue line) there are large fluctuations among the results that are obtained from the individual wave functions. This is reflected by the noisiness of the curves plotted in different colors as well as through the error bars that are shown for 110°, 120° and 130°. These large fluctuations among the results obtained using different wave functions reflect the correlation between this angle and the higher frequency HOH bend involving the bound OH bonds in the hydronium core. It also reflects changes in the optimized value of these angles as the free OH bond is displaced out of the plane of the three oxygen atoms (basically the umbrella motion of the hydronium core). When the hydronium core is planar, the optimized value of the OOO angle is close to 120°, while in the equilibrium geometry the OOO angle is closer to 113°. These couplings combined with the difference in the frequencies of these vibrations makes the projection of the ground state probability amplitude onto this coordinate difficult to capture using standard DMC approaches. Increasing the ensemble size mitigates the problem somewhat, as can be seen by comparing the size of the error bars for the curves plotted in green (40 000 walkers) and blue (20 000 walkers). The distributions from individual wave functions obtained in these calculations are provided in the middle panel of Figure 3.4. Interestingly, although increasing the time step used for the simulations with 20 000 walkers to 10 a.u. leads to better convergence behavior for the energy, the projections of the probability amplitudes obtained using the 10 a.u. time step with 20 000 walkers looks very similar to those reported in Figure 3.3A. If we compare these results to those obtained when the guided DMC approach is used with only 10 000 walkers (red curve in Figure 3.4B), we find that the smaller ensemble provides a further improvement to the results.

As with the neutral water clusters, the origins of the improved convergence behavior when guided DMC approaches are used can be traced to the fact that displacements of the

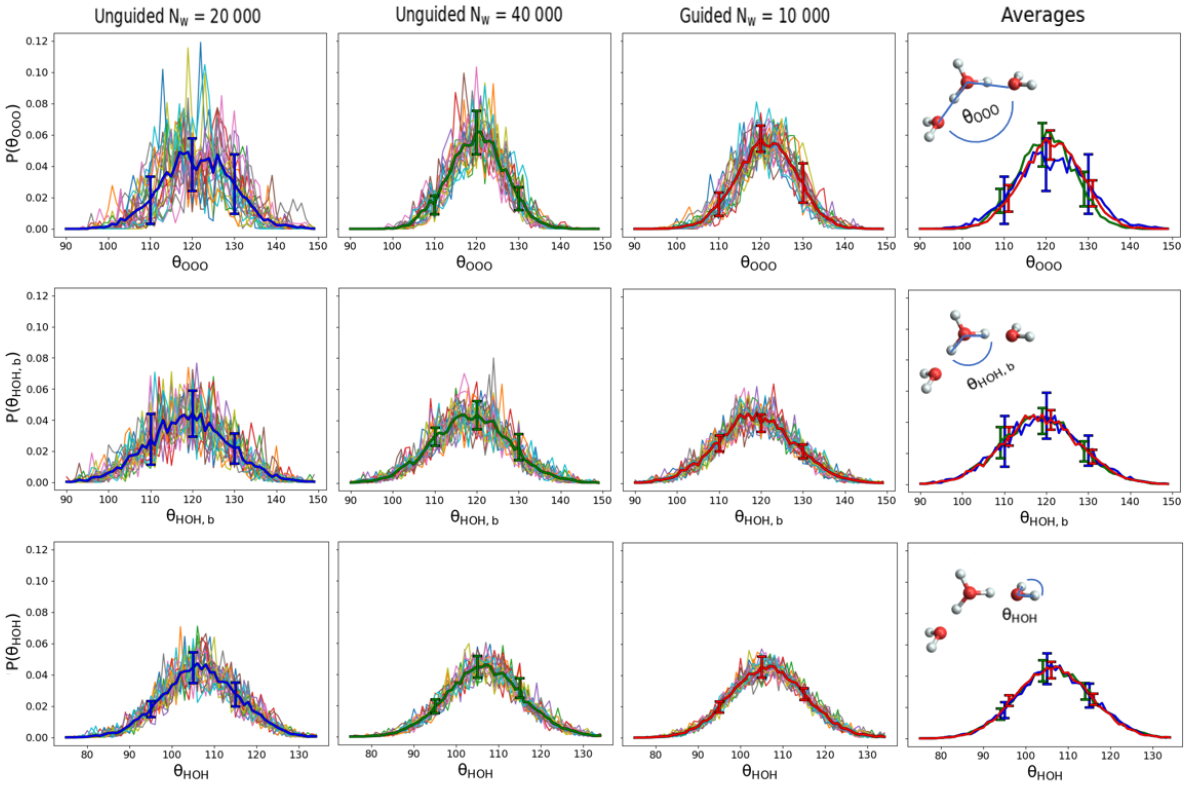


Figure 3.4: Projections of the ground state probability amplitude onto the OOO angle, θ_{OOO}, the bound HOH angle, θ_{HOH_b}, and the outer HOH angle, θ_{HOH}, in $H^+(H_2O)_3$. Twenty projections are shown using the unguided approach with 20 000 walkers, the unguided approach with 40 000 walkers, and the guided approach based on the OH stretches of the outer water molecules and unbound OH bonds in the hydronium core as well as the HOH angles in the outer water molecules with 10 000 walkers. The thicker lines show the average of these 20 projections with standard deviations when θ_{OOO} or $\theta_{\mathrm{HOH}_b} = 110°$, 120°, and 130° and when $\theta_{\mathrm{HOH}} = 95°$, 105°, and 115°. The last column shows the averages of these projections in the same plot to directly compare the various methods.

high frequency OH stretches lead to large fluctuations in the potential energy, particularly when compared to the lower-frequency vibrations. This makes simultaneous sampling of the high and low-frequency modes less efficient. By introducing guiding functions for the high frequency vibrations, particularly the unbound OH stretches, the low-frequency motions are sampled on what is effectively an adiabatic potential surface in which the full potential has been averaged over these high frequency motions.

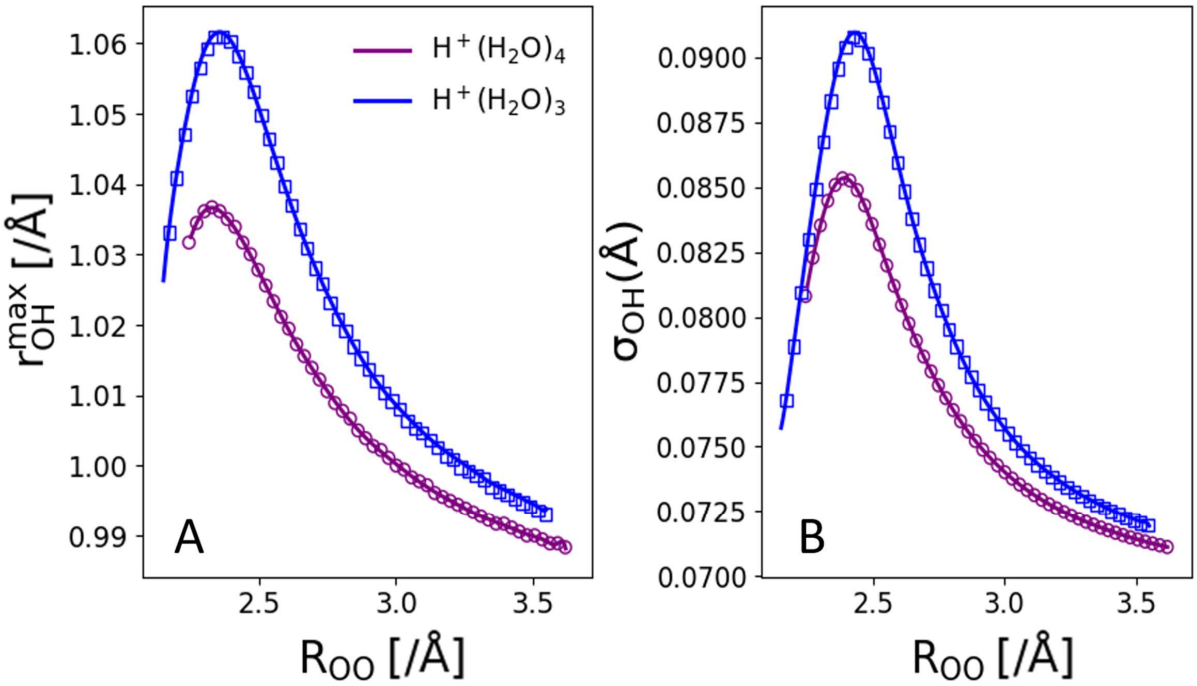


Figure 3.5: The maximum (A) and width (B) of the ground state wave functions for a water bound OH bond in the hydronium core of $H^+(H_2O)_3$ (blue curves) and $H^+(H_2O)_4$ (purple curves) are plotted as functions of the distance between the oxygen atoms in the hydronium core and the bound water molecule. The curves provide a $(n-1)$th order polynomial representation of the n data points that are plotted.

The question naturally arises as to whether the approach could be improved by incorporating the bound OH stretches in the hydronium core in the importance sampling scheme. As noted above, the frequency of these OH oscillators is sensitive to the bonding environment. A recent study from our group of the effects of solvation environment on the vibrational frequencies of the bound OH stretches in the hydronium core of protonated water clusters illustrated that the frequencies of these OH oscillators can be correlated to the distance between the oxygen atoms in the hydronium core and the associated water molecule.[7] Additionally, as with neutral water molecules, stronger hydrogen bonds are associated with both lower OH stretch frequencies and longer OH bonds in the donor hydronium molecule. In the case of $H^+(H_2O)_3$ and $H^+(H_2O)_4$, the equilibrium OH bond lengths differ by 0.025 Å, based on the potential used in this study, while the reported frequencies of these vibrations differ by roughly 550 cm^{-1}. Based on these observations, we cannot expect that the strategy of using a single trial wave function to describe an arbitrary environment for the OH bonds in water clusters will work for the water solvated OH bonds in hydronium core.

To address the above observations and concerns, we developed a modified strategy in which we drew on the correlations between the distance between the oxygen atoms in the accepting water molecule and the donating hydronium core, and the location of the maximum and the value of the width of this OH stretch wave function. For these calculations we evaluated one-dimensional scans of the potential as a function of the OH bond length for various OO distances keeping all other coordinates in their geometries based on the equilibrium structure of the ion. Using these scans, we calculated the ground state wave function for the OH stretch using a discrete variable representation (DVR).[119] The results of this analysis are provided in Figure 3.5, where in panel A, the maximum in the ground state wave function, $r_{\mathrm{OH}}^{\mathrm{max}}$ is plotted, while in panel B, we plot the width of the distribution, σ. To obtain a trial wave function for the bound OH stretches in the hydronium core, we use the instantaneous OO distance along with the curves in Figure 3.5 to shift and scale the water-bound OH stretch wave function evaluated at the equilibrium geometry of the cluster.

The results that are obtained when we introduced guiding functions to describe the water

bound OH stretches in the hydronium core are provided with blue and purple curves in panels C and D of Figure 3.2. The differences between these two sets of calculations are the parameters used to describe the wave function. For the blue curve, the parameters and reference wave function are based on $H^+(H_2O)_3$, while the parameters and reference wave function for $H^+(H_2O)_4$ are used to generate the purple curve. As the results provided in Figure 3.2 and Tables 3.3 to 3.5 show, the convergence properties of these calculations are very similar to those obtained when guiding functions are only used to describe the OH stretches and HOH bends of the outer water molecules and the free OH stretch in the hydronium. While at one level the lack of significant improvement in the convergence properties of the results when we incorporated trial wave functions for the bound OH stretches in the hydronium core is disappointing, the fact that the convergence behavior has not deteriorated with the introduction of the trial wave function indicates that such a strategy should be successful as we look to larger clusters, where a single cluster displays a range of strengths of hydrogen bonded interactions.

3.3.2 CH_5^+

We next turn our attention to CH_5^+ and its deuterated analogues. While in the water and hydronium molecules, all of the OH bonds are equivalent by symmetry, the CH bonds in CH_5^+ are not equivalent in its minimum energy structure. Specifically, for the minimum energy geometry the CH bonds range in length from 1.09 to 1.20 Å. The frequencies that are obtained from potential cuts along each of these CH distances vary from 2384 to 3070 cm^{-1} (see Table 3.1 and Figure 3.1). On the other hand, once zero-point energy is introduced the five CH bonds become equivalent. This is due to the fact that the barriers that separate the 120 equivalent minima on the potential are lower than the zero-point energy in the low-frequency vibrations that are responsible for the isomerization. To account for this, we first consider a guiding function of the form

$$\Phi_T(\mathbf{x}) = \prod_{i=1}^{5} \phi\left(r_{\mathrm{CH}}^{(i)}\right) \tag{3.1}$$

where $r_{\mathrm{CH}}^{(i)}$ represents the $\mathrm{CH}^{(i)}$ bond length in CH_5^+ and ϕ represents the average of five wave functions that are obtained by solving the one-dimensional Schrödinger equation using one dimensional cuts of the potential energy surface along the five CH bond lengths. The resulting wave function is plotted as a purple dotted line in Figure 3.6, and the wave functions that correspond to the longest ($\mathrm{CH}^{(5)}$) and shortest ($\mathrm{CH}^{(2)}$) CH bond lengths are plotted with red dashed and gold solid lines, respectively.

When we use the average CH stretch wave function as the trial wave function in the DMC calculations, we find that the convergence behavior of the zero-point energy is no faster than the unguided approach. It also appears that the calculations are converging to an energy that is slightly too large. Specifically, a calculation based on 20 000 walkers yielded a zero-point energy of 10 926(4) cm^{-1}, which is approximately 10 cm^{-1} above the zero-point energy obtained using an unguided calculation with 20 000 walkers, 10 918(6) cm^{-1}. The reason the trial wave function provided by Eq. 3.1 is not appropriate for this problem is illustrated in the results provided in Figure 3.6. First, the 0.11 Å difference between the lengths of the $\mathrm{CH}^{(2)}$ and $\mathrm{CH}^{(5)}$ bonds in the minimum energy geometry means that the average wave function does not provide a very good approximation to the ground state wave function associated with either of these oscillators. This can be seen by comparing the three wave functions plotted in Figure 3.6A. For comparison, in a recent paper from our group involving the study of equilibrium geometries of water clusters containing two to six water molecules, we found that these equilibrium OH bond lengths differ by at most 0.04 Å.[104] The difficulties in using a single wave function to describe the five CH oscillators in CH_5^+ is further illustrated in a comparison of the local energy function obtained when E_{L} in Eq. 1.1 is evaluated using the potential cut along $r_{\mathrm{CH}}^{(5)}$ (solid black curve) in Figure 3.6B based on the the three wave functions shown in panel A. When we use the average wave function to calculate the local energy it shows a sizable increase at larger values of r_{CH}. In contrast, the local energy

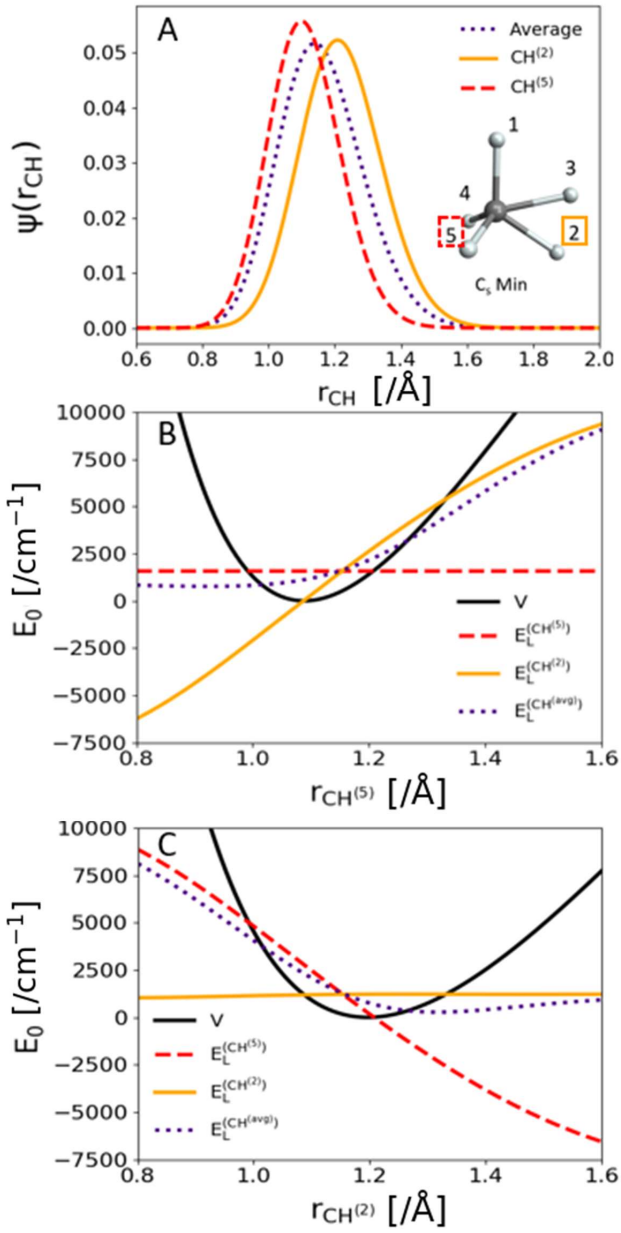


Figure 3.6: (A) Wave functions for the longest and shortest CH bonds in CH_5^+ (gold solid and red dashed line) as well as the average of the ground state wave functions for the five CH bonds (purple dotted line). (B and C) Local energies obtained from one-dimensional calculations using the potential along (B) $r_{\mathrm{CH}}^{(5)}$ and (C) $r_{\mathrm{CH}}^{(2)}$ (black line) based on the three wave functions shown in panel A.

plotted in the red dashed line is constant as the corresponding wave function is an eigenstate of this Hamiltonian. Similar behavior is found when we calculate the local energy using a cut through the potential through $r_{\mathrm{CH}}^{(5)}$ (panel C), although now the purple trace deviates from zero at small values of $r_{\mathrm{CH}}^{(2)}$. Based on these plots and the poor performance of guided DMC when the guiding function in Eq. 3.1 is used, we conclude that this approach that was effective for water clusters will not work well for CH_5^+. On the other hand, while the value of $\langle r_{\mathrm{CH}}^{(i)} \rangle$ depends on which CH bond is being considered, the widths of these distributions are less sensitive to the molecular environment.

To address this challenge, we need to find a procedure to relate $\langle r_{\mathrm{CH}}^{(i)} \rangle$ to the instantaneous structure of CH_5^+, just as we correlated the position and the widths of the wave functions for the bound OH stretches in $H^+(H_2O)_n$ to the OO distances. It has long been recognized that the structure of CH_5^+ can be described as a CH_3^+ part, in which the three CH bond lengths are roughly equal and the distances between these three hydrogen atoms are also equal, and a H_2 part, which is characterized by a shorter H-H distance.[124] In Tables 3.6-3.8, for each hydrogen atom, we report the H-H distances to each of the other four hydrogen atoms, evaluated at each of the three stationary point geometries on the potential of CH_5^+. The values based on the equilibrium structure of CH_5^+ are also plotted as functions of the corresponding CH bond lengths in the upper panel of Figure 3.7. For the CH bond lengths that involve hydrogen atoms in the H_2 group, the H-H distances range from 0.95 to approximately 2 Å. These are the two longest CH bonds and the breadth of the distances is illustrated by the blue and gold distribution in Figure 3.7A. For the hydrogen atoms in the CH_3 group the H-H distances only range from 1.7 to 1.9 Å, as is seen in the green, red and purple distributions. Similar behavior is seen for the other stationary point geometries. Based on this observation, we find that there is a correlation between the standard deviation among the four H-H distances involving a chosen hydrogen atom to the value of $\langle r_{\mathrm{CH}}^{(i)} \rangle$, which we have plotted in panel B

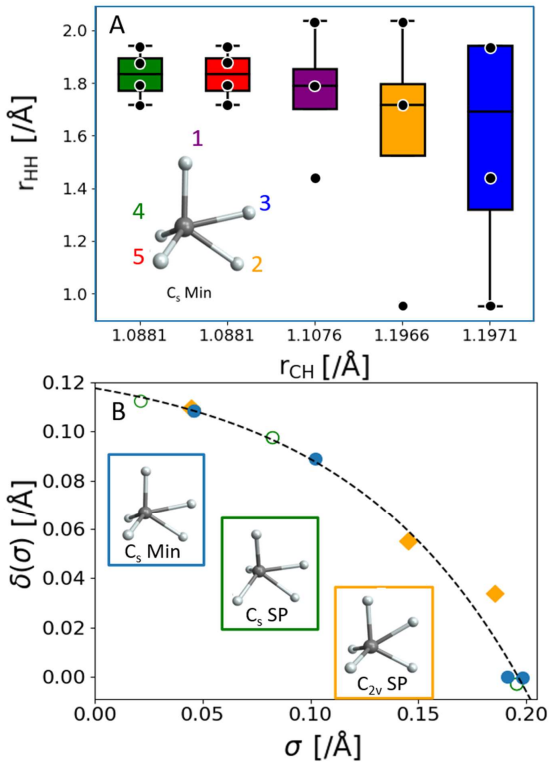


Figure 3.7: (A) Box and whisker plot showing the standard deviation of H-H distances for each of the five CH bonds on CH_5^+, plotted as a function of the CH bond length for the equilibrium structure of CH_5^+. (B) Plot of the displacement of the CH bond length from the value of $r_{CH}^{(2)}$ in the equilibrium geometry of CH_5^+, δ, as a function of standard deviation of the H-H distances to the hydrogen atom of interest σ, which is also plotted in panel A. These results are plotted for each of the three low-energy stationary point structures of CH_5^+, shown in the insets. The values for the C_s minimum are plotted with blue filled circles, the values for the C_s saddle point are plotted with green open circles, while the values for the C_{2v} saddle point are shown with filled gold diamonds. The dotted line provides a fit of these values to an exponential function of the form $\delta(\sigma_i) = Aexp(B\sigma_i) + C$. A, B, and C are -0.02389 Å, 6.29099 $Å^{-1}$, and 0.24620 Å respectively.

Table 3.6: HH Distances and Standard Deviations for the Minimum Energy Structure of CH_5^+

atom	$H^{(1)}$	$H^{(2)}$	$H^{(3)}$	$H^{(4)}$	$H^{(5)}$	σ (Å)^a	$\overline{\sigma}$ (Å)^b	$r_{CH,e}$ (Å)^c
$H^{(1)}$	-	2.0367	1.4413	1.7912	1.7912	0.2449	0.1023	1.1076
$H^{(2)}$	2.0367	-	0.9521	1.7182	1.7182	0.4613	0.1910	1.1966
$H^{(3)}$	1.4413	0.9521	-	1.9428	1.9428	0.4748	0.1983	1.1971
$H^{(4)}$	1.7912	1.7182	1.9428	-	1.8789	0.0985	0.0458	1.0881
$H^{(5)}$	1.7912	1.7182	1.9428	1.8789	-	0.0985	0.0458	1.0881

[a] Standard deviation of the four H-H distances.
[b] Normalized standard deviation among the four H-H distances obtained by scaling the CH bond lengths to 1 a_0.
[c] Equilibrium bond length.

Table 3.7: HH Distances and Standard Deviations for the C_s Saddle Point Structure of CH_5^+

atom	$H^{(1)}$	$H^{(2)}$	$H^{(3)}$	$H^{(4)}$	$H^{(5)}$	σ (Å)^a	$\overline{\sigma}$ (Å)^b	$r_{CH,e}$ (Å)^c
$H^{(1)}$	-	1.8471	1.8471	1.8527	1.8527	0.0033	0.0214	1.0842
$H^{(2)}$	1.8471	-	0.9444	2.0141	1.5485	0.4706	0.1956	1.1996
$H^{(3)}$	1.8471	0.9444	-	1.5485	2.0141	0.4706	0.1956	1.1996
$H^{(4)}$	1.8527	2.0141	1.5485	-	1.7564	0.1946	0.0825	1.0991
$H^{(5)}$	1.8527	1.5485	2.0141	1.7564	-	0.1946	0.0825	1.0991

[a] Standard deviation of the four H-H distances.
[b] Normalized standard deviation among the four H-H distances obtained by scaling the CH bond lengths to 1 a_0.
[c] Equilibrium bond length.

of Figure 3.7. To incorporate this observation into our trial wave function, we define

$$\Phi_T = \prod_{i=1}^{5} \phi\left(r_{CH}^{(i)} - \delta(\sigma_i)\right) \tag{3.2}$$

where $\phi(r_{CH})$ is the wave function associated with the lowest frequency CH oscillator in the equilibrium geometry, the gold curve in Figure 3.6A.

To obtain σ_i, we first scale all of the CH bond lengths so they are all the same value (1 a_0), and σ is the standard deviation of the four H-H distances involving the hydrogen atom

Table 3.8: HH Distances and Standard Deviations Among the C_{2v} Saddle Point Structure of CH_5^+

atom	$H^{(1)}$	$H^{(2)}$	$H^{(3)}$	$H^{(4)}$	$H^{(5)}$	σ (Å)^a	$\overline{\sigma}$ (Å)^b	$r_{CH,e}$ (Å)^c
$H^{(1)}$	-	1.1785	1.1785	1.9383	1.9383	0.4387	0.1852	1.1626
$H^{(2)}$	1.1785	-	2.0065	1.7485	1.7485	0.3498	0.1454	1.1414
$H^{(3)}$	1.1785	2.0065	-	1.7485	1.7485	0.3498	0.1454	1.1414
$H^{(4)}$	1.9383	1.7485	1.7485	-	1.9021	0.1002	0.0446	1.0870
$H^{(5)}$	1.9383	1.7485	1.7485	1.9021	-	0.1002	0.0446	1.0870

[a] Standard deviation of the four H-H distances.
[b] Normalized standard deviation among the four H-H distances obtained by scaling the CH bond lengths to 1 a_0.
[c] Equilibrium bond length.

of interest. The scaling is introduced to remove the effects of changes in the instantaneous CH bond lengths on the value of σ. In Figure 3.7B, we plot the difference between the associated CH bond length and $r_{CH}^{(2)}$ in the equilibrium geometry, δ, as a function of these scaled σ values. As is seen, δ increases monotonically with σ. We then fit this data to a shifted exponential,

$$\delta(\sigma_i) = A\exp(B\sigma_i) + C \tag{3.3}$$

Using this relationship, we have allowed the trial wave function to account for variations in the equilibrium CH bond length on the structure of the ion while keeping all of the hydrogen atoms equivalent. This does not add significant complexity to the DMC calculation.

The results of guided (blue) and unguided (red) DMC calculations of the ground states of CH_5^+, CD_5^+, and $CH_3D_2^+$ are provided in Figure 3.8 for various ensemble sizes. As is seen, in all of the isotopologues, when 2000 to 5000 walkers are used in the unguided simulations, the energies and associated uncertainties are comparable to the values obtained from unguided simulations with 20 000 walkers. Both values are also in good agreement with previously reported zero-point energy for this potential energy surface (black line with grey shading).[5]

In addition to confirming that the procedure is effective, these calculations also allowed us to explore the transferability of the trial wave functions with partial deuteration. The

only change in Φ_T in Eq. 3.2 when one or more of the hydrogen atoms are replaced with deuterium is the trial wave function that is used is the one appropriate for the ground state of the CD stretch based on the potential cut along $r_{\mathrm{CH}}^{(2)}$. We also explored how well this wave function performs on the various deuterated isotopologues of CH_5^+. One interesting feature of partially deuterated forms of CH_5^+ is that the probability amplitude becomes localized in a subset of the 120 minima on the potential. This localization reflects differences among the zero-point energies of the CH vibrations in the various bonding environments, as is illustrated by the wave functions plotted in Figure 3.1.

Overall, the agreement between the results of these calculations and previously reported zero-point energies for CH_5^+ is very good, further validating the approach. In the cases of CH_4D^+ and $CH_3D_2^+$ the energies obtained in the present chapter are between five and ten cm^{-1} higher than previously reported values based on both guided and unguided calculations. These differences are likely due to difficulties in sampling the high frequency CH stretches when a 10 a.u. time step is used, as was done in the previous studies of CH_5^+.[5] Studies of water monomer showed that the use of a 10 a.u. time step in DMC calculations resulted in zero-point energies that are about five cm^{-1} lower than the zero-point energy obtained from a converged variational calculation.[35,125] Interestingly, for D_2O the 10 a.u. time step provided an accurate value for the zero-point energy. The fact that we get slightly different zero-point energies when we reduce the time step from 10 a.u. to 1 a.u. for only CH_4D^+ and $CH_3D_2^+$ likely reflects fact that these are the mixed isotoplogues that have more hydrogen atoms than deuterium atoms.

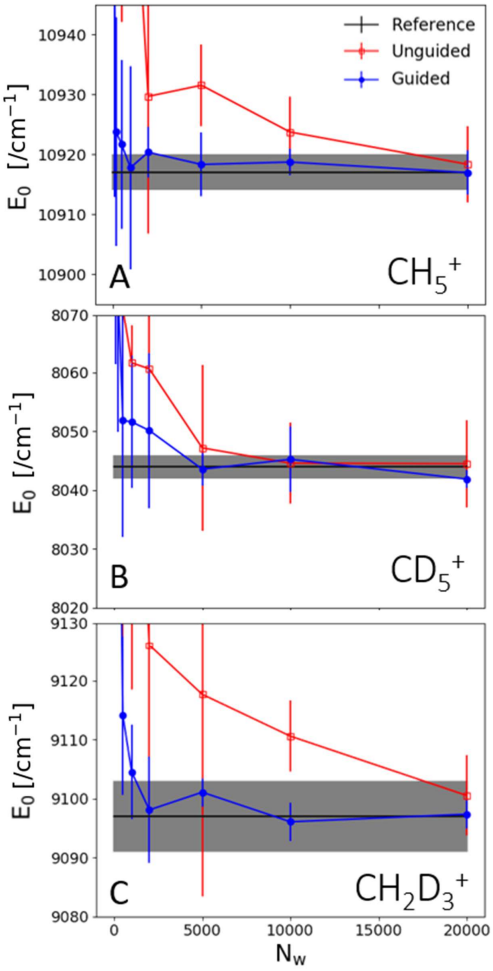


Figure 3.8: Calculated zero-point energies for (A) CH_5^+, (B) CD_5^+, and (C) $CH_2D_3^+$ plotted as a function of the number of walkers used in the simulation (N_W) for unguided (blue) and guided (red) DMC simulations. The black lines are the previously reported values of these zero-point energies, and the grey shading indicates their reported uncertainties.[5] The energies used to generate these plots are also provided in Table 3.9.

Table 3.9: Convergence of Zero-Point Energy for Deuterated Analogues of CH_5^+ in cm^{-1} Also Plotted in Figure 3.8

CH_5^+		
N_W	Guided	Unguided
100	10 949 (40)	11 228 (142)
200	10 924 (21)	11 096 (30)
500	10 922 (16)	10 971 (33)
1000	10 918 (19)	10 981 (25)
2000	10 920 (5)	10 930 (26)
5000	10 918 (6)	10 932 (8)
10 000	10 919 (2)	10 924 (7)
20 000	10 917 (4)	10 918 (7)
Reference[a]	10 917 (3)	
CD_5^+		
N_W	Guided	Unguided
100	8101 (45)	8332 (88)
200	8076 (29)	8158 (22)
500	8052 (22)	8071 (27)
1000	8052 (13)	8062 (7)
2000	8050 (15)	8061 (20)
5000	8044 (3)	8047 (16)
10 000	8045 (6)	8045 (8)
20 000	8042 (2)	8045 (8)
Reference[a]	8044 (2)	
$CH_2D_3^+$		
N_W	Guided	Unguided
100	9219 (56)	9488 (111)
200	9186 (62)	9299 (83)
500	9114 (15)	9167 (52)
1000	9105 (9)	9184 (73)
2000	9098 (10)	9126 (30)
5000	9101 (3)	9118 (38)
10 000	9096 (4)	9111 (7)
20 000	9097 (3)	9101 (8)
Reference[a]		9097 (6)

[a] Based on an unguided simulation with 20 000 walkers and $\Delta\tau = 10$ a.u.[5]

3.4 Conclusions

In this chapter, we have demonstrated that the guided DMC approach, which we recently developed for studies of water clusters,[27,35] can be used to study systems where the form of the vibrational wave function that describes the XH stretches depends on the local environment experienced by that bond. We applied the approach to studies of protonated water clusters with four or fewer water molecules and to CH_5^+. We showed that we could obtain substantial savings in the computational demands of the DMC simulations for these systems, in some cases as much as an order of magnitude, compared to unguided simulations. We also showed that the wave functions obtained from these smaller ensembles resulting from guided simulations were better converged than the wave functions obtained using unguided approaches with substantially larger ensemble sizes.

By being able to converge the zero-point energies and wave functions with smaller ensembles, we were able to show that for the protonated water systems, the ensemble sizes used in previous studies are likely not large enough to obtain accurate ground state properties.[8,30] With the approach validated and the savings that were achieved we are positioned to explore larger cluster ions, specifically $H^+(H_2O)_{5,6}$ where experimental studies have demonstrated that multiple isomers are sampled in the low-temperature experiments, and the most stable forms appear to be affected by partial deuteration.[7,11] The ability to separate the CH stretch vibrations from the lower frequency motions in CH_5^+ also provides an opportunity to further explore the five dimensional rotor model, which was successfully used by Schlemmer and co-workers to analyze their rotationally resolved spectrum.[106,126,127]

Chapter 4

ISOTOPE EFFECTS IN THE ZUNDEL-EIGEN ISOMERIZATION OF $H^+(H_2O)_6$

Reproduced in part with permission from [Jacob M. Finney, Tae Hoon Choi, Rachel M. Huchmala, Joseph P. Heindel, Sotiris S. Xantheas, Ken D. Jordan, Anne B. McCoy. Isotope Effects in the Eigen-Zundel Isomerization of $H^+(H_2O)_6$. *J. Phys. Chem. Lett.* **2023**, 14, 4666–4672] Copyright [2023] American Chemistry Society.

4.1 Introduction

One of the most fundamental questions concerning the nature of an excess proton in liquid water is how its structure depends on the local environment. Specifically, is there a persistent local structure that corresponds to either a solvated hydronium ion (e.g., the Eigen (E) form) or is the proton strongly associated with two water molecules (e.g., the Zundel (Z) form), or do these two limiting structures dynamically interconvert?[128–133] Experimental measurements of the vibrational spectra of protonated water clusters have proven especially valuable for elucidating the nature of an excess proton in various hydration environments.[7,130,134–144] These clusters are held together by strong hydrogen bonds, which in turn induce large red shifts of the bands associated with the OH stretches in their infrared (IR) vibrational spectra compared to the frequencies of the OH stretches in an isolated water molecule. The assignment of the vibrational spectra of protonated water clusters in the OH stretching region is further complicated by the presence of overtone and combination bands, which result from strong couplings among the vibrational motions. Even establishing which isomers are expected to have sizable populations at low temperature can be arduous due to the large differences in the vibrational zero-point energies (ZPEs) between the various isomers and the

effort required to accurately quantify the ZPEs beyond the harmonic approximation.

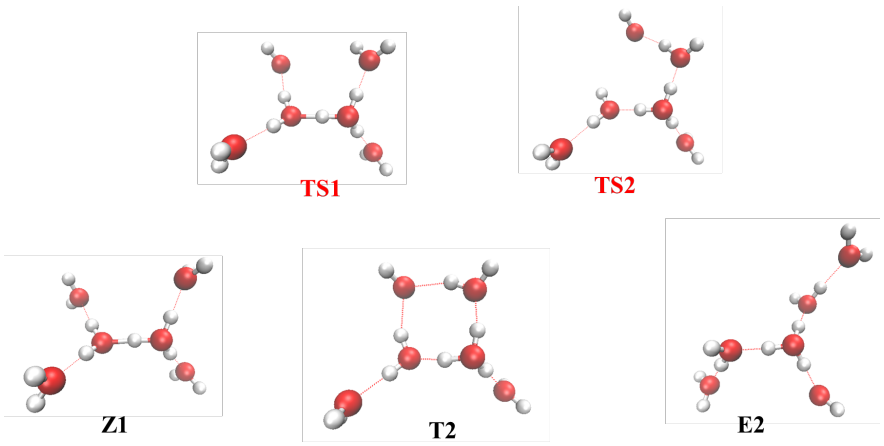


Figure 4.1: Structures of local minima (Z1, T2, E2) and transition states (TS1, TS2) of $H^+(H_2O)_6$ that are the focus of this study. Z1 and E2 represent Zundel- and Eigen-type isomers, respectively, while T2 corresponds to a stable intermediate structure that lies on the isomerization pathway between the Z1 and E2 isomers. TS1 is the transition state between Z1 and T2, and TS2 is the transition state between T2 and E2.

The present chapter addresses the above challenges by examining the isomerization pathway of the protonated water hexamer, $H^+(H_2O)_6$, which has been the subject of several experimental and theoretical studies.[7,12,55,140,145] This cluster is known to have multiple low-energy isomers, with the two most stable being Eigen- and Zundel-type in nature when vibrational ZPE is accounted for.[55,146] The structures of these two isomers are shown in Figure 4.1. Following earlier work on these clusters they are labeled E2 for the Eigen-type and Z1 for the Zundel-type.[55] On the basis of double resonance experiments, Heine *et al.* showed that both of these isomers are observed in low-temperature experimental studies of $H^+(H_2O)_6$ tagged with a hydrogen molecule.[12] The vibrational spectrum from this chapter is reproduced in the upper panel of Figure 4.2 along with the D_2 tagged spectrum of $D^+(D_2O)_6$ (lower panel).[7] Interestingly, the D_2 tagged spectrum of $D^+(D_2O)_6$ shows a very

different band structure compared to the H_2 tagged spectrum of $H^+(H_2O)_6$, which cannot be explained by consideration of the Eigen and Zundel isomers alone.[7]

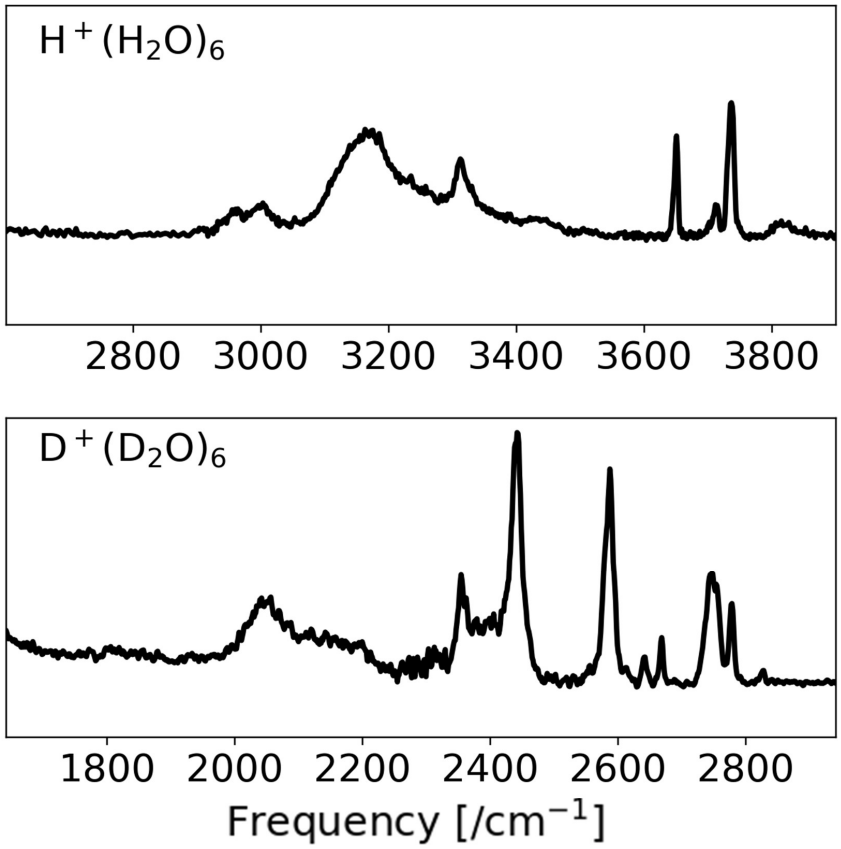


Figure 4.2: (Top) Experimental spectra of H_2 tagged $H^+(H_2O)_6$.[6] (Bottom) Experimental spectrum of D_2 tagged $D^+(D_2O)_6$.[7]

4.2 Methods

In this work, we use high-level electronic structure methods to accurately characterize the Born-Oppenheimer potential energy surface corresponding to the Z1-E2 isomerization path-

way of $H^+(H_2O)_6$ and apply advanced vibrational energy methods to probe the effects of zero-point energy and isotopic substitution on that isomerization pathway. We find that this pathway is more complex than previously assumed in that it involves a stable intermediate (T2, in Figure 4.1) with a four-membered ring and two single acceptor water molecules. The existence of this isomer may also help explain the large difference between the vibrational band structure of $H^+(H_2O)_6$ and $D^+(D_2O)_6$ shown in Figure 4.2.

4.2.1 Electronic Structure Calculations

A detailed investigation of the low-energy isomers of $H^+(H_2O)_6$ was performed in which the geometries of the Z1, T2, and E2 isomers as well as the TS1 and TS2 transition states, depicted in Figure 4.1, were optimized at the coupled cluster singles and doubles (CCSD) level[147] with the aug-cc-pVDZ basis set.[148] At each stationary point the harmonic vibrational frequencies of $H^+(H_2O)_6$ and $D^+(D_2O)_6$ were calculated at the same level of theory. The relative energies at the complete basis set (CBS) limit were inferred using both a three-parameter 4-5 polynomial extrapolation and double exponential extrapolation scheme.[149] These extrapolations were based on a series of CCSD(T) calculations,[150] which were performed using aug-cc-pVDZ, aug-cc-pVTZ, and aug-cc-pVQZ basis sets using the CCSD/aug-cc-pVDZ geometries. We also carried out the same extrapolation procedure where each calculation was performed using the geometries that were optimized at the MP2/aug-cc-pVnZ (n=D, T and Q) level of theory.[151–153] The resulting energies are summarized in Table 4.2. The two sets of calculations are in quantitative agreement with each other. All of these calculations were performed using Gaussian 16.[154] In addition, anharmonic ZPE corrections were evaluated using the diffusion Monte Carlo (DMC) method[22–24,114,155] together with the analytical potential for protonated water clusters developed by Yu and Bowman.[8] The DMC calculations performed in this work are based on the protocol employed in our earlier work on smaller protonated water clusters.[97]

4.2.2 Diffusion Monte Carlo Calculations

Our implementation of guided diffusion Monte Carlo has been described in Chapter 2 and only the details specific to this chapter are provided below.

All DMC calculations in this work are performed using the potential energy surface developed by Yu and Bowman.[8] In the previous chapter, we developed a general guiding function that can be used to describe the high-frequency intramolecular vibrations of protonated water clusters.[97] The guiding function is expressed as a product of one-dimensional wave functions that describe the OH and HOH bending vibrations of the water molecules and the OH stretching vibrations in the hydronium core. The one-dimensional wave functions that are used to describe vibrations in the water molecules are based on the ground state wave functions for the OH stretch in an isolated water molecule, while the HOH bend wave function is described by the ground state solution to the harmonic oscillator. Since the lengths of the OH bonds in the hydronium core and the frequency of the stretching vibrations depend strongly on the hydrogen-bonding environment, a flexible OH stretch wave function is used to describe the bound OH stretches in the hydronium core. This flexible guiding function is based on the OH stretch wave function used to describe the OH stretch in water, with the position of the maximum and the width of the wave function adjusted based on the instantaneous distance between the oxygen atom in the hydronium core and the oxygen atom of the solvating water molecule. This was achieved by first evaluating the ground state wave function for a bound OH stretching vibration in $H^+(H_2O)_3$ at various OO distances and developing a relationship between the maximum in the wave function and its width to the OO distance at which it was calculated. We found that the fit that was used in the previous chapter did not extrapolate well. For this work, we refit the data obtained in the previous chapter to

$$f(R_{\mathrm{OO}}) = \begin{cases} AR_{\mathrm{OO}}^2 + BR_{\mathrm{OO}} + C & R_{\mathrm{OO}} < t \\ \sum_{i=1}^{2} a_i \frac{R_{\mathrm{OO}} - x_{e,i}}{\sqrt{2\pi\sigma_i^2}} \exp\left[-\frac{1}{2}\left(\frac{(R_{\mathrm{OO}} - x_{e,i})}{\sigma_i}\right)^2\right] + c & R_{\mathrm{OO}} > t \end{cases} \tag{4.1}$$

where $f(R_{OO})$ represents either the maximum of the OH stretch wave function or the width of that function. A second fit was performed for $D^+(D_2O)_3$ as described. The parameters used to fit this data to Equation 4.1 are provided in Table 4.1 and the fits are plotted in Figure 4.3. Finally, since the shared proton in the Zundel core is highly delocalized, no guiding function is used to describe this bond.

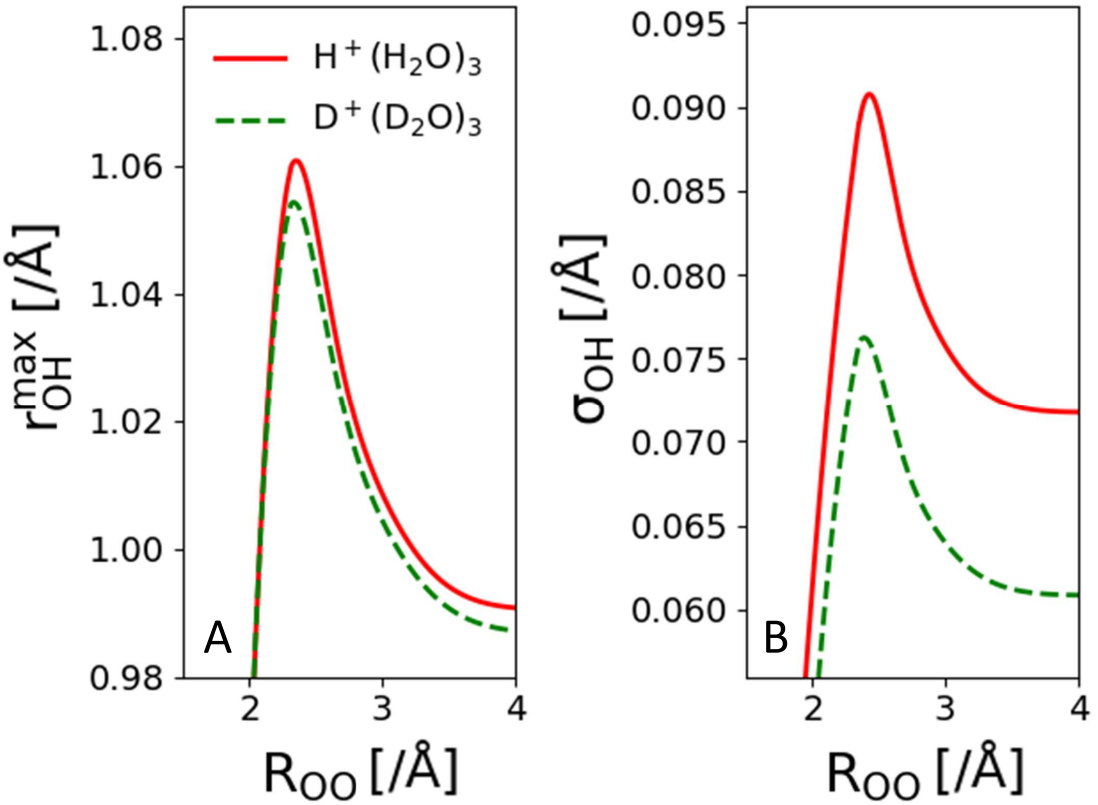


Figure 4.3: Fits to the maximum (r_{OH}^{max}, A) and widths (σ_{OH}, B) of the ground state wave functions for a water bound OH bond in the hydronium core of $H^+(H_2O)_3$ (red solid line) and $D^+(D_2O)_3$ (green dashed line) plotted as a function of the OO distance between the oxygen atoms in the hydronium core and the hydrogen bound water. Fits are based on Equation 4.1 and the parameters of these fits are reported in Table 4.1.

Table 4.1: Parameters Used to Construct the Fits in Figure 4.3 Based on Equation 4.1.

	OH stretch		OD stretch	
Parameter	$r_{\mathrm{OH}}^{\mathrm{max}}$	σ_{OH}	$r_{\mathrm{OH}}^{\mathrm{max}}$	σ_{OH}
A	-0.76	-0.06	-0.76	-0.09
B	3.57	0.35	3.55	0.48
C	-3.16	-0.38	-3.11	-0.53
a_1	0.14	0.04	0.13	0.03
a_2	0.16	0.04	0.15	0.03
$x_{e,1}$	1.90	1.95	1.83	1.89
$x_{e,2}$	2.06	2.23	2.03	2.17
σ_1	0.59	0.50	0.63	0.53
σ_2	0.26	0.19	0.28	0.21
c	0.99	0.07	0.99	0.06
t	2.31	2.35	2.31	2.35

For the DMC simulations, for each of the three isomers of $H^+(H_2O)_6$ and $D^+(D_2O)_6$ the initial ensemble of walkers was created using a shorter DMC calculation that started at the minimum energy geometry of that isomer. After the equilibration of the ensemble, three independent simulations were performed for each of the six isotopologues considered in this work. These calculations used 100 000 walkers and were run for 50 000 time steps with a time step of 1 a.u. Wave functions were collected every 1000 time steps and descendant weighting was run for an additional 250 time steps to collect Ψ^2. During these calculations, the walkers sometimes sampled unphysical geometries that were very low-energy based on the potential developed by Yu and Bowman[8] (e.g. holes in the fit potential). When this happened, the simulation was restarted from an earlier ensemble that was not affected by the hole in the potential.

4.2.3 VPT2 Calculations of the Spectra

The spectra described in the text were evaluated using second order vibrational perturbation theory, as implemented in Gaussian 16.[154] These calculations were performed at the MP2

level of theory, and using two basis sets. First we used the aug-cc-pVDZ basis. We also used an aug-cc-pVTZ basis in which we removed the d-orbitals from the description of the hydrogen atoms, which we will represent by aug-cc-pVTZ(-d). The removal of the d-basis functions on the hydrogen atoms was done because the size of the molecule and triple-zeta basis made the calculations too large to perform on the available computing resources. Additionally, for these VPT2 calculations, we used the perturbation theory to only calculate states with excitation in the 20 highest frequency modes. The modes that were considered include all of the OH stretching and HOH bending vibrations except for the low-frequency motion of the shared proton parallel to the OO axis in the Z1 isomer. The restriction to these states leads to significant reduction in the computational demands of these calculations without introducing significant approximation to the full VPT2 results.

It is well-known that VPT2 calculations of molecules, like the protonated water hexamers, that contain large-amplitude vibrations can suffer from numerical instabilities. These instabilities arise from the fact that the corrections to the wave function at nth order in perturbation theory are expressed in terms of the product of n ratios of the coupling between zero-order states to the harmonic energy difference between these states. If two (or more) states are close in energy at the harmonic level, these ratios can become large. In the calculations of energies, this is accounted for by the Martin test,[156–158] which identifies nearly degenerate states using a criterion that is based on energy considerations, and a small variational calculation is performed in the sub-space of nearly degenerate states. We have found that for the most part, this approach provides reasonable energies for the isotopomers of $H^+(H_2O)_6$ considered in this work. In the few cases where the perturbation theory yields unphysically large anharmonic shifts, these are readily identified, for example by comparing energies of levels associated with vibrational modes that have similar harmonic frequencies.

While the frequencies that are obtained by the VPT2 calculations are reasonable, for several isotopomers the intensities were found to be unphysically large. This reflects the fact that intensities for 1-0 transitions require the second order correction to the vibrational wave function, which involves products of two ratios of the coupling terms and the harmonic

energy difference. As a result, these intensities can be more sensitive to near degeneracies compared to the energies. As we expect that the intensities of the 1-0 levels should be similar to the harmonic intensities, we have used the harmonic intensities for the calculated spectra shown in Figure 4.6.

4.3 Results

Consistent with the study of Klein *et al.*,[146] the results of these calculations, which are provided in Figure 4.4, indicate that when vibrational zero-point energy corrections are included, the two lowest energy isomers of $H^+(H_2O)_6$ are the Eigen and Zundel species. Moreover, our calculations predict a local minimum intermediate, labeled as T2 in Figure 4.1, that is only a few tenths of a kcal/mol higher in energy compared to the E2 and Z1 isomers. Although the T2 isomer was considered in earlier studies,[55,146] it was not identified as an intermediate for the Zundel-Eigen isomerization pathway.

Depending on which bond of the 4-membered ring of T2 is broken, the pathway leads to either the E2 or the Z1 isomer. The relative energies of the stationary points are drastically altered upon inclusion of harmonic ZPE corrections. For $H^+(H_2O)_6$, when harmonic ZPE corrections are added to the CCSD(T)/CBS energies, the E2 and T2 isomers are predicted to lie 0.33 and 0.46 kcal/mol above the Z1 isomer, respectively. Moreover, the barrier for T2-Z1 isomerization is predicted to be only 0.06 kcal/mol (see Figure 4.4a).

For $D^+(D_2O)_6$ (Figure 4.4B), the CCSD(T) calculations with harmonic ZPE corrections predict the Z1, T2, and E2 isomers to be essentially isoenergetic, with the E2 isomer lying only 0.04 kcal/mol above Z1, and the T2 isomer being equal in energy to Z1. Moreover, for $D^+(D_2O)_6$ the calculations predict a 0.46 kcal/mol barrier for isomerization from Z1 to T2, and a barrier of 0.51 kcal/mol for isomerization from T2 to E2. These results would lead one to expect significant populations of all three of the Z1, T2, and E2 isomers of $D^+(D_2O)_6$ even at temperatures of a few tens of Kelvin. However, there remains the possibility that the inclusion of anharmonicity in the ZPEs could significantly alter the relative energies of the three isomers in question.

Table 4.2: Energies of the T2 and E2 Isomers and the Transitions States, Reported Relative to the Energy of the Z1 Isomer.[a,b]

Energy Differences (kcal/mol)			
Extrap. Method	MP2//MP2	CCSD(T)//MP2	CCSD(T)//CCSD
T2			
CBS (4-5 poly.)[c]	−1.13	−1.34	−1.29
CBS (double exp.)[d]	−1.11	−1.33	−1.31
CBS[e]	-1.12 ± 0.02	-1.33 ± 0.01	-1.30 ± 0.01 [f]
E2			
CBS (4-5 poly.)[c]	−0.70	−0.79	−0.75
CBS (double exp.)[d]	−0.62	−0.71	−0.70
CBS[e]	-0.66 ± 0.06	-0.75 ± 0.06	-0.73 ± 0.03 [f]
TS1			
CBS (4-5 poly.)[c]	0.50	0.47	0.31
CBS (double exp.)[d]	0.48	0.44	0.31
CBS[e]	0.49 ± 0.02	0.46 ± 0.03	0.32 ± 0.01 [f]
TS2			
CBS (4-5 poly.)[c]	0.02	−0.1	−0.21
CBS (double exp.)[d]	0.04	−0.05	−0.18
CBS[e]	0.03 ± 0.02	-0.05 ± 0.00	-0.198 ± 0.02 [f]

[a] See text for details.
[b] All raw energies used in these extrapolations are provided in the excel file along with results of additional calculations.
[c] Obtained from a fit to $E_n = E_{\mathrm{CBS}} + a/(n+1)^4 + b/(n+1)^5$, for $n = 2, 3$ and 4.
[d] Obtained from a fit to $E_n = E_{\mathrm{CBS}} + a\exp[-(n-1)] + b\exp[-(n-1)^2]$, for $n = 2, 3$ and 4.
[e] Average of the results of the two extrapolations.
[f] The results reported Figure 4.4.

This issue was addressed by carrying out DMC calculations of the ground states of the three isomers of $H^+(H_2O)_6$ and $D^+(D_2O)_6$. The results of this analysis are summarized in Figures 4.4C and 4.4D. In the absence of ZPE corrections, the relative energies from the Yu and Bowman potential[8] and from the *ab initio* calculations are in good agreement,

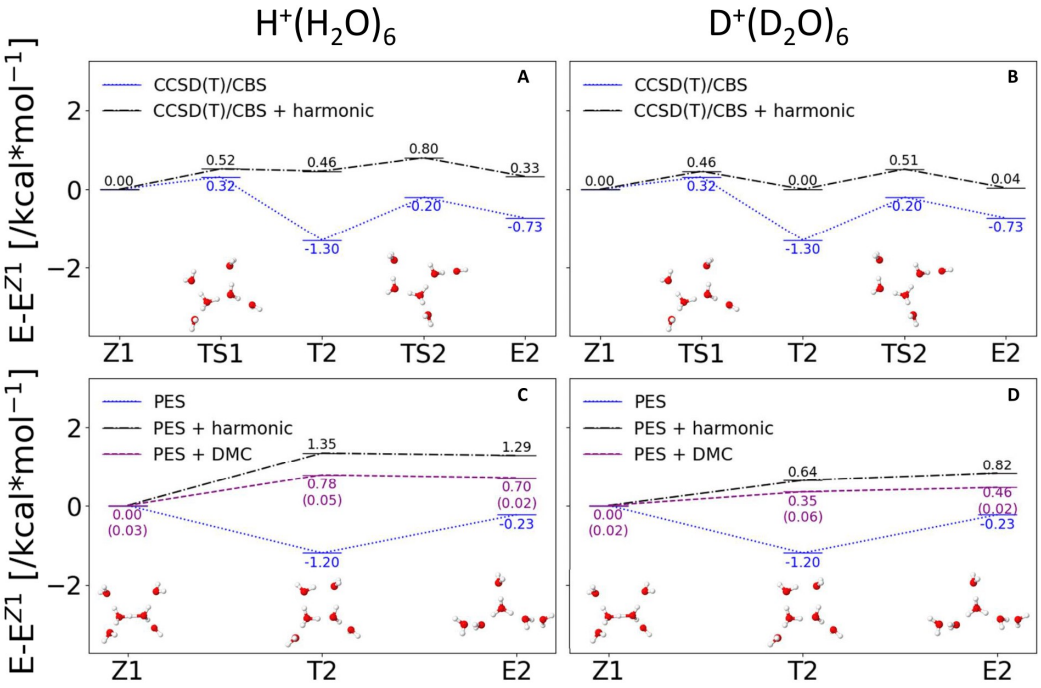


Figure 4.4: Pathway for the Z1-E2 isomerization both with and without the inclusion of harmonic and anharmonic vibrational zero-point energy corrections for $H^+(H_2O)_6$ and $D^+(D_2O)_6$. For panels A and B the electronic energies are calculated at the CCSD(T) level of theory, extrapolated to the CBS limit, based on the optimized structures obtained at the CCSD/aug-cc-pVDZ level of theory. The harmonic zero-point corrections were evaluated at the CCSD/aug-cc-pVDZ level of theory. For panels C and D, all geometries are optimized based on the potential of Yu and Bowman (PES),[8] and the harmonic and anharmonic (DMC) zero-point energy corrections were obtained using this potential surface. The relative energies, which are reported next to the corresponding line, are also provided in Table 4.3. Numbers in parentheses reflect the uncertainties in the reported values.

Table 4.3: Energies (in kcal/mol) of the Stationary Points for $H^+(H_2O)_6$ and $D^+(D_2O)_6$, Plotted in Figure 4.4.

Stationary Point	CCSD(T)/CBS	CCSD(T)/CBS + harmonic		CCSD(T)/CBS + harmonic (D)	
Z1	0.00	0.00		0.00	
TS1	0.32	0.52		0.46	
T2	−1.30	0.46		0.00	
TS2	−0.20	0.80		0.51	
E2	−0.73	0.33		0.04	
Stationary Point	PES	PES + harmonic	PES + DMC	PES + harmonic (D)	PES + DMC (D)
Z1	0.00	0.00	0.00 (0.03)	0.00	0.00 (0.02)
T2	−1.20	1.35	0.78 (0.05)	0.64	0.35 (0.06)
E2	−0.23	1.29	0.70 (0.02)	0.82	0.46 (0.02)

differing by less than 0.5 kcal/mol. For both $H^+(H_2O)_6$ and $D^+(D_2O)_6$, all of the calculations show that the T2 and E2 isomers become higher in energy than Z1 when harmonic zero-point energies are considered. Consistent with the CCSD(T) calculaitons, the T2 isomer is higher in energy than the E2 isomer for $H^+(H_2O)_6$, and the energy ordering is reversed for $D^+(D_2O)_6$. While there are quantitative differences between the energies evaluated at the CCSD(T) level and those obtained using the surface of Yu and Bowman,[8] the general trends are preserved. Additionally, while the use of anharmonic, rather than harmonic, ZPE corrections changes the values of the relative energies, it does not affect the energy ordering of the various isomers.

The low-energy barriers that separate the T2 isomer from the Z1 and E2 isomers could lead one to conclude that the vibrational wave function would be delocalized among the three isomers even at very low temperatures. Closer examination of the vibrational wave functions obtained from the DMC calculations shows that the wave function for each isomer remains localized in a single minimum in the potential energy surface. This can be seen most clearly when we project the corresponding probability amplitude onto the OO distances, providing the OO-radial distribution function for these clusters. The results are shown in Figure 4.5.

As can be seen in these plots, the OO-distance distributions each consist of three or four peaks, the centers of which are reported in Tables 4.4 - 4.6 along with their values in the optimized structures. Most notably the OO distance distributions for the three isomers are distinct. For OO distances below 5 Å, the distributions obtained from independent DMC simulations of each isomer are nearly identical. For larger OO distances, there are larger differences among these distributions, as indicated by the colored shading around these distributions. These larger differences between the results of different simulations reflect the large amplitude motions of the terminal water molecules in these clusters. Comparing the distributions for $H^+(H_2O)_6$ and $D^+(D_2O)_6$, the distributions for the deuterated systems show a slight decrease in the widths of the peaks, but overall the distributions are are unaffected by deuteration. The fact that the DMC calculations give a localized Z1, T2 and E2 structures for both $H^+(H_2O)_6$ and $D^+(D_2O)_6$ indicates that these isomers should exist as a well-defined forms of the protonated water hexamer with distinct spectral signatures. The existence of localized Z1 and E2 structures for $H^+(H_2O)_6$ is further supported by recent experimental studies by Khuu *et al.*[145] in which they excited the cold clusters in the OH stretch region and followed the relaxation of the vibrationally hot clusters back into the Z1 and E2 minima on the potential surface.

As noted above, the energy ordering of the isomers changes with the inclusion of zero-point effects and is also affected by deuteration. A similar energy reordering of isomers of water clusters was also noted for $(H_2O)_6$ based on microwave studies by Pate and co-workers[159,160] and subsequently through DMC calculations.[35,99] While qualitatively the relative electronic energies follow the number of hydrogen bonds (the T2 structure contains one more hydrogen bond than the other two structures and has a lower electronic energy), overall the relative zero-point energies will reflect both the strength of the hydrogen bonds and the number of dangling water molecules. Comparing the number of dangling water molecules, Z1 has four, E2 has three and T2 has only two. The strength of the hydrogen bonds can be related to the linearity of the HOO hydrogen bond angle. While the equilibrium structures have HOO angles that deviate from linearity by up to 3.5° in the Z1 and E2 structures and 15.5° in T2,

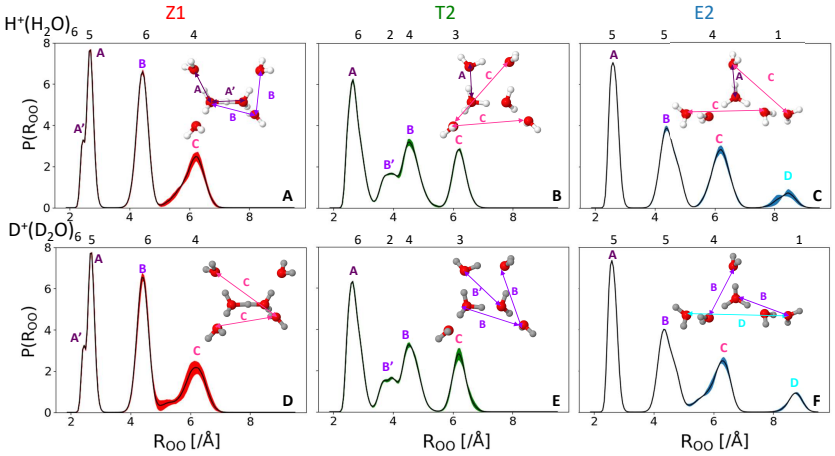


Figure 4.5: Probability amplitude projected onto the 15 OO distances ($P(R_{OO})$), of $H^+(H_2O)_6$ (A, B, and C) and $D^+(D_2O)_6$ (D, E, and F). The OO distance distributions were obtained from three independent simulations of each isomer, Z1 (A and D), T2 (B and E), and E2 (C and F). We obtained 49 wave functions from each simulation. The solid line in each plot is the average of these distributions over the three independent simulations, while the colored shading around the line represents the standard deviation from the three simulations. The numbers at the top of each plot provide the integrated probability amplitude over the peak(s) below the number, rounded to the nearest whole number. Representative OO distances that contribute to each peak are identified in the insets. The average OO distances and centers of the peaks of these distributions are provided in Tables 4.4, 4.5, and 4.6.

when anharmonic zero-point motions are included, the average values of these angles are reduced and the widths of these distributions are generally larger than the average value of the angle (see Tables 4.7 - 4.9). The additional constraints on the structure of T2 due to the four-membered ring result in the larger decrease in the ZPE of T2 relative to E2 and Z1 upon deuteration.

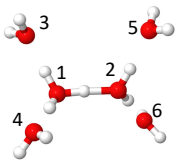


Table 4.4: Average OO Distances (in Å) for the Z1 Isomer of $H^+(H_2O)_6$ (Shown Above).

Atoms[a]	Equilibrium	$H^+(H_2O)_6$[b]	$D^+(D_2O)_6$[b]
O_1O_2	2.38	2.42	2.41
O_1O_3	2.65	2.69	2.69
O_1O_4	2.66	2.68	2.68
O_1O_5	4.24	4.34	4.32
O_1O_6	4.29	4.48	4.45
O_2O_3	4.29	4.48	4.45
O_2O_4	4.24	4.40	4.34
O_2O_5	2.66	2.68	2.68
O_2O_6	2.65	2.69	2.69
O_3O_4	4.41	4.40	4.44
O_3O_5	6.15	6.19	6.25
O_3O_6	6.20	5.68	5.45
O_4O_5	4.66	6.22	6.13
O_4O_6	6.15	6.31	6.29
O_5O_6	4.41	4.44	4.42
Peak[c]		$H^+(H_2O)_6$	$D^+(D_2O)_6$
A′		2.39 (0.8)	2.39 (0.8)
A		2.68 (4.2)	2.68 (4.2)
B		4.42 (6.0)	4.41 (6.1)
C		6.10 (4.0)	6.06 (3.9)

[a] The oxygen atoms involved in the reported OO distance.

[b] Reported averages are based on three independent simulations with 49 wave functions from each simulation.

[c] Positions of the peaks (in Å) in the OO distance distributions plotted in panels A and D of Figure 4.5. The numbers in parentheses provide the integrated area of the peak.

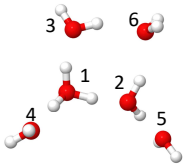


Table 4.5: Average OO Distances (in Å) for the T2 Isomer of $H^+(H_2O)_6$ (Shown Above).

Atoms[a]	Equilibrium	$H^+(H_2O)_6$[b]	$D^+(D_2O)_6$[b]
O_1O_2	2.47	2.46	2.47
O_1O_3	2.57	2.61	2.60
O_1O_4	2.60	2.64	2.63
O_1O_5	4.69	4.00	4.00
O_1O_6	3.93	4.66	4.71
O_2O_3	3.63	3.69	3.66
O_2O_4	4.27	4.41	4.40
O_2O_5	2.69	2.91	2.90
O_2O_6	2.87	2.72	2.72
O_3O_4	4.42	4.49	4.49
O_3O_5	6.10	2.97	2.93
O_3O_6	2.82	6.17	6.16
O_4O_5	6.29	6.14	6.17
O_4O_6	6.03	6.33	6.36
O_5O_6	4.74	4.90	4.84
Peak[c]		$H^+(H_2O)_6$	$D^+(D_2O)_6$
A		2.71 (6.0)	2.71 (6.0)
B		3.80 (1.8)	3.78 (1.8)
B′		4.59 (4.2)	4.59 (4.2)
C		6.21 (3.0)	6.23 (3.0)

[a] The oxygen atoms involved in the reported OO distance.

[b] Reported averages are based on three independent simulations with 49 wave functions from each simulation.

[c] Positions of the peaks (in Å) in the OO distance distributions plotted in panels B and E of Figure 4.5. The numbers in parentheses provide the integrated area of the peak.

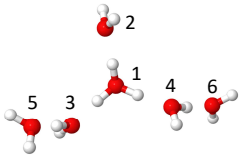


Table 4.6: Average OO Distances (in Å) for the E2 Isomer of $H^+(H_2O)_6$ (Shown Above).

Atoms[a]	Equilibrium	$H^+(H_2O)_6$[b]	$D^+(D_2O)_6$[b]
O_1O_2	2.61	2.64	2.64
O_1O_3	2.51	2.52	2.52
O_1O_4	2.50	2.51	2.51
O_1O_5	4.76	4.80	4.75
O_1O_6	4.49	4.59	4.58
O_2O_3	4.30	4.39	4.35
O_2O_4	4.22	4.33	4.30
O_2O_5	5.62	6.03	5.61
O_2O_6	6.16	6.35	6.20
O_3O_4	4.16	4.25	4.23
O_3O_5	2.68	2.72	2.72
O_3O_6	6.27	6.26	6.40
O_4O_5	6.06	6.10	6.27
O_4O_6	2.67	2.72	2.71
O_5O_6	8.50	8.38	8.73
Peak[c]		$H^+(H_2O)_6$	$D^+(D_2O)_6$
A		2.62 (5.0)	2.62 (5.0)
B		4.47 (5.0)	4.45 (5.1)
C		6.19 (4.0)	6.15 (3.9)
D		8.38 (1.0)	8.73 (1.0)

[a] The oxygen atoms involved in the reported OO distance.

[b] Reported averages are based on three independent simulations with 49 wave functions from each simulation.

[c] Positions of the peaks (in Å) in the OO distance distributions plotted in panels C and F of Figure 4.5. The numbers in parentheses provide the integrated area of the peak.

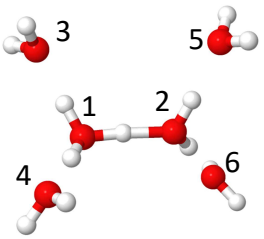


Table 4.7: Average Value of the HOO Angles (in Degrees) for the Z1 Isomer of $H^+(H_2O)_6$ (Shown Above) and the Width of the Ground State Probability Distribution When Projected onto the HOO Angle is Provided in Parentheses.

Atoms[a]	Equilibrium	$H^+(H_2O)_6$[b]	$D^+(D_2O)_6$[b]
O_1O_2	1.09	0.61 (6.07)	0.64 (5.34)
O_1O_3	2.56	0.51 (9.58)	0.76 (8.45)
O_1O_4	3.23	0.61 (9.55)	0.85 (8.42)
O_2O_5	−3.22	−0.32 (9.47)	−0.53 (8.36)
O_2O_6	−2.57	−0.52 (9.56)	−0.65 (8.41)

[a] The oxygen atoms involved in the reported HOO or DOO angle.

[b] Reported averages are based on three independent simulations with 49 wave functions from each simulation.

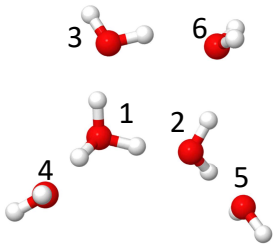


Table 4.8: Average Value of the HOO Angles (in Degrees) for the T2 Isomer of $H^+(H_2O)_6$ (Shown Above) and the Width of the Ground State Probability Distribution When Projected onto the HOO Angle is Provided in Parentheses.

Atoms[a]	Equilibrium	$H^+(H_2O)_6$[b]	$D^+(D_2O)_6$[b]
O_1O_2	5.76	1.52 (7.64)	1.6 (7.22)
O_1O_3	8.95	3.09 (10.71)	3.06 (9.93)
O_1O_4	−2.60	−0.55 (8.93)	−0.92 (7.87)
O_2O_5	−3.74	0.10 (9.96)	−0.66 (8.92)
O_2O_6	13.36	10.60 (11.88)	10.42 (10.69)
O_3O_6	15.51	15.27 (14.03)	14.73 (12.09)

[a] The oxygen atoms involved in the reported HOO or DOO angle.

[b] Reported averages are based on three independent simulations with 49 wave functions from each simulation.

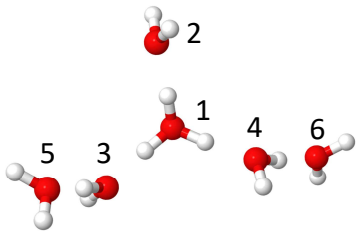


Table 4.9: Average Value of the HOO Angles (in Degrees) for the E2 Isomer of $H^+(H_2O)_6$ (Shown Above) and the Width of the Ground State Probability Distribution When Projected onto the HOO Angle is Provided in Parentheses.

Atoms[a]	Equilibrium	$H^+(H_2O)_6$[b]	$D^+(D_2O)_6$[b]
O_1O_2	−2.11	−0.73 (8.97)	−0.71 (7.84)
O_1O_3	−3.26	−1.12 (7.84)	−0.65 (6.89)
O_1O_4	−2.52	−0.62 (7.68)	−0.77 (6.81)
O_3O_5	−2.05	−0.82 (10.44)	−0.18 (9.09)
O_4O_6	1.94	0.24 (10.20)	0.42 (8.93)

[a] The oxygen atoms involved in the reported HOO or DOO angle.

[b] Reported averages are based on three independent simulations with 49 wave functions from each simulation.

On the basis of our best estimates of the relative energies of the Z1, T2 and E2 isomers of $H^+(H_2O)_6$, the E2 and T2 isomers are higher in energy than Z1 by 0.46 and 0.33 kcal/mol, respectively. For $D^+(D_2O)_6$ the analogous calculations predict the Z1, E2 and T2 isomers to be roughly equal in energy. On the basis of these results, one would expect only the Z1 isomer to be important at low (e.g., $T = 40$ K) temperatures for $H^+(H_2O)_6$, while for $D^+(D_2O)_6$, it is difficult to anticipate the relative contributions of these isomers to the spectrum. Indeed, double resonance studies of $H^+(H_2O)_6$, tagged with a hydrogen molecule by Heine *et al.*[12]

showed evidence that the Zundel form of $H^+(H_2O)_6$ contributes to the spectrum. This was concluded on the basis of the strong transition near 1000 cm^{-1}, close to the frequency of the 1-0 transition of the shared proton in Zundel form of $H^+(H_2O)_2$.[161] Further analysis of this spectrum also showed evidence of contributions of the Eigen form of $H^+(H_2O)_6$.

To further understand the contributions of the three isomers to the spectra, we performed vibrational perturbation theory (VPT2) calculations[156–158,162,163] for the three stable structures shown in Figure 4.1 at the MP2/aug-cc-pVDZ and MP2/aug-cc-pVTZ(-d) levels of theory, focusing on the OH stretching and HOH bending vibrations (see the Supporting Information for additional details of these calculations). The VPT2 spectra obtained at the MP2/aug-cc-pVTZ(-d) level of theory are plotted in Figure 4.6, where they are compared to the observed spectra. The strong couplings between vibrational motions in these protonated water clusters that result from the strong ionic hydrogen bonds means that one needs to be careful both in performing and interpreting the results of VPT calculations for these ions. For this reason, in the calculations presented in Figure 4.6 we only consider states with excitation in the OH stretching or HOH bending vibrations. While this makes the calculations more reliable, states with excitation in low-frequency vibrations are not considered in this analysis, and such transitions have been shown to contribute to this spectral region in similar systems.[7,9,164] As it is likely that our VPT2 analysis will not capture all of the transitions that contribute to the spectrum, in determining which isomers may be contributing to the spectra of $H^+(H_2O)_6$ and $D^+(D_2O)_6$ we focus on transitions that have large intensities in the calculated spectrum in regions where there are no peaks in the corresponding experimental spectrum.

As seen in Figure 4.6, the calculated spectrum for the Z1 isomer of $H^+(H_2O)_6$ contains peaks that are consistent with most of the features in the corresponding experimental spectrum. Since the calculated spectrum focuses on the 1-0 transitions, it is not able to capture the breadth of many of the features. Likewise, the calculated spectrum for the E2 isomer

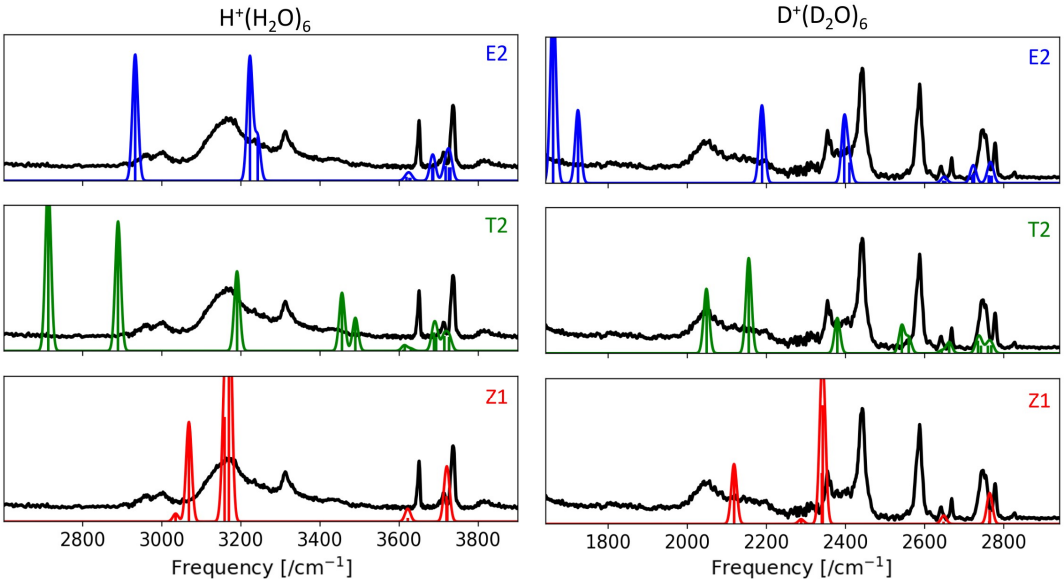


Figure 4.6: (Left) Comparison of the calculated $H^+(H_2O)_6$ spectra of the E2, T2, and Z1 (top to bottom) isomers to the experimental spectrum for H_2 tagged $H^+(H_2O)_6$ [6] previously shown in Figure 4.2. The calculated peaks of the Z1 isomer at 3159 cm^{-1} and 3068 cm^{-1} are shifted from their VPT2 frequencies. (Right) Comparison of $D^+(D_2O)_6$ spectra of the E2, T2, and Z1 (top to bottom) isomers to the experimental spectrum for D_2 tagged $D^+(D_2O)_6$,[7] previously shown in Figure 4.2. The calculated peak of the Z1 isomer at 2339 cm^{-1} is shifted from its VPT2 frequency. All VPT2 calculations are performed at the MP2/aug-cc-pVTZ(-d) level of theory and basis set.

has intensity in spectral regions in which there is intensity in the measured spectrum. This observation is consistent with the double resonance studies of Heine *et al.*,[12] in which they identified two isomers of $H^+(H_2O)_6$ that contributed to the spectrum and characterized them as Eigen and Zundel. In contrast, the calculated spectrum for the T2 isomer has strong transitions near 2700 and 3500 cm^{-1}, where there is no signal in the experimental spectrum. This is consistent with the T2 isomer being higher in energy compared to the E2 and Z1 isomers when ZPE is considered.

The corresponding spectrum of $D^+(D_2O)_6$ tagged with a D_2 molecule has a very different spectral envelope compared to the corresponding spectrum for $H^+(H_2O)_6$. This contrasts with the smaller protonated water clusters, for which the main spectral features of the deuterated spectrum map onto the spectrum for the all-H species, albeit shifted by roughly a factor of $\sqrt{2}$ in frequency.[7,9] If we make a similar comparison of the reported D_2 tagged spectrum of $D^+(D_2O)_6$ to the spectra of the three isomers calculated using VPT2, we find that the spectrum that shows the best agreement is the T2 isomer, with possible contributions from the Z1 isomer. Now it is the E2 isomer that has calculated peaks that are not observed (e.g., near 1700 cm^{-1}). Again, this observation is consistent with the energetic ordering of the three isomers once zero-point energy is considered, specifically the lowering of the energy of the T2 isomer relative to E2.

4.4 Conclusion

In conclusion, we have explored the isomerization pathway between the Zundel (Z1) and Eigen (E2) minima on the potential for $H^+(H_2O)_6$ and $D^+(D_2O)_6$. We identified a stable intermediate between the Z1 and E2 structures that contains a four-membered ring (T2). Based on just the electronic energies, the T2 structure is lower in energy than either the Z1 or E2 isomers. Once zero-point energy is considered, the Z1 structure becomes the most stable in $H^+(H_2O)_6$, while for $D^+(D_2O)_6$ the three structures differ in energy by less than 0.05 kcal/mol based on CCSD(T)/CBS+ZPE calculations. While the energy differences between these structures are much smaller than their zero-point energies, the fully anharmonic calculations

of the ground state wave function using DMC indicate that at low temperatures the wave function for each isomer is localized in a single minimum on the potential energy surface. This is consistent with the findings of double-resonance experiments by Heine *et al.*[12] and by Khuu *et al.*[145] Finally, the changes in the energy landscape with deuteration provide an explanation for the large differences in the spectra of H_2-tagged $H^+(H_2O)_6$ and D_2-tagged $D^+(D_2O)_6$.

Chapter 5

CORRELATIONS BETWEEN STRUCTURE AND SPECTRA OF PROTONATED WATER CLUSTERS

5.1 Introduction

Assigning vibrational spectra of medium-sized clusters that are held together by hydrogen bonds has proven to provide a long-standing challenge for theoretical and computational approaches. Early studies focused on using scaled harmonic calculations to provide assignments.[165,166] The challenge in applying this approach to systems that include strong ionic hydrogen bonds comes from the large variability of the anharmonicities of the OH bonds depending on the strength of the hydrogen-bonding interaction. Second order vibrational perturbation theory has evolved into a method of choice for including anharmonicity in semi-rigid systems. While this approach is effective in reproducing the anharmonic frequencies and intensities for vibrations that are well-described as perturbed harmonic oscillators, the approximation has been demonstrated to break down in hydrogen-bonded systems. This is particularly notable for ionic hydrogen bonds between a hydronium ion and water molecules where the large anharmonicity of the OH stretching vibrations and strong couplings of these vibrations to of low-frequency vibrations that affect the strength of the hydrogen bonds can result in energies of vibrationally excited states being calculated to be lower than the calculated zero-point energy.[40,167] This problem can be avoided by performing large-scale variational calculations, but the large density of vibrational states can make such calculations very challenging. Even in cases where such calculations can be performed either using MCTDH or VSCF/VCI approaches,[8,40,168–170] the assignment of the spectra remains challenging.

An attractive alternative is to use reduced dimensional approaches to explore the spec-

trum.[122,171] While studies of this type may be able to approximately capture the position of the band origins for the OH stretching vibrations, they will fail to fully capture the couplings between the OH stretching vibrations and the lower-frequency vibrations. There will be similar challenges for the NEO-based approaches, which have been applied to the protonated water tetramer by Fetherolf *et al.*[172]

In this chapter, we apply a mapping approach to the assignment of the measured spectra for protonated water clusters. The focus will be on the protonated water hexamer, where observed changes in the spectral envelope upon deuteration indicate that different combinations of isomers are contributing to the $H^+(H_2O)_6$ and the $D^+(D_2O)_6$ spectra. The method is calibrated for smaller clusters, where the assignment of the spectrum is more straightforward, and the results are compared to those obtained using scaled harmonic and VPT2 approaches.

5.2 Methods

5.2.1 Diffusion Monte Carlo

The approach, described below requires accurate ground state probability amplitudes, which can be used to obtained vibrationally averaged values for structural parameters, for example OH distances. These will be obtained using diffusion Monte Carlo. For the protonated water clusters with three or four water molecules, e.g. **3H**, **3D**, **4H-E**, and **4D-E** we use the probability amplitudes obtained as part of a previous study,[97] while for clusters containing five and six water molecules, we follow the protocol used in our recent study of the **6H** clusters.[13] In the above and the discussion that follows, we will denote the protonated water cluster using n**H-isomer label** and n**D-isomer label** for a particular isomer of $H^+(H_2O)_n$ and $D^+(D_2O)_n$, respectively. In this notation, **E** will be used to denote an Eigen-form of protonated water clusters, in which a single hydronium molecule is solvated, while **Z** denotes the Zundel-form in which the excess proton is shared between two water molecules.[173,174] Finally **T** is used to denote a four-member ring structure.

The details of DMC and our implementation of guided DMC for water clusters can be found elsewhere.[13,22,24,25,27,97,113,114,155] In this work, we use the potential energy surface that was developed by Yu and Bowman.[8] At the end of the simulation, we obtain a distribution of structures of the ion of interest, which provides a Monte Carlo sampling of the ground state. We refer to the members of this distribution as *walkers*. The calculation of expectation values requires a way to evaluate the value of the wave function at each of the Monte Carlo sampling points. Suhm *et al.* and Barnett *et al.* have shown that this can be achieved using descendent weighting,[24,114]in which the number of descendants of each walker are tabulated over a specified simulation time. The number of descendents for the ith walker will be represented by w_i. Using these weights, the expectation values of atom-atom distances are obtained by evaluating

$$\langle r_{ij} \rangle = \sum_{k=1}^{N} r_{ij}(\mathbf{x}_k) w_k \tag{5.1}$$

where the summation is over the individual walkers, and $\mathbf{x}_i$ provides the coordinates of the kth walker. Additional details about the implementation of DMC for the present study are provided in the Supporting Information.

5.2.2 Electronic Structure Calculations

In developing the mappings and generate spectra based on scaled harmonic and vibrational perturbation theory (VPT2) calculations, we optimized the structures and evaluated the harmonic frequencies and VPT2 frequencies and intensities at the MP2/aug-cc-pVTZ(-d) level of theory and basis. Due to the size of the calculations of the larger clusters, the $d-$orbitals were removed from the basis set for the hydrogen atoms. For the scaled harmonic calculations, scaling factors of 0.9531 and 0.9696 were used for the ions with hydrogen and deuterium atoms, respectively. These values were chosen to match the positions of the calculated and measured peaks associated with the free OH stretching vibration in the E2 isomer of **6H** and **6D** at 3735 cm^{-1} and 2750 cm^{-1}, respectively. A similar approach has

been followed in previous studies of protonated water clusters.[175]

To obtain the VPT2 spectra, calculations were preformed in the subspace that included the OH stretching and HOH bending vibrations as the inclusion of the low-frequency vibrations can cause the expansions used in perturbation theory to diverge. The intensities are more sensitive to this behavior than the frequencies. As in perturbation theory, the corrections to the intensities of 1-0 transitions come in at second order, we have elected to use the harmonic intensities in plots of the spectra. All of these calculations were preformed using the Gaussian 16[154] program package.

5.3 Results and Discussion

5.3.1 Evaluation of Relationships Between OH Bond Lengths and Transition Frequencies and Intensities

We start by considering the correlations between the frequencies and intensities of OH stretching transitions and the calculated OH bond lengths. As noted above, previous work of Xantheas and co-workers demonstrated a linear relationship between the change in the OH bond length and the vibrational frequency of the OH stretching vibration from their values for an isolated water molecule, evaluated at the same level of electronic structure theory.[52,53] This work also showed that $d\omega_{OH}/dr_{e,\mathrm{OH}}$ is independent of the level of electronic structure theory used to evaluate the equilibrium structure and the harmonic frequency. Recently they extended this exploration to hydronium, and showed that the value of $d\omega_{OH}/dr_{e,\mathrm{OH}}$ is the same as that found for water.[55] That investigation explored a variety of levels of electronic structure theory and basis sets as well as the potential of Yu and Bowman used in the present study.[8] Additionally, the work of Boyer, *et al.*[4] showed that for hydrogen bonded OH stretches in water clusters $d\nu_{\mathrm{OH}}/dr_{0,\mathrm{OH}} = d\omega_{\mathrm{OH}}/dr_{e,\mathrm{OH}}$, indicating that the anharmonic frequency and vibrationally averaged OH bond length follow the same trends as the harmonic frequency and the equilibrium OH bond length.

Based on the above observations, we calculated the shifts of the harmonic frequencies of

the OH and OD oscillators and the corresponding equilibrium bond lengths for hydrogen-bound OH bonds in water and hydronium for a variety of water and protonated water clusters with up to six water molecules. The chosen clusters are illustrated in Figures 5.2 and 5.3As noted above, these calculations were performed at the MP2/aug-cc-pVTZ(-d) level of theory and basis, with the $d-$functions removed from the basis used to describe the hydrogen and deuterium atoms. Based on earlier studies, we expect that the value of $d\omega_{\mathrm{OH}}/dr_{e,\mathrm{OH}}$ obtained at this level of theory can be used to obtain $\Delta\nu_{\mathrm{OH}}$ from $\Delta r_{0,\mathrm{OH}}$ values obtained from DMC calculations that were preformed using the potential of Yu and Bowman,[8] which was based on energies calculated at the CCSD(T)-F12/aug-cc-pVDZ level of theory and basis.

One challenge in this approach comes from the fact that the OH stretching vibrations are often collective in nature. In other words the normal modes involve displacements of several OH oscillators. On the other hand, these correlations, which are based on Badger's rule,focus on isolated OH oscillators. As such, rather than using calculations based on clusters in which all of the hydrogen atoms have the same mass number, for calculations of the OH stretch frequency, we replace all but one of the hydrogen atoms with deuterium. Likewise to determine the frequency of the local OD stretching vibrations, the masses of all the hydrogen atoms except the one of interest are multiplied by four. This differentiation in the masses of the hydrogen or deuterium atom of interest from the remaining hydrogen atoms provides an effective way to generate modes that are localized on a single OH stretching vibration. A similar approach was taken by Sibert and co-workers in their development of localized normal modes, and we use their approach to obtain quadratic coupling terms among the OH oscillators, which are used later in this study.[122,176,177]

Another challenge is found when the OH bond length in hydronium is extended beyond 0.1 Å, and, as in earlier studies, we only consider cases where the displacement of the OH bond length in hydronium is smaller than 0.1 Å. This restriction was introduced because longer displacements often correspond to Zundel-like structures, where the excess proton becomes equidistant from the donor and acceptor water molecule. The correlation is expected to break down in these cases due to the change in the zero-order description of the OH vibrational

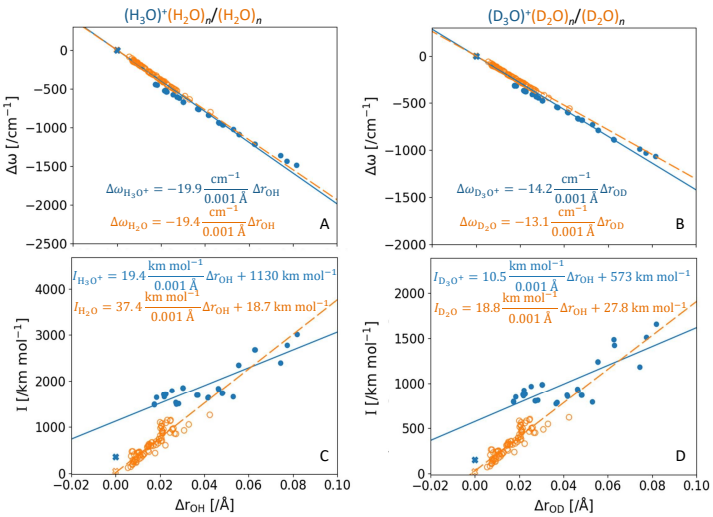


Figure 5.1: Plots showing the relationships between $\Delta\omega$ (A and B) and the harmonic intensity (B and D) and Δr for the hydrogen bonded OH and OD stretches in isotopologues of $(H_2O)_n$, n=2-6, and $H^+(H_2O)_n$, n=3-6 shown in Figures 5.2 and 5.3, excluding OH bonds that are extended by more than 0.10 Å as noted in the text. The values of Δr and $\Delta\omega$ are obtained by subtracting the corresponding values for an OH bond in an isolated water molecule (orange) or hydronium ion (blue). The data is fit to the equations provided in the same color in the insets. The solid blue and open orange X in each panel represent the value of $\Delta\omega$ (A and B) and Intensity (C and D) for the hydronium molecule (blue) and the isolated water molecule (orange).

motions in these structures. This restriction will affect the description of the shared proton in the **6H-Z** structure as well as the OH bond labeled 1 in Figure 5.6 for **5H-E** and **6H-T2**.

Finally, while the analysis described above allows us to determine the frequencies and intensities of transitions involving OH bonds that are in hydrogen-bonding environments, it does not provide insights into the assignment of transitions involving the free OH oscillators. In most cases, these frequencies are close to those of the OH frequencies in an isolated water molecule or hydronium ion. To model the spectrum in this region, we assume that the vibrational frequencies of the free OH oscillators are the average of the symmetric and antisymmetric stretches of the isolated water molecule or hydronium ion, and reintroduce the quadratic couplings between these oscillators that were obtained when we evaluated the localized normal modes. For this spectral region, all transitions are given equal intensity, and the value of the intensity was chosen to allow comparison with the experimental spectrum. As such, the relative intensities of the transitions associated with the free OH oscillators are not expected to match the observed pattern in the measured spectra.

The results of this analysis are provided in Figure 5.1. Open orange circles are used to report results for water, while closed blue circles show the results for hydronium. In panels A and B, we report the change in the harmonic frequency of the isolated OH oscillators from their values in an isolated water molecule or hydronim ion, plotted against the shift in the corresponding value of the OH bond length in the optimized structure of the cluster. These data are fit to a line that goes through the origin, and the equation for the fit line is reported in the inset. As expected from previous studies,[4,55] the slope for both hydronium and water, plotted in Figure 5.1A are close to -19 cm^{-1}/(0.001 Å), which is consistent with earlier studies. Figure 5.1B provides the corresponding information for the OD stretching vibrations. As with the OH stretching vibrations, the slopes obtained for D_2O and D_3O^+ are nearly identical. Additionally, they are roughly $1/\sqrt{2}$ times the values obtained for the OH stretches. This is consistent with the decrease in the OD harmonic frequency (compared to OH) by a factor of $\sqrt{2}$ while, based on the Born-Oppenheimer approximation, the equilibrium bond lengths are independent of mass. Anharmonic calculations yield similar slopes, and

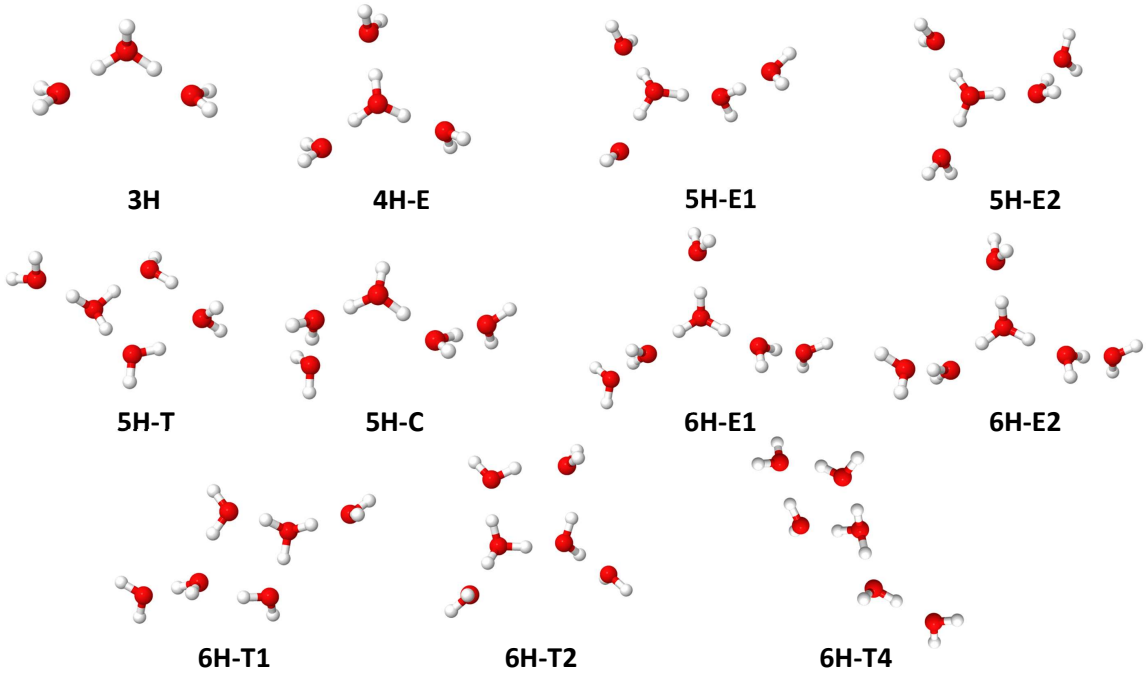


Figure 5.2: Structures of the local minima of the protonated water clusters that are used in Figure 5.1. Structures are optimized at the MP2/aug-cc-pVTZ(-d) level of theory and basis set.

the resulting plots are provided in 5.4.

In the remaining two panels of Figure 5.1, we plot the calculated harmonic intensities, obtained from the calculations on the partially deuterated clusters as functions of the displacement of the OH bond lengths relative to an isolated water molecule or hydronium ion. In agreement with previous studies[62,63] the intensity of the OH stretching vibrations in water increase roughly linearly with the hydrogen bond length. Since the OH bond lengths also show a linear dependence on the hydrogen bond length, the results plotted in orange and blue in panels C and D of Figure 5.1 can be fit to a line, and the parameters are provided in the plot. While this data was evaluated at the MP2/aug-cc-pVTZ(-d) level, we expect that the resulting correlations will provide at least a qualitative description of the intensities

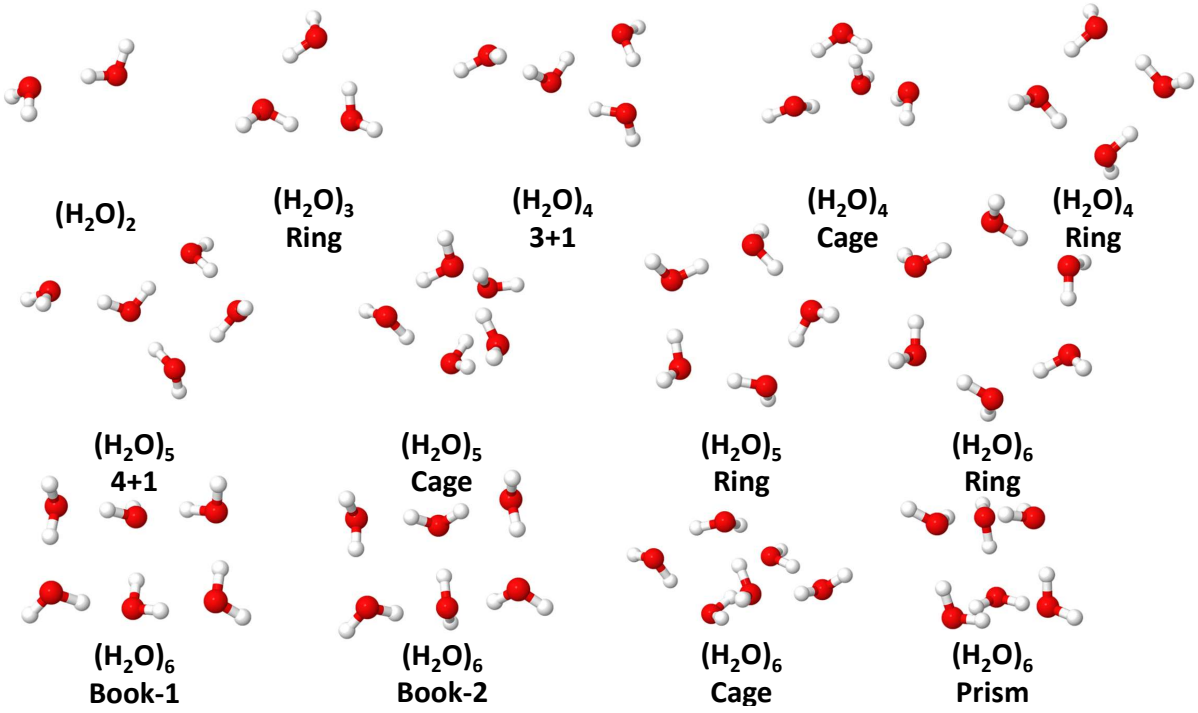


Figure 5.3: Structures of the local minima of the neutral water clusters that are used in Figure 5.1 . Structures are optimized at the MP2/aug-cc-pVTZ(-d) level of theory and basis set.

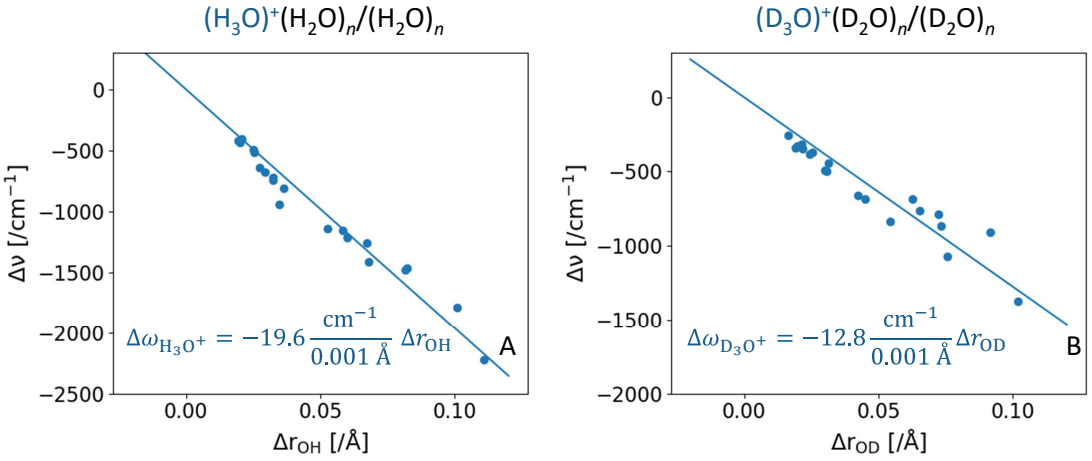


Figure 5.4: Plots showing the relationships between $\Delta\nu$ and Δr for the hydrogen bonded OH (A) and OD (B) stretches in isotopologues of $H^+(H_2O)_n$, n=3-6 shown in Figure 5.2. The values of Δr and $\Delta\nu$ are obtained by subtracting the corresponding values for an OH bond in a hydronium ion. The data is fit to the equations provided in the same color in the insets.

based on higher levels of theory.

5.3.2 Comparison of Spectra

With the above correlations in hand, we can use the calculated vibrationally averaged OH bond lengths to determine the expected frequencies and intensities of the OH transitions in water and protonated water clusters. The calculated values of $\langle r_{\mathrm{OH}} \rangle_0$, obtained from the DMC calculations, are provided in Tables 5.1-5.4. It should be noted that because the data used in the data collection for hydronium only considers displacements of the OH bond by up to 0.1 Å we cannot use this approach to predict frequencies of the shared proton stretch in conformers that contain larger displacements of the OH bond, for example **5H-E**, **6H-Z**, and **6H-T2**. These are labeled with orange numbers in the structures in Figures 5.5 and 5.6. Likewise, this model treats all of the OH stretching vibrations as localized oscillators (as opposed to collective normal modes). With the exception of **3H** and **4H**, where several OH oscillators are equivalent by symmetry, each of the hydrogen-bound OH oscillators feels a unique OH bonding environment, and the OH oscillators are expected to remain localized. Likewise, because the model focuses on hydrogen-bonding environments, it cannot describe the free OH oscillators, and this region is modeled by introducing bilinear couplings between the OH oscillators, using the frequencies obtained for OH oscillators in a water moleucle or hydronium ion. Finally, the model cannot describe overtones or combination transitions involving lower-frequency vibrations. This final limitation also affects the scaled harmonic and VPT2 spectra. With this limitation in mind, in the discussion that follows we will focus on agreement between calculated and measured peaks and eliminate structures from contributing to the spectra based on predicted transitions in regions where there is no intensity in the measured spectrum.

Table 5.1: Frequencies (in cm^{-1}), Intensities (in km/mol), and Expectations Values of r_{OH} (in Å) of the Hydronium Core from the DMC Wave Functions of the Isomers Used to Calculate Spectra in Figures 5.5 and 5.6.

System	$\langle r_{OH} \rangle_0$ [a]	ν [b]	Intensity [c]
3H	1.060	2191.64	1261.92
	1.060	2191.64	1261.92
4H-E	1.028	2828.44	1199.84
	1.028	2828.44	1199.84
	1.029	2808.54	1201.78
5H-E	1.015	3087.14	1174.62
	1.015	3087.14	1174.62
	1.096	1475.24	1331.76
5H-T	1.019	3007.54	1182.38
	1.038	2629.44	1219.24
	1.039	2609.54	1221.18
6H-E	1.012	3146.84	1168.80
	1.052	2350.84	1246.40
	1.056	2271.24	1254.16
6H-T2	1.108	1236.44	1355.04
	1.022	2947.84	1188.20
	1.011	3166.74	1166.86
6H-T4	1.056	2271.24	1254.16
	1.026	2868.24	1195.96
	1.024	2908.04	1192.08

[a] Standard deviations for these values are < 0.001

[b] Calculated from the blue correlation in Figure 5.1A.

[c] Calculated from the blue correlation in Figure 5.1C.

Table 5.2: Frequencies (in cm^{-1}), Intensities (in km/mol), and Expectations Values of r_{OD} (in Å) of the Hydronium Core from the DMC Wave Functions of the Isomers Used to Calculate Spectra in Figures 5.5 and 5.6

System	$\langle r_{OD} \rangle_0^a$	ν^b	Intensity^c
3D	1.051	1697.79	639.15
	1.051	1697.79	639.15
4D-E	1.021	2123.79	607.65
	1.021	2123.79	607.65
	1.022	2109.59	608.70
5D-E	1.014	2223.19	600.30
	1.013	2237.39	599.25
	1.091	1129.79	681.15
5D-T	1.014	2223.19	600.30
	1.030	1995.99	617.10
	1.032	1967.59	619.20
6D-E	1.008	2308.39	594.00
	1.043	1811.39	630.75
	1.047	1754.59	634.95
6D-T2	1.086	1200.79	675.90
	1.019	2152.19	605.55
	1.006	2336.79	591.90
6D-T4	1.049	1726.19	637.05
	1.021	2123.79	607.65
	1.020	2137.99	606.60

[a] Standard deviations for these values are < 0.001

[b] Calculated from the blue correlation in Figure 5.1B.

[c] Calculated from the blue correlation in Figure 5.1D.

Table 5.3: Frequencies (in cm^{-1}), Intensities (in km/mol), and Expectations Values of r_{OH} (in Å) of the Hydrogen-Bonded Water Molecules from the DMC Wave Functions of the Isomers Used to Calculate Spectra in Figures 5.5 and 5.6.

System	$\langle r_{OH} \rangle_0^a$	ν^b	Intensity^c
5H-E	1.003	3182.70	1029.58
5H-T	0.987	3493.10	430.54
	0.987	3493.10	430.54
6H-E	0.998	3279.70	842.38
	0.997	3299.10	804.94
6H-Z	1.002	3202.10	992.14
	1.003	3182.70	1029.58
	1.004	3163.30	1067.02
	1.002	3202.10	992.14
6H-T2	0.999	3260.30	879.82
	0.987	3493.10	430.54
	0.983	3570.70	280.78
6H-T4	1.000	3240.90	917.26
	0.990	3434.90	542.86
	0.987	3493.10	430.54

[a] Standard deviations for these values are < 0.001

[b] Calculated from the orange correlation in Figure 5.1A.

[c] Calculated from the orange correlation in Figure 5.1C.

Table 5.4: Frequencies (in cm^{-1}), Intensities (in km/mol), and Expectations Values of r_{OD} (in Å) of the Hydrogen-Bonded Water Molecules from the DMC Wave Functions of the Isomers Used to Calculate Spectra in Figures 5.5 and 5.6.

System	$\langle r_{OD} \rangle_0^a$	ν^b	Intensity^c
5D-E	1.001	2145.19	591.82
5D-T	0.982	2513.79	234.58
	0.982	2513.79	234.58
6D-E	0.992	2319.79	422.60
	0.991	2339.19	403.80
6D-Z	0.997	2222.79	516.61
	0.998	2203.39	535.41
	0.999	2183.99	554.22
	0.997	2222.79	516.61
6D-T2	0.992	2319.79	422.60
	0.982	2513.79	234.58
	0.979	2571.99	178.18
6D-T4	0.993	2300.39	441.40
	0.981	2533.19	215.78
	0.981	2533.19	215.78

[a] Standard deviations for these values are < 0.001

[b] Calculated from the orange correlation in Figure 5.1B.

[c] Calculated from the orange correlation in Figure 5.1D.

The results of this analysis for $n\mathbf{H}$, $n = 3 - 5$, is plotted with the blue traces in Figure 5.5. These results are compared to the results of scaled harmonic calculations (green trace) and the results of VPT2 calculations that focus on the OH stretching and HOH bending vibrations, as described above (red trace). The bottom trace is the previously reported measured spectrum.[9–11] The geometries of the isomers used in the construction of these spectra are shown at the top of the Figure.

The plot in panel A of Figure 5.5 shows the comparison of the spectra of **3H**. The arrow shows the peak in the experimental spectrum that has previously been assigned to the 1-0 transition of the antisymmetric stretching vibration of the shared protons in the hydronium

Table 5.5: Reference Values of the OH and OD Stretch Frequencies (in cm^{-1}) and Expectation Values of r_{OH} and r_{OD} (in Å), Used in the Spectral Mapping Method in Figures 5.5 and 5.6.

System	$\langle r_{OH} \rangle_{ref}$ [a,b]	ν_{ref}
H_2O [c]	0.976	3706.50
D_2O [c]	0.971	2727.19
H_3O^+ [d]	0.992	3544.84
D_3O^+ [d]	0.987	2592.39

[a] Calculated from the wave functions obtained from five independent guided DMC simulations with 5000 walkers, run for 20 000 time steps with a time step size of 1 a.u.
[b] Standard deviations for these values are < 0.001
[c] Based on the potential by Partridge and Schwenke.[115]
[d] Based on the potential by Huang, Carter, and Bowman.[120]

core.[51] The analysis that is based on the relationships between OH bond lengths and the frequencies and intensities provided in Figure 5.1 (blue), referred to as the spectral mapping method in the remainder of the discussion, produces a peak, peak 1-2, that is about 100 cm^{-1} higher than the transition in the experimental spectrum . Compared to the scaled harmonic calculation (green) the spectral mapping method prediction is significantly closer to the experimental transition. The calculated transition from VPT2 overestimates the anharmonicity and predicts a peak that is about 100 cm^{-1} lower than the transition in the experiment.

Panel B in Figure 5.5 shows the calculated and experimental specrtra of **4H-E**. The spectral mapping method (blue) predicts that the shared proton transition, peak 1-3, is 2848 cm^{-1}, which agrees well with the scaled harmonic calculation (green) but is still about 100 cm^{-1} higher than the transition in the experimental spectrum. The VPT2 spectrum overestimates the anharmonicty in this structure as well with the two intense features being

lower in frequency than the most intense feature in the experimental spectrum.

Panels C and D in Figure 5.5 compare the spectra of two low energy isomers of **5H**. In the study where the experimental spectrum of this species was obtained, DFT calculations showed that two isomers of the **5H-E** (referred to as the branched isomers in their study) and the **5H-T** isomer (referred to as the ring isomer in their study) differed by less than 1 kJ/mol in the harmonic zero-point corrected energies.[11] Table 5.6 shows the relative energies of these isomers based on calculations done at the MP2/aug-cc-pVTZ(-d) level of theory and basis and the trend of the energy ordering for **5H** is consistent with the previous study. The **5H-T** isomer is the minimum energy isomer but **5H-E** becomes lower with the addition of zero-point energy. Overall, the agreement between the calculated spectra of **5H-E** and the experimental spectrum in the lower left panel of Figure 5.5 is significantly better than the calculated spectra of **5H-T**. This is consistent with the results of the harmonic and VPT2 calculations of the previous study.[11] For the spectrum of **5H-E** in the bottom left panel of Figure 5.5 that is predicted based on the spectral mapping method (blue), the large peak at 3107 cm^{-1}, peak 2-3, is about 140 cm^{-1} higher than the experimental peak, higher than both the scaled harmonic (green) and VPT2 (red) calculations predict. There is one OH bond length in this isomer that is longer than the equilibrium OH bond length in hydronium by 0.1 Å. This is outside the range of the correlations plotted in Figure 5.1 and is therefore more akin to a shared proton stretch in a Zundel-like isomer, which is where this model starts to break down. The model predicts this peak to be 1495 cm^{-1}, which is about 400 cm^{-1} lower than the experimental transition at 1890 cm^{-1}. In the calculated spectrum of **5H-T** (lower right panel) using the spectral mapping method (blue), the splitting between the peaks at 3027 cm^{-1} and 2640 cm^{-1} is significantly larger than what is seen in the scaled harmonic (green) and VPT2 (red) calculations. These peaks represent the three OH stretches from the hydronium core and the spectral mapping method predicts a larger change in local environment between the one OH stretch that is hydrogen-bonding to the water off the ring, peak 3, to the two OH stretches that are hydrogen-bonding within the ring, peak 1-2. This method also predicts that that the OH stretch that is associated with the hydrogen-bond to

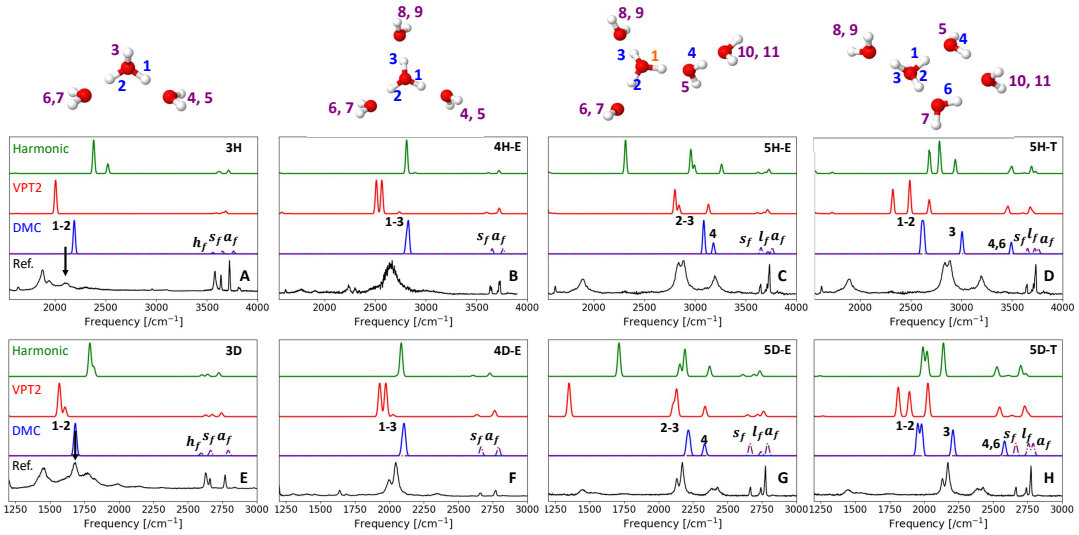


Figure 5.5: Comparisons of the calculated spectra of **3H** (A), **4H-E** (B), and two isomers of **5H**, **5H-E** (C), and **5H-T** (D). The spectra of the deuterated species are included in panels E-H. The spectra in blue are calculated with the $\langle r_{\mathrm{OH}} \rangle_0$ and $\langle r_{\mathrm{OD}} \rangle_0$ from the DMC wave functions of each isomer, the correlations from Figure 5.1, and the reference OH and OD distances and frequencies in Table 5.5. The data for these spectra are provided in Tables 5.1-5.4. The structures above the spectra indicate the specific OH stretch vibrations that are associated with the calculated frequencies shown in the spectrum (blue), as well as the free OH stretches (purple) where the dotted purple line indicates the use of the harmonic coupling between the OH and OD stretches from the harmonic calculations in the spectral mapping method. The orange OH stretch indicates one where Δr is greater than 0.1 Å and is therefore not shown in the range of the spectrum. The free OH region is labeled with h_f, for free hydronium OH/OD stretches, s_f, for the free OH/OD symmetric stretches, a_f, for the free OH/OD antisymmetric stretches, and l_f, for local free OH/OD stretches such as the one labeled 5 in **5H-E**. The other two calculated spectra are scaled harmonic (green) and calculated anharmonic (VPT2) frequencies with harmonic intensities (red). The black curve is the experimentally measured spectra.[9–11] All harmonic and VPT2 calculations are performed at the MP2/aug-cc-pVTZ(-d) level of theory and basis set.

the water off of the ring, peak 3, is higher in frequency than both the scaled harmonic and VPT2 calculations.

Table 5.6: Energies (in cm^{-1}) of **5H-E** and **5H-T**, Reported Relative to the Energy of **5H-E**.

Stationary Point	MP2/aug-cc-pVTZ(-d)	MP2/aug-cc-pVTZ(-d) + harmonic	MP2/aug-cc-pVTZ(-d) + harmonic (D)
5H-E	0.00	0.00	0.00
5H-T	−287.93	158.29	40.38

The deuterated isotopologues in panels E-H of Figure 5.5 show similar results to those shown in the spectra of $n\mathbf{H}$, $n = 3-5$. For **3D**, the spectral mapping method (blue) predicts a peak, peak 1-2, between the scaled harmonic (green) and VPT2 (red) calculations. However, this peak is closer in energy to the experimental transition, denoted by the arrow which was previously assigned to the 0-1 transition of the antisymmetric shared proton stretch from the hydronium core,[51] than the peak for the spectrum of **3H** in panel A of Figure 5.5. The spectra of **4D-E** in the panel F of Figure 5.5 follow the same trend that is seen in the spectra of **4H-E**. For the spectra of **5D-E** in panel G, the peak at 2230 cm^{-1}, peak 2-3, calculated using the spectral mapping method (blue), aligns closer to experiment than what was seen in the spectrum of **5H-E**. There is a 60 cm^{-1} shift in the frequency compared to experiment.

*5.3.3 Isomerization of **6H***

For the spectrum of $H^+(H_2O)_6$ the calculated spectra of the **6H-E** isomer (panel A) and **6H-Z** isomer (panel B) in Figure 5.6 show good agreement to experiment for all of the methods. This is expected as these two isomers have previously been observed by Heine et al. in low temperature double resonance experiments of this cluster.[12] For the spectrum of **6H-E** the spectral mapping method shows good agreement to experiment in the region between 3000 and 3500 cm^{-1}. The peaks at 2371 and 2291 cm^{-1} peaks 1 and 2, show a similar trend to that seen in $H^+(H_2O)_3$, where the peaks are closer to experiment than the scaled harmonic

spectrum, but not as low frequency as the one calculated in the VPT2 spectrum. However, in this case, the corresponding peak in the spectrum is at 1950 cm^{-1}, so the spectral mapping method is showing worse agreement than what was seen in the spectrum of **3H** in Figure 5.5. In the study by Heine et al. where the double resonance spectra are obtained they showed that spectra obtained from PBE+vdW AIMD simulations showed a similar shift of the calculated transition to the experimental transition of about 400 cm^{-1} in frequency.[12] VSCF/VCI calculations of the spectrum of the E2 isomer performed by Yu et al. show better agreement in this transition to experiment compared to the VPT2 calculations.[178] The calculated spectra of **6H-Z** in panel B using all three methods shows approximately the same level of agreement to experiment, with the scaled harmonic and spectral map method showing comparable spectra. For the spectral mapping method, the four OH stretches in **6H-Z** that are not involved in the central shared proton stretch are treated with the correlations that correspond to hydrogen-bonded water molecules as opposed to the correlations that correspond to the hydronium core since there is no definitive hydronium core in this isomer. The spectra of the other two isomers shown in Figure 5.6 have peaks around 3500 cm^{-1}, peak 6 for **6H-T2** and peaks 6 and 8 for **6H-T4**, where there is no intensity in the experimental spectrum. The calculated spectra of **6H-T4** in panel D of Figure 5.6 also have a doublet around 2900 cm^{-1}, peaks 2 and 3, with the spectral mapping method (blue) and the scaled harmonic calculation (green) and a doublet at 2700 cm^{-1} in the VPT2 spectrum (red) that does not correspond to a transition in the experimental spectrum.

There are two key features in the experimental spectrum **6D** that are not present in the spectrum of **6H**. There is one feature at 2050 cm^{-1} and another at 2588 cm^{-1}. These features are used as strong indicators for agreement to the experimental spectrum in the following discussion. The calculated spectrum of **6D** in Figure 5.6 shows that **6D-T2** in the plot in panel G and **6D-T4** in panel H has the best match to experiment. The spectral mapping method shows that the spectra of these two isomers reproduce the two key features present in the experimental spectrum, while the spectra of the other two isomers in Figure 5.6 do not show intensity in those transitions. The VPT2 spectrum of **6D-T2** in our previous study

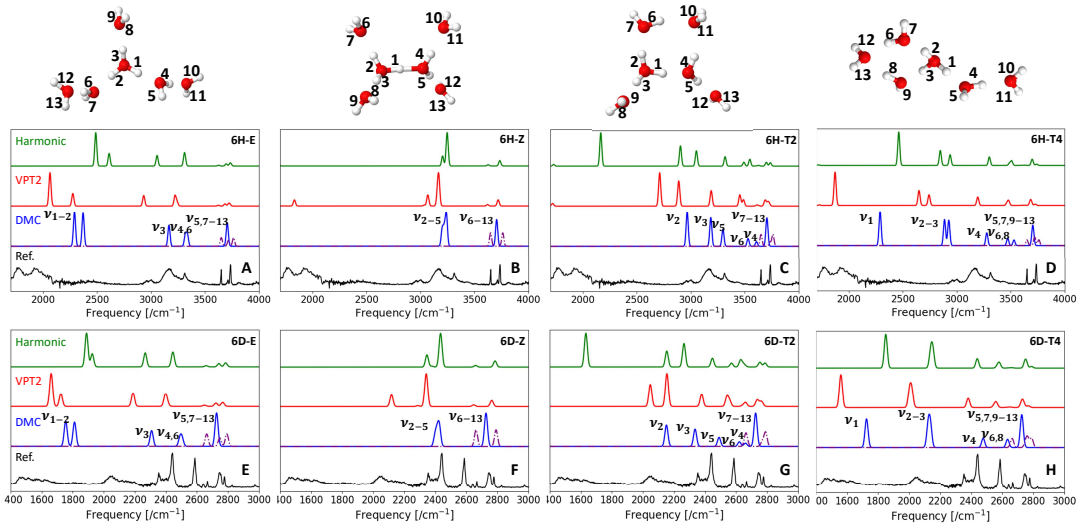


Figure 5.6: Comparisons of the calculated spectra of **6H-E** (A), **6H-T2** (B), **6H-Z** (C), and **6H-T4** (D). The spectra of the deuterated species are included in panels E-H. The spectra in blue are calculated with the $\langle r_{\mathrm{OH}} \rangle_0$ and $\langle r_{\mathrm{OD}} \rangle_0$ from the DMC wave functions of each isomer, the correlations from Figure 5.1, and the reference OH and OD distances and frequencies in Table 5.5. The data for these spectra are provided in Tables 5.1-5.4. The structures above the spectra indicate the specific OH stretch vibrations that are associated with the calculated frequencies shown in the spectrum (blue), as well as the free OH stretches (purple) where the dotted purple line indicates the use of the harmonic coupling between the OH and OD stretches from the harmonic calculations in the spectral mapping method. The orange OH stretches indicates ones where Δr is greater than 0.1 Å and is therefore not shown in the range of the spectrum. The free OH region is labeled with s_f, for the free OH/OD symmetric stretches, a_f, for the free OH/OD antisymmetric stretches, and l_f, for local free OH/OD stretches such as the one labeled 5 in **6H-T4**. The other two calculated spectra are scaled harmonic (green) and calculated anharmonic (VPT2) frequencies with harmonic intensities (red). The black curve is the experimentally measured spectra.[7,12] The calculated peaks of **6H-Z** at 3159 cm^{-1} and 3068 cm^{-1} as well as the calculated peak of **6D-Z** at 2339 cm^{-1} are shifted from their VPT2 frequencies in the same manner as our previous study.[13] All harmonic and VPT2 calculations are performed at the MP2/aug-cc-pVTZ(-d) level of theory and basis set.

showed good agreement to experiment when compared to **6D-E** and **6D-Z**.[13] In the plot depicting the spectrum of **6D-T4**, the spectral mapping method shows a consistent trend that has been observed for the other systems, where the lowest frequency feature, peak 1, is between the scaled harmonic transition and the VPT2 transition corresponding to the experimental feature around 1450 cm^{-1}. However, similar to the spectrum of **6H-E2** in panel A of Figure 5.6, the spectral mapping method is significantly higher in frequency than the experimental transition.

5.4 Conclusion

In conclusion, we have continued the exploration of the isomerization of **6H**, and how it changes upon deuteration. This work utilized a spectral map that utilizes the vibrationally averaged hydrogen-bonded OH distances obtained from diffusion Monte Carlo to predict the spectral features that are unique to each isomer of **6D** to determine the isomer(s) that contribute to the spectrum. The spectral map was created using simple relationships between the hydrogen-bonded OH distances in water clusters and the frequencies and intensities calculated from harmonic calculations of those clusters. We calibrated this method against scaled harmonic and VPT2 calculations for smaller clusters to show the limits of the method. For high frequency vibrations, the method showed little difference to the scaled harmonic calculations. However, for modes with large anharmonicity, the spectral mapping method showed closer agreement to experiment and in the case of **3H** and **3D**, the agreement was closer than VPT2. When applying the spectral map to four isomers of **6D** we showed that the isomers that display the closest match to experiment are **6D-T2** and **6D-T4**. **6D-T2** was shown in the previous chapter to be an intermediate in the isomerization pathway between **6D-Z** and **6D-E**, but **6D-T4** is another isomer that requires more investigation to determine its possible contributions to the spectrum of **6D**.

Chapter 6

ASSESSING THE COGNITIVE COMPLEXITY OF GENERAL CHEMISTRY EXAM QUESTIONS USING MARZANO'S TAXONOMY

6.1 Introduction

6.1.1 What is Marzano's Taxonomy?

Marzano and Kendall's New Taxonomy of Educational Objectives (henceforth referred to as Marzano's taxonomy or MT),[14] was developed as a replacement to Bloom's taxonomy. Marzano's taxonomy was developed as a theoretical model in order to predict specific phenomena related to a person's learning as opposed to a framework that describes learning. A key part of this theoretical model is that it is tied to the psychology principle that an action is independent of how complex a task is, but more about how practiced the task is to the person performing it.[179–181] Marzano gives the example of driving a car, which when broken down into steps is a complicated process, but can be done innately by people that have a lot of experience driving.[15] Marzano's taxonomy has two dimensions: the Mental Processing dimension describes the level of cognitive processing a student exhibits during a task seen in Figure 6.1, and the Domains of Knowledge dimension describes the type of knowledge the student is processing during the task.

Mental Processing Levels

Marzano describes three systems within the Mental Processing dimension in order of *decreasing* cognitive control exerted: the *Self System*, the *Metacognitive System*, and the *Cognitive System*. The Self System governs motivation and attention, and determines whether a student decides to engage in a task, and how much energy to bring to the task. This system

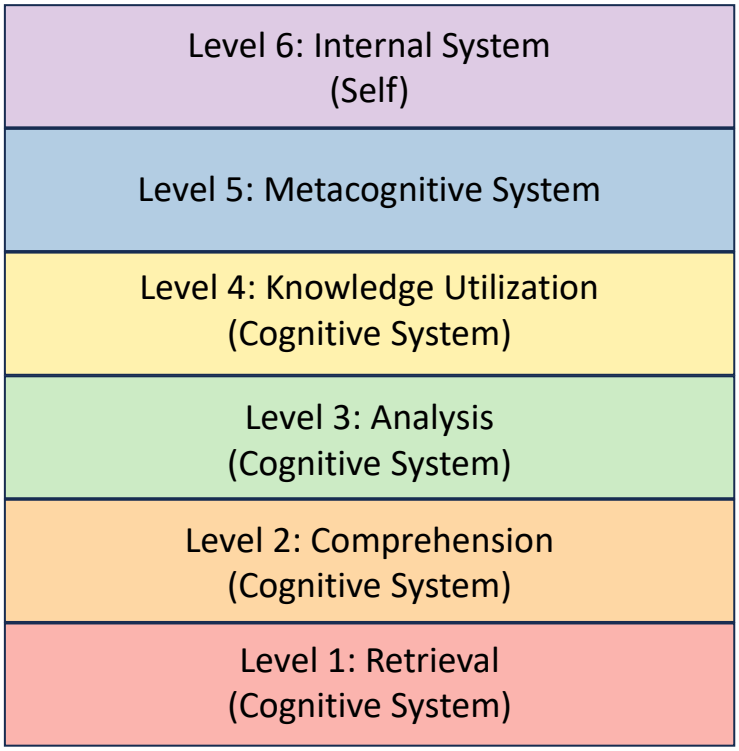


Figure 6.1: A representation of the Mental Processing dimension of Marzano's Taxonomy.[14,15]

includes all previous experience tied to a subject that influence the confidence that a student has when approaching a topic. Once engaged in a task, the Metacognitive System allows the student to set goals for the task, and to monitor their progress towards those goals. This governs how the student thinks about how they are learning and influences their study patterns. The Cognitive System then processes information related to the task, allowing the student to remember, classify, solve, compare/contrast, etc.

The Cognitive System itself contains four levels. In order of *increasing* cognitive control exerted, they are: *Retrieval*, *Comprehension*, *Analysis*, and *Knowledge Utilization*. In this work we use only *Retrieval*, *Comprehension*, *Analysis*, but we briefly describe all four levels below for completeness.

Retrieval: This level focuses on immediate recall or recognition of learned information, or execution of a learned algorithm. It is the lowest level of cognitive thinking, and has three sub-levels; Remembering, Recognizing, and Executing. This level focuses on accessing information that has been stored in a person's permanent memory in order to answer a question or solve a problem. Note that Marzano's Taxonomy places "execution of a learned algorithm" in this category, meaning that a student can perform a rote mathematical exercise without any deeper comprehension required of them. For the execution of an algorithm, this is the only process used in this work which is not included in the Information domain. Certainly, such a student may be able to explicate their deeper conceptual understanding of the algorithm, but since chemistry assessments largely tend to emphasize lower-level cognitive skills, students are rarely pushed to do so.[68,69,72] We don't actually know the full extent of what students know, based on our current assessment structure.

Comprehension: Comprehension involves inferring a new (to the student) relationship based on learned information. This level has two sub-levels; Integration and Symbolizing. It may be tempting to think of Integration as the integration of multiple topics in order to answer a question, but this level actually involves the integration of conceptual material into permanent memory from working memory, in which facts or concepts are actively being processed. This usually takes the form of explaining or symbolizing a learned concept in

a student's own words, such as creating a diagram that explains how a process works or explaining a trend in a different way than was taught.

Analysis: The Analysis level focuses on elaborating and extending comprehended information in a new (to the student) manner. Analysis has five sub-levels; Matching, Classifying, Analyzing Errors, Generalizing, and Specifying. This level is where Marzano describes the extension of knowledge to new questions or problems. This is where a person is taking what they have learned and applying it to new situations. The intent of this level is to draw new conclusions about the concepts that have previously been learned. The student is asked to elaborate and extend the information that was previously Integrated (moved from the working memory into the students permanent memory) in a new (to the student) manner. In this work, we only observed Analyzing Errors and Specifying.

Knowledge Utilization: Processes in this level focus on using comprehended information to accomplish a new (to the student) task. This level is made up of four sub-levels; Decision Making, Problem Solving, Experimenting, and Investigating. Knowledge Utilization is focused on using knowledge of a particular topic in order to perform a specific task or finish a project. These sub-levels focus on actions that are used when working on a larger project, similar to work done on a lab report, particularly one where the experimentation is guided by the student and not a provided step-by-step recipe. This also includes the literature review that is required to provide background information on certain projects and determining what information relates to that project and supports the claims within.

The 2D structure of MT is reminiscent of the revised Bloom's taxonomy (RBT),[2] but MT is distinct in important ways. One key differnce is the placement of the Metacognitive aspect in each taxonomy. In MT this is placed above each of the cognitive processes as shown in Figure 6.1, because this process informs the success of the levels below due to the goals that are formed in this level that inform how the cognitive levels will be pursued. In RBT, the Metacognitive aspect is in the same dimension as factual, conceptual, and procedural knowledge. This places it as an aspect of each cognitive level as opposed to a step before cognition can begin. Another key aspect is the placement of the Self system. In RBT this is

made a part of Metacognitive knowledge, indicating the influence it has on metacognition. However, in MT, Marzano uses arguments by Csikszentmihalyi about the separation of the self-system and the metacognitive system to place the Self system at the top of the taxonomy. This indicates the influence it has on all levels, including the hierarchy of how we create goals and pursue them.[14,182]

6.1.2 Using Marzano's Taxonomy as a Rubric

Marzano's taxonomy has been used recently to guide the construction of course materials, learning objectives, and assessment questions in different contexts.[183–187] However, in these examples, the courses were developed such that the assessment materials were aligned with specific, assessable, learning objectives. This "backwards" course design is often not employed, especially in higher-education settings, and therefore many instructors have an abundance of test questions that are not explicitly tied to any learning objectives, but rather test some aspect of the current topic that the class is covering. One way that an educational taxonomy can be used is to classify existing exam problems according to the cognitive effort required of the student. An educational taxonomy can also be used to guide the construction of problems to target the specific cognitive level that the instructor is requiring a student to use when solving problems. These have both been done with Bloom's taxonomy in biology[70] and chemistry,[76] as well as with Marzano's taxonomy in physics.[187]

In this study, we developed a rubric from Marzano's taxonomy that is specific to multiple choice questions from the first quarter of a three-quarter general chemistry sequence at the University of Washington. The rubric is called the Marzano's Taxonomy rubric for chemistry multiple-choice exam items (CheMT/MC). Typically, Marzano's taxonomy places all multiple choice questions in the Retrieval category so we are pushing the descriptions of the levels in this rubric to include multiple choice questions that can target higher level thinking in our classification. The biggest change that we are making in the inclusion of multiple choice questions in the higher levels is that the process involved in these higher levels typically involving explained the reasoning behind an answer. This can't be done in

multiple-choice questions, so we instead rely on assumptions made by the team to assume the process that a student uses to answer a question. We used this rubric to categorize a large bank of multiple choice questions collected over several quarters of the first quarter of general chemistry at the University of Washington. The focus on multiple choice questions was due to the high abundance of this question type in the assessments, a common practice at large-enrollment universities like UW. We intend that instructors use this tool to assess the cognitive effort required by their current multiple choice questions, and to construct new questions that target a specific cognitive level. Aligning the cognitive level of the question to the intent of the instructor gives a clearer way to determine the areas or concepts that students are doing well on or need help with as well.

6.2 Methods

To develop our implementation of Marzano's taxonomy for general chemistry, we employed a set of 404 exam items from the first course in the three-quarter, non-honors General Chemistry sequence at UW (GC1). All exam items were multiple-choice and instructor-authored, with one correct answer and three or four distractors. The set includes items from three professors who taught GC1 during the autumn quarters of the 2016-17, 2017-18, and 2018-19 academic years, and one professor who taught GC1 during the winter quarters of the 2017-18 and 2018-19 academic years. The items cover the following topics: quantum mechanics and electronic structure; localized-electron bonding model, Lewis structures, and VSEPR theory; mass-to-mass stoichiometry; reaction classes and solution-phase stoichiometry; chemical kinetics; and gas laws.

The 404 exam items were randomized, and all metadata (correct answer, author, topic, exam instance, and quarter) was removed and retained in a separate file. The items were separated into subsets of 25 items each (except for subset #16, which contained 29 items). Subsets 1-8 (200 items) were used to develop the Marzano's Taxonomy rubric for chemistry multiple-choice exam items (the "CheMT/MC" rubric); subsets 9-16 (204 items) were used to confirm the stability of the CheMT/MC rubric. Each subset contained the correct answers

on a separate page, but complete worked-out solutions to the questions were not provided to the rubric-development or -confirmation teams.

6.2.1 Development of the CheMT/MC rubric

The first eight subsets (comprising 200 items) were used to develop the CheMT/MC rubric. A rubric development team rated the items in subsets 1-8 according to Marzano's Taxonomy, in order to develop a chemistry specific language for the first three cognitive levels from MT: Retrieval, Comprehension, and Analysis. (Due to the multiple-choice format of the items we considered, Knowledge Utilization was not observed in our dataset.) The members of the development team varied over time as follows:

- *Jan 2020 - May 2021, subsets 1-2*: Development team is comprised of one UW teaching professor (Rater 1) and three UW undergraduates (Raters 2-4). Rater 1 has taught all courses in the main general chemistry sequence at UW multiple times. Rater 2 took the honors general chemistry sequence at UW; Raters 3 and 4 took the main general chemistry sequence at UW.

- *April - May 2021, subsets 3-4*: A chemistry professor from a small liberal arts college and an expert in Marzano's Taxonomy joins the development team (Rater 5).

- *Jun - Dec 2021, subsets 5-8*: Raters 1, 2, and 3 depart the development team; Raters 4 and 5 remain. A UW chemistry graduate student and experienced general chemistry TA (JF, Rater 6) joins the team.

Throughout the development period, the rubric development team worked on one subset at time, and met regularly to discuss their item ratings. The raters were blind to the ratings of the other raters until they met as a team to discuss a particular subset. If there there was not complete agreement among the individual ratings for an item, that item was discussed until consensus was achieved among the majority of raters. Raters were free to change their

ratings based on the discussion. In 85% of the discussions of items from sets 1-8, complete consensus was achieved among the members of the development team.

All development meetings took place via Zoom, which allowed us to record discussions. Starting in June 2021, Rater 6 viewed recordings of past and ongoing meetings to identify recurring points of inconsistencies between ratings. These were subsequently brought to the attention of the group to further refine the metrics used in rating the items. This led to the development of a four-step procedure for raters when assigning a Marzano's level to the items in each subset.

- **STEP 1: Attempt to solve the problem to the best of your ability**, to ensure full engagement with the content covered in the item.

- **STEP 2: The rater records their assumptions about what topics/skills the student had practiced or experienced in the course prior to encountering the problem on an exam**. This could include whether a student had seen a similar question solved in class, or had frequently worked such questions in their homework. During subsequent discussions of the ratings, details about the raters' assumptions helped us to diagnose the cause of inconsistencies. For example, one rater assumed that students had enough prior experience with electron configurations to have developed mastery, and therefore rated a question on this topic at a lower cognitive level than another rater who did not assume such experience.

- **STEP 3: The rater explicitly wrote out the steps that the problem requires the student to do in as much detail as possible, employing action verbs consistent with Marzano's descriptions of the taxonomic levels**. For example, "This question requires a student to recall the Lewis dot structure and formal charge algorithms and execute them to discern the proper representation of N3-." This can include identifying what we call a "quantitative escape hatch", in which a question can be answered through a method that is of a lower level on the taxonomy than the

intended approach to the problem, often involving the execution of a previously learned algorithm. This could include using a brute force calculation for each multiple choice option when a higher order relationship could have been found by analyzing differences in the choices.

- **STEP 4: Specifying what is strictly required to solve a given problem allowed the rater to complete the fourth step, assigning the appropriate level and sublevel of Marzano's taxonomy to the question**. If a rater identified multiple Marzano's levels for a particular problem, the highest cognitive level was reported.

This four-step procedure was used by Raters 4, 5, and 6 to rate subsets 5-8 during the CheMT/MC rubric development, and subsets 9-16 during rubric confirmation, described below.

6.2.2 Confirmation of the CheMT/MC rubric

Subsets 9-16 (204 items) were used to confirm the stability of the CheMT/MC rubric. The rubric confirmation team was comprised of Raters 4, 5, and 6.

Similar to the workflow during the rubric development period, the confirmation team worked on one subset at time, and met regularly to discuss their item ratings. The raters were blind to the ratings of the other raters until they met as a team to discuss a particular subset. If there was complete disagreement among the individual ratings for an item, it was discussed until consensus was achieved among the majority of raters. The inter-rater reliability (IRR) for three out of three raters agreeing *prior* to discussing the items was 65.7%; the IRR for two out of three raters agreeing was 92.2

6.3 Results and Discussion

6.3.1 Results of the coding

Using the rubric that we developed, the coding team (JF, RO, and ST) assigned 204 questions a level and sublevel from Marzano's taxonomy. The inter-rater reliability for this process indicated that 92% of the time, 2 out of 3 raters agreed on the level and sublevel assigned to a question. The results of this process are shown in Table 6.1.

Table 6.1: Perecentage of Questions at Each Sublevel of the CheMT/MC Rubric.

Level	Before Discussion[a]	After Discussion[b]
1.1	1.1%	0.7%
1.2	13.8%	14.2%
1.3	67.7%	63.0%
2.1	12.4%	14.9%
2.2	0.4%	0.7%
3.3	3.2%	3.3%
3.5	1.4%	3.3%

[a] Rating of questions determined by agreement of two out of three raters.

[b] Rating of questions determined by agreement of two out of three raters after discussion of questions with complete disagreement

The overwhelming majority of questions, 82.6% were assigned by the coding team to the first level of Marzano's taxonomy: Retrieval. The first sublevel, *1.1: Recognizing*, was made up of questions where the answer could be identified without deeper conceptual understanding of the topic.

1.1 Recognize Example

The formula for bromic acid is:

(a) $HBrO_4$

(b) $HBrO_2$

(c) $HBrO_3$

(d) HBr

(e) HBrO

This example from the test bank shows a question where the student is identifying the correct answer based on memorized information about the formulas of polyatomic molecules. There is not a need for the student to demonstrate any more knowledge about bromic acid or any of the other acids shown in options. These types of questions are easily practiced with flash cards or other memorization aides. There were few questions of these types that were assigned by the raters, showing a trend toward more questions where the student has to at least apply the knowledge about the concept to answer the question.

A slightly more common type of question that was assigned by the raters was *1.2: Recalling*. These types of questions require the student to recall a trend, fact, or relationship and then use it to answer the question.

1.2 Recall Example

In the gaseous phase, which of the following diatomic molecules would be the most polar?

(a) CsCl

(b) LiF

(c) CsF

(d) NaF

(e) NaCl

In this example the student uses periodic trends of electronegativity to determine the polarity of each of the molecules to find the most polar. There is still no need for the student to demonstrate a more conceptual understanding of polarity and what impacts it has on other topics in chemistry. The raters also determined that questions where the student only needs to recall a definition to answer the question would be placed in this level. The assumptions that are made for those types of questions are that the definitions have been taught in that wording, or that the student has encountered that wording in the textbook. If the wording is new to the student, then that would increase the conceptual load the student would need to use in order to answer the question, but this was unknown to the raters.

Most of the questions in the set were assigned to *1.3: Execution*. This is due to the large number of questions that—in the raters' estimation—only require the student to use an algorithm to solve the question whether or not the student actually comprehends the content at a deeper level.

1.3 Execute Example

How many of the following molecules and ions contain double or triple bonds?

N2 H_2CO C_2H_4 C_2H_6 SCN^-

(a) 1

(b) 2

(c) 3

(d) 4

(e) 5

This question requires the student to draw Lewis dot structures for molecules where they are unfamiliar with the number of double or triple bonds within the molecule. The process of drawing Lewis dot structures is typically taught in an algorithmic way where there are set steps that the student takes in order to create the structure. This meant, to the raters, that if a question only required this process to be carried out by the student, the question would be placed into *1.3: Execution*. Other examples of *1.3: Execution* questions we categorized in the set include: drawing Lewis dot structures, "plug and chug" type calculations that are solved with a given formula, or other problems that use a practiced step-by-step method. This means that these questions primarily test students on their use of well rehearsed algorithms, rather than their higher level conceptual understanding of the topics. This also includes many questions that, at first glance, seemed complex, but were instead either worded in a confusing manner, or were able to be solved by a quantitative approach that didn't include a deeper conceptual effort on the part of the student. In the former case, this presents issues of inequity for students where English is a second language.

The remaining 17.4% of questions are split between Comprehension and Analysis, with the former having a larger percentage of questions. The questions within the Comprehension

level are mainly classified as *2.1: Integrating*. Since multiple-choice questions do not allow the student to give voice to their explanations of certain topics in their own words, the rater team choice to classify questions of this type by analyzing whether the question required the student to demonstrate a "conceptual leap." We defined this term to be a task that the student had to do that it just outside of learned information. For example, a question that requires the student to demonstrate a further analysis after the execution of an algorithm.

2.1 Integrate Example

Consider the following nitrogen-oxygen bond lengths: Given these data, predict the nitrogen-oxygen bond lengths in the nitrite ion, NO_2^-.

Bond	Length (pm)
N–O	140
N=O	120

(a) 120 pm, 140 pm

(b) 130 pm, 130 pm

(c) 120 pm, 120 pm

(d) 140 pm, 140 pm

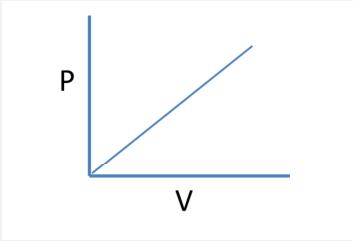
In the example above, the student utilizes their knowledge of resonance after likely drawing the Lewis dot structure of the molecule, to connect the ideas of resonance to bond lengths of molecules. In this example, the student is required to take an extra step after an established algorithm to analyze these concepts in tandem.

The relatively few cases of *2.2: Symbolizing* within multiple-choice questions is not surprising since these problems must include a symbolic representation that is unfamiliar to the student. By "unfamiliar to the student", we mean a symbolic representation that that was never explicitly shown in practice problems or lectures, but which the student can interpret correctly if they have integrated the concept or trend depicted in the symbolic

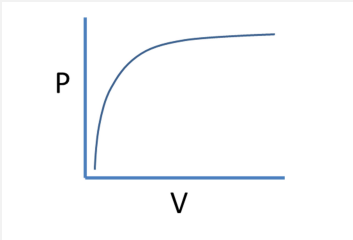
2.2 Symbolize Example

Which of the following plots shows the relationship between pressure and volume for an ideal gas?

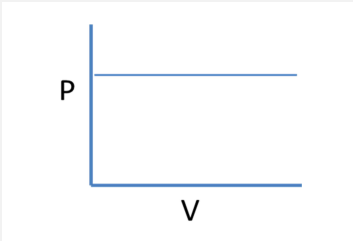
(a)



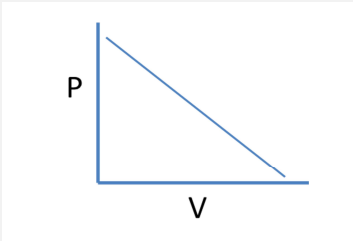
(b)



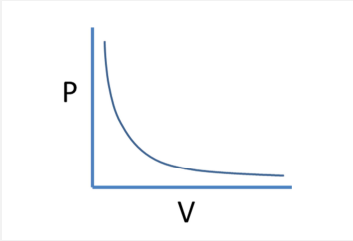
(c)



(d)



(e)



representation. However, this is all part of the assumptions that the coding team had to make, based on their knowledge of the curriculum, and typical support materials (textbook, lecture notes, etc.). In the example above, the student has to analyze the relationship between pressure and volume in a graphical representation. The assumption that the raters made is that the student had not observed this graphical relationship before and therefore had to think through this relationship mathematically. We did observe numerous questions that contained symbolic representations, but were conservative in the designation of *2.2: Symbolizing* since we did not have information about the specific course context for these questions.

There are several Analysis sublevels that the coding team did not rate among the questions present in the question bank: *3.1: Matching*, *3.2: Classifying*, and *3.4: Generalizing*. This does not mean that a multiple choice problem of these types can't be written, just that they were not rated in the set. Sublevels that were categorized in this set include *3.3: Analyzing Errors* and *3.5: Specifying*. Questions that were categorized within *3.3: Analyzing Errors* were ones where the raters found the student needed to analyze the validity of each answer choice within the question. The raters also found that these problems were not able to be solved algorithmically, so a "conceptual leap" was needed to answer the question correctly. The question below requires the student to determine the elementary steps for the rate law of the given reaction. For this question, the raters assumed that this way to determine the mechanisms for a reaction is not taught algorithmically. The answer choices also reinforce the need to determine if multiple mechanisms could be responsible for this reaction. Often, questions that were categorized in this level consisted of multiple-choice questions where each answer could be its own true/false question about a particular concept.

3.3 Analyze Errors Example

The rate law for the reaction $2A + B \rightarrow C + D$ is found to be Rate $= k[A]^2[B]$. Which of the following mechanisms gives this rate law?

I. $A + B \rightleftharpoons E$ (fast)
$E + B \rightleftharpoons C + D$ (slow)

II. $A + B \rightleftharpoons E$ (fast)
$E + A \rightleftharpoons C + D$ (slow)

III. $A + A \rightleftharpoons E$ (slow)
$E + B \rightleftharpoons C + D$ (fast)

(a) I only

(b) II only

(c) III only

(d) two of these

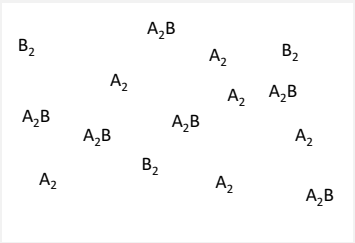
(e) none of these

3.5 Specify Example

Which 1 liter container below best represents the reaction $2A_2B(g) \rightarrow 2A_2(g) + B_2(g)$ after 30 seconds if the rate law is rate = 0.20 molecules A_2B/L second, if 10 molecules of A_2B is present initially in the one liter container?

(a)

(b)



(c)

(d)

The question above is an example of *3.5: Specifying.* The raters classified questions into

this sublevel when the student is required to predict a new trend, or relationship between concepts. The example problem has the student symbolize the reaction based on numbers of molecules in a container. The student then has to determine how that representation connects to rate laws and determine the instantaneous makeup of the container at a specified time. This requires the student to connect concepts of rates to this assumed new symbolism for the student. Upon further reflection by the team, some problems could have possibly been categorized as *3.5: Specifying* when they were more akin to a *3.4: Generalizing* problem. This is because the key features of *3.5: Specifying* problems are to predict new outcomes and new knowledge based on previously integrated concepts, while for *3.4: Generalizing* problems, the student is asked to make inferences and rules based on learned concepts. Therefore, a problem could have asked the student to infer a new relationship or rule, but then the student needed to use that relationship to predict new information. With so few examples of questions within the Analysis level, there was not a clear distinction that was put into the rubric that delineated these sublevels clearly.

We should note that the original MT states that multiple-choice questions could only be aligned with the *1.1: Recognizing* sublevel of Level 1: Retrieval, the lowest level in the taxonomy. However, we contend that if carefully crafted, multiple-choice type questions can be aligned with all the sublevels in Level 1, with both sublevels in Level 2: Comprehension, and at least some of the sublevels in Level 3: Analysis. To be sure, we think that multiple-choice questions for Levels 2 and Level 3 are more challenging to write than those for Level 1, and are sensitive to the local course context. For example, consider a question on the periodic trend of ionization energy, and how it can be rationalized using electron configuration. This question could align with Level 2 or perhaps Level 3, if the connection between these topics was not explicitly taught in the course, and the question is scaffolded appropriately to allow the student to connect these ideas in the moment they encounter the question on an exam. Such a question could also allow the student to "report" on the higher level thinking they had performed in their learning prior to the exam. If the connection *was* explicitly taught, then this question would align with either *1.1: Recognizing* or Recall in Level 1. (We

do not think that effective MC questions are possible in Level 4: Knowledge Utilization, since these sublevels require extended reasoning, synthesis, and experimentation that would be extremely challenging to implement in a multiple-choice question with only one correct answer.) The following sections use questions that were written for this study as examples to illustrate the specific attributes that each sub level has.

6.3.2 Rubric

The CheMT/MC rubric we developed is given in Table 6.2. The rubric includes a description of the sublevels within the first three levels of Marzano's taxonomy and how we interpreted them in the context of the multiple choice, GC1 questions in our question bank, as well as examples of the types of problems we assigned to these sublevels. We did not observe any questions in the set that could be categorized as *3.1: Matching*, *3.2: Classifying*, or *3.4: Generalizing*, so there are no example questions for these sublevels. It is important to remind the reader that the items in our question bank were instructor-authored—or modified from an existing test bank—with the goal of "covering topics" on the exam, and were not written to align with assessable student learning goals or even with Marzano's Taxonomy. Therefore, we find it unsurprising that we did not observe questions that aligned with the *3.1: Matching*, *3.2: Classifying*, or *3.4: Generalizing* sublevels of Level 3: Analysis. In the following sections, questions about one topic (quantum mechanics) are written for each sublevel to demonstrate key aspects of each sublevel in the rubric.

Table 6.2: Chemistry using Marzano's Taxonomy Applied to Multiple Choice Questions (CheMT/MC)

Level	Level 1:Retrieval			Level 2: Comprehension	
Sub-level	1: Recognizing	2: Recalling	3: Executing	1: Integrating	2: Symbolizing
Student is asked to ...	Identify an idea or concept immediately	Remember and apply the appropriate idea, concept, or trend in order to solve the problem	Recall and carry out the steps to the algorithm appropriate to the problem	Perform a conceptual leap in order to infer a relationship based on recalled information	Create or identify a new visual representation based on recalled information
General Examples	Immediately identify correct definitions	Periodic trends, relationships in electromagnetic spectrum, identifying bond polarity	Stoichiometry, Drawing Lewis dot structures, balancing equations/ redox reactions	What new representation best explains a concept/trend	Which figure/ plot/diagram best represents a concept/trend

Level	Level 3:Analysis				
Sub-level	1: Matching	2: Classifying	3: Analyzing Errors	4: Generalizing	5: Specifying
Student is asked to ...	Compare and contrast integrated concepts and topics	Organize or sort integrated concepts and topics	Evaluate validity of various statements about a topic or series of topics	Create a new principle based on integrated concepts and topics	Predict new relationships or concepts from integrated concepts
General Examples	Not seen is this work	Not seen is this work	Which of the following is true given known concepts	Not seen is this work	Deduce what happens based on known concepts

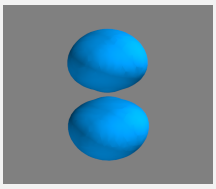
Level 1: Retrieval

The sublevels of Level 1: Retrieval are *1.1: Recognizing*, *1.2: Recalling*, and *1.3: Executing*. In the context of multiple choice questions in our question set, *1.1: Recognizing* questions are those in which the answer can be identified among the choices without deeper analysis or understanding. These questions ask the student to identify something that is well memorized such as the molecular formula for a polyatomic ion given its name, or the charge on a sodium ion. The first question in Table 6.3 has the student identify the orbital from its shape; they are not required to demonstrate any further knowledge of the quantum mechanics that pertain to these orbitals.

1.2: Recall questions require the student to recall a fact, trend, relationship, or concept and apply that to answer the question. This can include questions where student is required to use a memorized periodic trend to answer questions about bond polarity, correctly remembering the definition for a particular concept, or identifying the correct solute using memorized solubility rules. The second question of Table 6.3 has the student recall the formula to determine the number of nodes in an orbital and determine the nodes based on the quantum numbers in the question. There is no need for the student to demonstrate any knowledge of orbitals beyond this formula.

An *1.3: Execution* question requires the student to execute a memorized algorithm. Algorithms can include calculations involving unit conversions and stoichiometry, or that use a formula such as PV=nRT or the Rydberg formula. Algorithms can also be non-mathematical, such as drawing Lewis dot structures, or writing electron configurations. In the third question of Table 6.3, the student is asked to use the rules of quantum numbers to determine which values of ml and ms are allowed. No deeper analysis of what those rules mean and why they exist are required. *1.3: Execution* is the only sublevel in our rubric for which the mental procedures domain of knowledge is necessary to describe the cognitive process required, because *1.3: Execution* does not have an information domain.[15] We assigned questions to this sublevel if all they require of the student is the execution of a

Table 6.3: Example General Chemistry Multiple-Choice Questions About Quantum Mechanics Written at the Retrieval Level According to the ChemMT/MC Rubric

Sublevel	Example Problem	Key Aspects
1.1: Recognize	 What type of orbital is pictured above? (a) s (b) p (c) d (d) f	This problem only requires the student to identify, without further analysis.
1.2: Recall	How many total nodes are present in a 3p orbital? (a) 0 (b) 1 (c) 2 (d) 3 (e) 4	The student is now required to recollect the rules of nodes present in orbitals and use that in this situation.
1.3: Execute	An electron has the following quantum numbers: n=2, l=1. Which of the following values of m_l and m_s are allowed with the above values of n and l? (a) $m_l = 0$; $m_s = 0$ (b) $m_l = 2$; $m_s = -1/2$ (c) $m_l = 1$; $m_s = 1$ (d) $m_l = -1$; $m_s = 1/2$	This question now requires the student to recollect the rules of quantum numbers and apply those rules algorithmically.

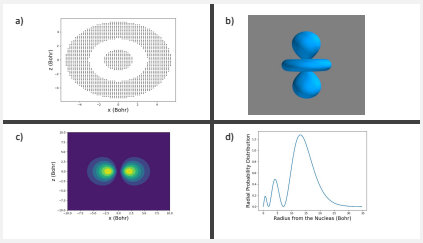
well-rehearsed algorithm, even if the question could also be answered by taking a more conceptually complex route. The sublevels within the Retrieval level are hierarchical; in order for the student to *Execute* an algorithm, they need to *Recall* the algorithm, after first *Recognizing* that the problem requires that algorithm. Therefore, the Retrieval sublevels build upon each other.

Level 2: Comprehension

There are two sublevels under Comprehension: *Integrating* and *Symbolizing*. In Comprehension, the student must demonstrate that the concept has been moved from their working memory into their permanent memory.[15] This is what we define in the rubric as a *conceptual leap*: the student demonstrates that they have taken a topic or concept that was in their working memory and reiterated it in a new way (to the student) that demonstrates the conceptual understanding of that topic. For *2.1: Integrating*, this means that the student has taken the concept that they have learned from lecture or studying, and are inferring some new relationship or trend based upon that concept in an explanation.

In the first question in Table 6.4, the student must recall that the m_s quantum number refers to spin, and that a noble gas contains a full valence shell of electrons. Further, the student must also recall the Pauli exclusion principle and its stipulation that all the electrons in an orbital must have unique spin. Finally, the student must put all the pieces together to determine which of the elements would have a complete valence shell in this alternative universe. This question requires the student to demonstrate their comprehension of the spin quantum number and what it represents. Note that it does not matter if the student has made this conceptual leap in the moment during the assessment, or while they were studying before the test. What the assessment should be testing is the students ability to demonstrate their proficiency in the topic. This is because, with increased proficiency in these topics and concepts, the action that the level at which the student is approaching the topic should decrease. This is a key feature of Marzano's taxonomy and is why the raters had to state the assumptions that they made about the students practice with that type of question.

Table 6.4: Example General Chemistry Multiple-Choice Questions About Quantum Mechanics Written at the Comprehension Level According to the ChemMT/MC Rubric

Sublevel	Example Problem	Key Aspects
2.1: Integrate	In an alternate universe, the spin quantum number, m_s, can be -1, 0, and 1 instead of -1/2 and 1/2. Assume all other aspects of quantum mechanics in our universe, such as the Pauli exclusion principle, still hold. In this universe, what would the first three Noble gases be? (a) He, Ne, Ar (b) Li, P, Co (c) Li, Mg, Se (d) Li, C, P	This problem requires the student to demonstrate a "conceptual leap" to determine the implications of changing the rules of quantum numbers.
2.2: Symbolize	 Which figure best represents the electron density for the quantum numbers n=3 and l=0?	This problem requires the student to determine a new (to the student) representation of the concept.

2.2: Symbolizing requires the student to identify a symbolic representation (an image, graph, schematic, etc.) that best represents a concept that they have Integrated. Crucially, though the symbolic representation(s) employed in the question should be *new* to the student. That is, they should be a version of a symbolic representation that has not been part of formal teaching. Otherwise, such a question would just be *1.2: Recall*. This was an assumption that the coding team had to make when they encountered questions that contained symbolic representations, as we only had the questions and not their specific course context. The second question in Table 6.4 asks the student to determine which of the images (some more familiar to the student than others) depicts an orbital with the given quantum numbers. This requires the student to identify the number and type of nodes that the orbital can have, as well as demonstrate that they can analyze how they would visualize an orbital with that number of nodes. An assumption in this problem is that the student is unfamiliar with at least some of these representations of orbitals, or at least that they are not heavily represented in the homework and practice problems available to the student.

Level 3: Analysis

Within the Analysis level there are five sublevels; *Matching*, *Classifying*, *Analyzing Errors*, *Generalizing*, and *Specifying*. Within the question set that the raters had, the raters did not categorize any questions within the sublevels of *3.1: Matching*, *3.2: Classifying*, or *3.4: Generalizing*. Therefore, for the purpose of this discussion, only *3.3: Analyzing Errors* and *3.5: Specifying* will be discussed.

3.3: Analyzing Errors questions have the student analyze what is true or false within a statement or procedure and why. These problems are often questions where each of the multiple-choice options act as true or false questions based on the problem where the student also has to identify why those options are true or false. In the first example of Table 6.5, the student has to execute the algorithm of the rules of quantum numbers in order to determine the number of electrons that follow those rules. The student then has to determine what rule specifically makes the answer correct as well as the validity of those rules. The difference

between this and a multiple-choice question at a lower level can be seen when comparing the example for *1.3: Execution* to the above question. In the latter case, the student is required to analyze each option and determine its validity in relation to the question stem and the other options.

In a *3.5: Specifying* question the student, must predict the outcome to a situation that is unfamiliar to them by creating a new (to the student) relationship, trend, or concept from previously learned concepts. Similar to the *2.1: Integrating* example, the student has to make the "conceptual leap", or demonstrate that they can connect the ideas of a full octet of electrons to the spin of electrons. However, they then have to relate that to the concept of stability which is a new (to the student) connection to the spin quantum numbers. The difference between this and *2.1: Integrating* is the creation of new knowledge as opposed to a reframing of existing knowledge. This is a key distinction that was used in the creation of the rubric and when rating the questions.

Table 6.5: Example General Chemistry Multiple-Choice Questions About Quantum Mechanics Written at the Analysis Level According to the ChemMT/MC Rubric

Sublevel	Example Problem	Key Aspects
3.3: Analyze Errors	An electron has the following quantum numbers; n=4 and m_l=-1. How many different electrons can have these quantum numbers and why? (a) 6; l has to be less than n and greater than -n so there are six orbitals the electrons can be in (b) 8; m_s can be -1/2 or 1/2 so there should be twice as many electrons as possible orbitals (c) 6; m_l is -1 so l has to be 1 or greater (d) 4; l can be between 0 and n-1 so there should be four possible orbitals	This problem requires the student to analyze each answer for both the validity of the number of unique quantum numbers and the rules of quantum numbers.
3.5: Symbolize	In an alternate universe, the spin quantum numbers, m_s, can be -1, 0, and 1 instead of -1/2 and 1/2. Assume all other aspects of quantum mechanics in our universe, such as the Pauli exclusion principle, still hold. In this universe, which atom would be the most stable? (a) H (b) He (c) Li (d) Be	This problem requires the student to demonstrate a "conceptual leap" to determine the implications of changing the rules of quantum numbers, while predicting stability under new conditions.

6.3.3 Implications for Practice

Although multiple choice questions are not the most effective way to assess higher order thinking, at large enrollment institutions ease of grading becomes a limiting factor. Therefore, having a rubric to assess the cognitive load of exam questions gives instructors the opportunity to target specific cognitive targets for a particular learning goal. This can also allow instructors to assess the cognitive level of their own questions in the development of student learning outcomes (SLOs) so that there are specific questions that test each SLO.[70,183,184,188] This can then influence the in-class activities to encourage the higher order thinking that would be required in order to answer the questions pertaining to those learning goals.

In the development of multiple choice questions that do target high level thinking, this work particularly highlights the need to analyze what level of thinking is *required* of the student to answer a question, as opposed to the level of thinking that the instructor might have *intended* for the student to exhibit. Indeed, one of the authors (CC, who was not involved in rating the questions) was routinely surprised at how many of their questions ended up in Level 1. The question rating scheme that emerged from conversations among the raters When investigating what the cognitive level a question is at, it is essential to determine the minimum set of things that a student needs to know in order to get the answer right, not necessarily through a deeper understanding of the content. According to the rating team, many questions that ostensibly targeted higher order thinking could actually be solved via a series of algorithmic calculations, without an exhibition of deeper understanding or the creation of any new knowledge in the process. We referred to this option as a "quantitative escape hatch" (QEH), since it allowed the student to avoid higher order thinking and solve the question using algorithms alone.[70,71]

Instructors could also teach students to use the rubric to analyze their own learning.[70,184] By assigning clear levels to problems that the students are aware of, this can help them analyze their own learning and take steps in their further studying practices to focus on areas that need further proficiency. This allows the student to focus on the Metacognitive

level of Marzano's taxonomy and determine the ways that help them focus their studying on concepts that they are struggling with as opposed to the ones that they are already showing proficiency in.

6.4 Conclusion

In conclusion, we have developed a rubric using Marzano's taxonomy to classify the cognitive level of general chemistry multiple choice questions. We used the rubric to classify a large set of multiple choice questions from the first quarter of a three quarter general chemistry class. The majority of questions were rated at a low cognitive level. Many questions that seemed complex or higher order, were either poorly worded, which would unfairly negatively bias the results from students where English is their second language, or the question could be approached through a fully quantitative method that circumvented any attempt for the instructor to determine if the student could show a deeper understanding of the concept. This came primarily from analyzing what the minimum set of requirements to answer the question correctly were. This highlights the need to analyze the questions included in assessments and development assessments with questions that target higher order thinking to align with the learning goals of the class. A limitation of this work was that the raters had to make numerous assumptions had due to their limited knowledge of the course context of the questions. This naturally leads to future work where we will conduct cognitive interviews with actual students who have taken GC1 at UW to elicit the approaches that students in the class would take when given these problems.[189,190] This will allow us to assess the validity of the assumptions that the raters made when classifying the problems as well as given us deeper insight into various approaches that students will use for these questions.

Chapter 7

SUMMARY AND FUTURE WORK

This dissertation has investigated both the use of diffusion Monte Carlo to study systems with large amplitude motions and the use of Marzano's taxonomy to study the cognitive level of general chemistry questions. For the former, a major goal of the study was to develop methods that investigated the spectral signatures in large protonated water clusters using the high quality wave functions that can be obtained from DMC. This required the construction of a general guiding function that could describe the highly anharmonic OH stretches in these systems to perform simulations of these large clusters. In this work, we used guided diffusion Monte Carlo to explore the effects of hydrogen-bonding in protonated water clusters. The last part of this dissertation focused on analyzing multiple choice questions using a rubric developed using Marzano's taxonomy. This work promotes the use of questions that require higher-order thinking to answer correctly as opposed to questions that can be answered using more quantitative, algorithmic, methods that don't require students to demonstrate a deeper conceptual proficiency of the topic.

After the review of guided diffusion Monte Carlo in Chapter 2, this approach was used in Chapter 3 to develop a guiding function for two systems that exhibit large amplitude motions; $H^+(H_2O)_{1-4}$ and CH_5^+. Both of these systems have vibrations that depend on their local environment, which was incorporated into the construction of the guiding functions for these modes. The use of these guiding functions showed large reductions in the simulations size that was required to obtain converged ground state energies and wave functions for these systems. By projecting the ground state probability amplitude onto a large amplitude motion of $H^+(H_2O)_3$ that is not explicitly described in the guiding function, we showed that the reduced ensemble sizes used in simulations that contain a guiding function showed an

improvement in the quality of the wave functions obtained.

The general guiding function for protonated water clusters developed in Chapter 3 was used in DMC calculations in Chapter 4 to investigate the isomerization of $H^+(H_2O)_6$ upon deuteration. Comparison of the zero-point energy corrected energy among the deuterated froms of the three isomers showed differences <0.5 kcal/mol. This indicates that each isomer could have significant populations at low temperatures. The trends in energy are consistent between levels of theory and between harmonic and anharmonic calculations. Vibrational perturbation theory was used to investigate the possible contributions to the spectrum of $H^+(H_2O)_6$ and $D^+(D_2O)_6$ from each of the three isomers. The perturbation theory showed consistent results for $H^+(H_2O)_6$ where the calculated spectrum of the T2 isomer, which contains a four membered ring, displayed peaks in the calculated spectrum that are not in the experimental spectrum however, the T2 isomer has peaks in the calculated $D^+(D_2O)_6$ spectrum that are present in the experimental spectrum. This leads to the conclusion that the $D^+(D_2O)_6$ spectrum has population of an isomer that is similar to the T2 isomer.

Continuing the investigation of the isomerization of $H^+(H_2O)_6$ upon deuteration, a spectral map was developed that utilizes the vibrationally averaged hydrogen-bonded OH distances obtained from diffusion Monte Carlo to predict spectral features. This method was compared against scaled harmonic and vibrational perturbation theory calculations, as well as experiment, to determine the information that can be gained from this method compared to others. The approach was calibrated against two smaller protonated water clusters, $H^+(H_2O)_{3-4}$ and $D^+(D_2O)_{3-4}$. For $H^+(H_2O)_3$ and $D^+(D_2O)_3$ the map showed slightly closer agreement to experiment than the other two calculated methods. For $H^+(H_2O)_4$ and $D^+(D_2O)_4$ the spectral map showed similar results to scaled harmonic, indicating that the map shows better agreement to experiment for vibrational modes with large anharmonicities. When applying the spectral map to four isomers of $D^+(D_2O)_6$ we showed that the isomer that displays the closest match to experiment in the high frequency region of the spectrum are the T2 and T4 isomers.

Looking forward, Diffusion Monte Carlo is a powerful method for studying systems that

exhibit large amplitude motions, but one limitation is the computational cost of evaluating the potential energy of every walker at every time step. Guided DMC has been used with great success to significantly reduce the ensemble size that is required to obtain high quality ground state wave functions and ground state energies.[13,27,35,97] This has allowed exploration of systems that were previously intractable.[13,35] However, even with this reduction in ensemble size, the large number of potential energy evaluations greatly limit the size of system that we can study. Recent research in the group has explored the use of neural network potential energy surfaces to fit the potential energy surface which leads to substantial reductions in the cost of evaluating the potential energy.[117,118] This has also lead to studies of $C_2H_5^+$ that does not have an established potential energy surface by using electronic structure calculations from the MOB-ML model as inputs to the neural network.[45,118]

In Chapter 2, we showed extensions of the guided DMC method where an excited state wave function wave used in the guiding function to obtain excited state information, such as frequencies and matrix elements involving the ground state and the excited state. In that study there was no assumption that an excited state guiding function allows a reduced ensemble size but it is likely that the guiding function approach could yield comparable reductions in the required ensemble size as seen in the ground state. If a reduced ensemble size is shown to yield reductions in ensemble size compared to fixed-node DMC then this method could be used to calculate vibrational frequencies in molecules that were intractable without that reduction.

Finally, the last chapter utilizes Marzano's taxonomy to develop a rubric that categorizes multiple choice questions in general chemistry based on the cognitive effort that a student would require to correctly answer the question. This rubric was then applied to a large test set of questions by a team of raters. The results of this analysis showed a large percentage of questions that were categorized as low-level, algorithmic problems. Many questions that at first glance may have seemed complex were found to be categorized at a low-level due to the possibility of using a quantitative approach to answer the question without requiring a deeper conceptual understanding of the topic. The work highlights the need to analyze

questions for what the student is required to do as opposed to what the instructor expects the student to do.

The limitation of the work using Marzano's taxonomy is the assumptions that the raters make about the students' experience when answering the questions. This is due to the limited knowledge that the raters had about the amount of practice a student has with a certain type of question, how familiar a student is with a specific representation of a concept, what equations the student is given in the exam, and other factors that influence how a student approaches a problem. Due to how the proficiency is described within Marzano's taxonomy, a question that may be conceptually complex can become a lower level question with enough practice. Therefore, an extension of the work would be to conduct cognitive interviews to gather qualitative data about how students in the class would approach these problems.[189] This would allow us to compare the process that a student would take to answer a problem to the work that was proposed by the rater team. This could be a way to analyze the validity of the various assumptions that were made as part of the rating process.

BIBLIOGRAPHY

[1] Bloom, B. S. *Taxonomy of Educational Objectives: The Classification of Educational Goals.;* Longmans: London, 1956.

[2] Anderson, L. W.; Krathwohl, D. R.; Airasian, P. W.; Cruikshank, K. A.; Mayer, R. E.; Pintrich, P. R.; Raths, J.; Wittrock, M. C., Eds.; *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives;* Addison Wesley Longman, Inc.: New York, 2001.

[3] Kreitzer, A. E.; Madaus, G. F. *Bloom's taxonomy: A forty-year retrospective: Ninety-third yearbook of the National Society for the Study of Education* 1994 pages 64–81 Publisher: University of Chicago Press Chicago, IL.

[4] Boyer, M. A.; Marsalek, O.; Heindel, J. P.; Markland, T. E.; McCoy, A. B.; Xantheas, S. S. *J. Phys. Chem. Lett.* 2019 *10*(5), 918–924.

[5] Jin, Z.; Braams, B. J.; Bowman, J. M. *J. Phys. Chem. A* 2006 *110*, 1569–1574.

[6] Heine, N.; Fagiani, M. R.; Rossi, M.; Wende, T.; Berden, G.; Blum, V.; Asmis, K. R. *J. Am. Chem. Soc.* 2013 *135*(22), 8266–8273.

[7] Wolke, C. T.; Fournier, J. A.; Dzugan, L. C.; Fagiani, M. R.; Odbadrakh, T. T.; Knorke, H.; Jordan, K. D.; McCoy, A. B.; Asmis, K. R.; Johnson, M. A. *Science* 2016 *354*(6316), 1131–1135.

[8] Yu, Q.; Bowman, J. M. *J. Chem. Phys.* 2017 *146*(12), 121102.

[9] Duong, C. H.; Yang, N.; Johnson, M. A.; DiRisio, R. J.; McCoy, A. B.; Yu, Q.; Bowman, J. M. *J. Phys. Chem. A* 2019 *123*(37), 7965–7972.

[10] Duong, C. H.; Yang, N.; Kelleher, P. J.; Johnson, M. A.; DiRisio, R. J.; McCoy, A. B.; Yu, Q.; Bowman, J. M.; Henderson, B. V.; Jordan, K. D. *J. Phys. Chem. A* 2018 *122*(48), 9275–9284.

[11] Fagiani, M. R.; Knorke, H.; Esser, T. K.; Heine, N.; Wolke, C. T.; Gewinner, S.; Schöllkopf, W.; Gaigeot, M.-P.; Spezia, R.; Johnson, M. A.; Asmis, K. R. *Phys. Chem. Chem. Phys.* 2016 *18*, 26743–26754.

[12] Heine, N.; Fagiani, M. R.; Rossi, M.; Wende, T.; Berden, G.; Blum, V.; Asmis, K. R. *J. Am. Chem. Soc.* 2013 *135*(22), 8266–8273.

[13] Finney, J. M.; Choi, T. H.; Huchmala, R. M.; Heindel, J. P.; Xantheas, S. S.; Jordan, K. D.; McCoy, A. B. *J. Phys. Chem. Lett.* 2023 *14*(20), 4666–4672.

[14] Marzano, R. J.; Kendall, J. S. *The New Taxonomy of Educational Objectives;* Corwin Press: Thousand Oaks, California, 2 ed.; 2007.

[15] Marzano, R. J.; Kendall, J. S. *Designing and Assessing Educational Objectives : Applying the New Taxonomy;* Corwin Press: Thousand Oaks, California, 2008.

[16] Sibert, E. L. *J. Chem. Phys.* 1988 *88*(7), 4378–4390.

[17] Barone, V. *J. Chem. Phys.* 2005 *122*(1), 014108.

[18] Bowman, J. M. *Acc. Chem. Res.* 1986 *19*, 202–208.

[19] R. B. Gerber, V. Buch and M. A. Ratner, *J. Chem. Phys.* 1982 *77*, 3022–3030.

[20] Bowman, J. M.; Carter, S.; Huang, X. *Int. Rev. Phys. Chem.* 2003 *22*, 533–549.

[21] Wang, Y.; Bowman, J. M. *J. Chem. Phys.* 2011 *134*(15), 154510.

[22] Anderson, J. B. *J. Chem. Phys.* 1975 *63*(4), 1499–1503.

[23] Anderson, J. B. *J. Chem. Phys.* 1976 *65*(10), 4121–4127.

[24] Suhm, M. A.; Watts, R. O. *Phys. Rep* 1991 *204*(4), 293 – 329.

[25] Reynolds, P. J.; Ceperley, D. M.; Alder, B. J.; Lester, W. A. *J. Chem. Phys.* 1982 *77*(11), 5593–5603.

[26] McCoy, A. B. *Int. Rev. Phys. Chem.* 2006 *25*(1-2), 77–107.

[27] Lee, V. G. M.; McCoy, A. B. *J. Phys. Chem. A* 2019 *123*(37), 8063–8070.

[28] Wang, X. G.; Carrington Jr., T. *J. Chem. Phys.* 2008 *129*, 234102.

[29] Vendrell, O.; Gatti, F.; Meyer, H.-D. *J. Chem. Phys.* 2007 *127*, 184303.

[30] Esser, T. K.; Knorke, H.; Asmis, K. R.; Schöllkopf, W.; Yu, Q.; Qu, C.; Bowman, J. M.; Kaledin, M. *J. Phys. Chem. Lett.* 2018 *9*(4), 798–803.

[31] Thompson, K. C.; Crittenden, D. L.; Jordan, M. J. T. *J. Am. Chem. Soc.* 2005 *127*, 4954–4958.

[32] McCoy, A. B.; Braams, B. J.; Brown, A.; Huang, X.; Jin, Z.; Bowman, J. M. *J. Phys. Chem. A* 2004 *108*(23), 4991–4994.

[33] Headrick, J. M.; Diken, E. G.; Walters, R. S.; Hammer, N. I.; Christie, R. A.; Cui, J.; Myshakin, E. M.; Duncan, M. A.; Johnson, M. A.; Jordan, K. D. *Science* 2005 *308*(5729), 1765–1769.

[34] Mallory, J. D.; Brown, S. E.; Mandelshtam, V. A. *J. Phys. Chem. A* 2015 *119*(24), 6504–6515.

[35] Lee, V. G. M.; Vetterli, N. J.; Boyer, M. A.; McCoy, A. B. *J. Phys. Chem. A* 2020 *124*(34), 6903–6912.

[36] Talbot, J. J.; Yang, N.; Huang, M.; Duong, C. H.; McCoy, A. B.; Steele, R. P.; Johnson, M. A. *J. Phys. Chem. A* 2020 *124*(15), 2991–3001.

[37] Schmidt, M.; Constable, S.; Ing, C.; Roy, P.-N. *J. Chem. Phys.* 2014 *140*(23), 234101.

[38] Kohno, B. H.; Mandelshtam, V. A. *The Journal of Physical Chemistry A* 2020 *124*(42), 8766–8777 PMID: 32960063.

[39] Acioli, P. H.; Xie, Z.; Braams, B. J.; Bowman, J. M. *J. Chem. Phys.* 2008 *128*, 104318.

[40] Hammer, N. I.; Diken, E. G.; Roscioli, J. R.; Johnson, M. A.; Myshakin, E. M.; Jordan, K. D.; McCoy, A. B.; Huang, X.; Bowman, J. M.; Carter, S. *J. Chem. Phys.* 2005 *122*(24), 244301.

[41] McCoy, A. B.; Diken, E. G.; Johnson, M. A. *J. Phys. Chem. A* 2009 *113*(26), 7346–7352.

[42] Lee, V. G. M.; Madison, L. R.; McCoy, A. B. *J. Phys. Chem. A* 2019 *123*(20), 4370–4378.

[43] Lin, Z.; McCoy, A. B. *J. Phys. Chem. A* 2013 *117*(46), 11725–11736.

[44] Lee, H.-S.; McCoy, A. B. *J. Chem. Phys.* 2002 *116*, 9677–9689.

[45] Moonkaen, P.; Finney, J. M.; McCoy, A. B. *The Journal of Physical Chemistry A* 2023 *127*(5), 1196–1205.

[46] Severson, M. W.; Buch, V. *J. Chem. Phys.* 1999 *111*, 10866–10875.

[47] Lee, H.-S.; Herbert, J. M.; McCoy, A. B. *J. Chem. Phys.* 1999 *110*(12), 5481–5484.

[48] Lee, H.-S.; Herbert, J. M.; McCoy, A. B. *J. Chem. Phys.* 1999 *111*(20), 9203–9212.

[49] Finney, J. M.; DiRisio, R. J.; McCoy, A. B. "Diffusion Monte Carlo Approaches for Studying Large Amplitude Vibrational Motions in Molecules and Clusters", submitted as a chapter in *A Computational Approach to Vibrational Dynamics of Molecules*, edited by J. M. Bowman, 2022.

[50] Guasco, T. L.; Johnson, M. A.; McCoy, A. B. *J. Phys. Chem. A* 2011 *115*(23), 5847–5858.

[51] DiRisio, R. J.; Finney, J. M.; Dzugan, L. C.; Madison, L. R.; McCoy, A. B. *J. Phys. Chem. A* 2021 *125*(33), 7185–7197.

[52] Xantheas, S. S.; Dunning, T. H. *J. Chem. Phys.* 1993 *99*(11), 8774–8792.

[53] Miliordos, E.; Aprà, E.; Xantheas, S. S. *J. Chem. Phys.* 2013 *139*(11), 114302.

[54] Badger, R. M. *J. Chem. Phys.* 1934 *2*(3), 128–131.

[55] Heindel, J. P.; Yu, Q.; Bowman, J. M.; Xantheas, S. S. *J. Chem. Theory Comput.* 2018 *14*(9), 4553–4566.

[56] Gruenbaum, S. M.; Tainter, C. J.; Shi, L.; Ni, Y.; Skinner, J. L. *J. Chem. Theory Comput.* 2013 *9*(7), 3109–3117 PMID: 26583990.

[57] Ni, Y.; Skinner, J. L. *J. Chem. Phys.* 2015 *143*(1) 014502.

[58] Kananenka, A. A.; Skinner, J. L. *J. Chem. Phys.* 2018 *148*(24), 244107.

[59] Tainter, C. J.; Ni, Y.; Shi, L.; Skinner, J. L. *J. Phys. Chem. Lett.* 2013 *4*, 12–17.

[60] Hayashi, T.; la Cour Jansen, T.; Zhuang, W.; Mukamel, S. *J. Phys. Chem. A* 2005 *109*(1), 64–82 PMID: 16839090.

[61] Eaves, J. D.; Tokmakoff, A.; Geissler, P. L. *J. Phys. Chem. A* 2005 *109*(42), 9424–9436.

[62] Yang, N.; Huchmala, R. M.; McCoy, A. B.; Johnson, M. A. *J. Phys. Chem. Lett.* 2022 *13*(34), 8116–8121.

[63] Hermansson, K. *J. Chem. Phys.* 1993 *99*(2), 861–868.

[64] Diken, E. G.; Robertson, W. H.; Johnson, M. A. *J. Phys. Chem. A* 2004 *108*(1), 64–68.

[65] Crandell, O. M.; Lockhart, M. A.; Cooper, M. M. *J. Chem. Educ.* 2020 *97*(2), 313–327.

[66] Momsen, J.; Offerdahl, E.; Kryjevskaia, M.; Montplaisir, L.; Anderson, E.; Grosz, N. *CBE Life Sci. Educ.* 2013 *12*(2), 239–249.

[67] Rappoport, L. T.; Ashkenazi, G. *Int. J. Sci. Educ.* 2008 *30*(12), 1585–1603.

[68] Stowe, R. L.; Scharlott, L. J.; Ralph, V. R.; Becker, N. M.; Cooper, M. M. *J. Chem. Educ.* 2021 *98*(8), 2490–2495.

[69] Ralph, V. R.; Scharlott, L. J.; Schafer, A. G. L.; Deshaye, M. Y.; Becker, N. M.; Stowe, R. L. *JACS Au* 2022 *2*(8), 1869–1880.

[70] Crowe, A.; Dirks, C.; Wenderoth, M. P. *CBE Life Sci. Educ.* 2008 *7*(4), 368–381.

[71] Momsen, J. L.; Long, T. M.; Wyse, S. A.; Ebert-May, D. *CBE Life Sci. Educ.* 2010 *9*(4), 435–440.

[72] Ralph, V. R.; Scharlott, L. J.; Schwarz, C. E.; Becker, N. M.; Stowe, R. L. *J. Res. Sci. Teach.* 2022 *59*(5), 841–875.

[73] Ralph, V. R.; States, N. E.; Corrales, A.; Nguyen, Y.; Atkinson, M. B. *Chem. Educ. Res. Pract.* 2022 *23*(3), 662–685.

[74] Shah, L.; Fatima, A.; Syed, A.; Glasser, E. *J. Chem. Educ.* 2022 *99*(1), 14–24.

[75] Prevost, L. B.; Lemons, P. P. *CBE Life Sci. Educ.* 2016 *15*(4), ar71.

[76] Geye, E. "The Process of Establishing a Blooming Chemistry Tool for Use in Undergraduate Chemistry Education and Research", Master's thesis, Portland State University, 2022.

[77] Zoller, U.; Dori, Y. *Int. J. Sci. Educ.* 2002 *24*(2), 185–203.

[78] Gierl, M. *J. Educ. Res.* 1997 *91*(1), 26–32.

[79] Stowe, R. L.; Cooper, M. M. *J. Chem. Educ.* 2017 *94*(12), 1852–1859 Type: Article.

[80] Biggs, J.; Collis, K. *Aust. J. Educ.* 1989 *33*(2), 151–163.

[81] Chan, C. C.; Tsui, M. S.; Chan, M. Y. C.; Hong, J. H. *Assess. Eval. High. Educ.* 2002 *27*(6), 511–527.

[82] Brabrand, C.; Dahl, B. *High. Educ.* 2009 *58*(4), 531–549.

[83] Smith, K. C.; Nakhleh, M. B.; Bretz, S. L. *Chem. Educ. Res. Pract.* 2010 *11*(3), 147–153.

[84] Zoller, U.; Lubezky, A.; Nakhleh, M. B.; Tessier, B.; Dori, Y. J. *J. Chem. Educ.* 1995 *72*(11), 987.

[85] Nurrenbern, S. C.; Pickering, M. *J. Chem. Educ.* 1987 *64*(6), 508.

[86] Stamovlasis, D.; Tsaparlis, G.; Kamilatos, C.; Papaoikonomou, D.; Zarotiadou, E. *Chem. Educ. Res. Pract.* 2005 *6*(2), 104–118.

[87] Dori, Y. J.; Hameiri, M. *J. Res. Sci. Teach.* 2003 *40*(3), 278–302.

[88] Wolfskill, T.; Hanson, D. *J. Chem. Educ.* 2001 *78*(10), 1417.

[89] Metropolis, N.; Ulam, S. *J. Am. Stat. Assoc.* 1949 *44*, 334–341.

[90] Benoit, D.; Lauvergnat, D.; Scribano, Y. *Faraday Discuss.* 2018.

[91] Sandler, P.; Buch, V.; Clary, D. C. *J. Chem. Phys.* 1994 *101*, 6353–5.

[92] Gregory, J. K.; Clary, D. C. *J. Chem. Phys.* 1995 *102*, 7817–7829.

[93] Huang, X.; Cho, H. M.; Carter, S.; Ojamäe, L.; Bowman, J. M.; Singer, S. J. *J. Phys. Chem. A* 2003 *107*(37), 7142–7151.

[94] Johnson, L. M.; McCoy, A. B. *J. Phys. Chem. A* 2006 *110*, 8213–8220.

[95] Lin, Z.; McCoy, A. B. *J. Phys. Chem. Lett.* 2012 *3*(24), 3690–3696.

[96] Kelleher, P. J.; Johnson, C. J.; Fournier, J. A.; Johnson, M. A.; McCoy, A. B. *J. Phys. Chem. A* 2015 *119*(18), 4170–4176.

[97] Finney, J. M.; DiRisio, R. J.; McCoy, A. B. *J. Phys. Chem. A* 2020 *124*(46), 9567–9577.

[98] Petit, A.; McCoy, A. B. *J. Phys. Chem. A* 2009 *113*, 12706–12714.

[99] Mallory, J. D.; Mandelshtam, V. A. *J. Chem. Phys.* 2016 *145*(6), 064308.

[100] Mallory, J. D.; Mandelshtam, V. A. *J. Phys. Chem. A* 2017 *121*(33), 6341–6348.

[101] Buch, V. *J. Chem. Phys.* 1992 *97*(1), 726–729.

[102] Valdès, A.; Prosmiti, R.; Delgado-Barrio, G. *J. Chem. Phys.* 2012 *136*(10), 104302/1–6.

[103] Cheng, T. C.; Jiang, L.; Asmis, K. R.; Wang, Y.; Bowman, J. M.; Ricks, A. M.; Duncan, M. A. *J. Phys. Chem. Lett.* 2012 *3*, 3160–3166.

[104] Boyer, M. A.; Chiu, C. S.; McDonald, D. C.; Wagner, J. P.; Colley, J. E.; Orr, D. S.; Duncan, M. A.; McCoy, A. B. *J. Phys. Chem. A* 2020 *124*(22), 4427–4439.

[105] Asvany, O.; Kumar, P.; Redlich, B.; Hegeman, I.; Schlemmer, S.; Marx, D. *Science* 2005 *309*, 1219–1222.

[106] Asvany, O.; Yamada, K. M. T.; Brünken, S.; Potapov, A.; Schlemmer, S. *Science* 2015 *347*(6228), 1346–1349.

[107] Kumar P, P.; Marx, D. *Phys. Chem. Chem. Phys.* 2006 *8*(5), 573–586.

[108] Braams, B. J.; Bowman, J. M. *Int. Rev. in Phys. Chem.* 2009 *28*, 577–606.

[109] Paesani, F. *Acc. Chem. Res.* 2016 *49*(9), 1844–1851.

[110] Viel, A.; Whaley, K. B.; Wheatley, R. J. *J. Chem. Phys.* 2007 *127*(19), 194303.

[111] Barone, V.; Alessandrini, S.; Biczysko, M.; Cheeseman, J. R.; Clary, D. C.; McCoy, A. B.; DiRisio, R. J.; Neese, F.; Melosso, M.; Puzzarini, C. *Nat. Rev. Methods Primers* 2021 *1*(1), 38.

[112] Petit, A. S.; Ford, J. E.; McCoy, A. B. *J. Phys. Chem. A* 2014 *118*(35), 7206–7220.

[113] B. L. Hammond, W. A. Lester, J.; Reynolds, P. J., Eds.; *Monte Carlo Methods in Ab Initio Quantum Chemistry;* volume 1 of *World Scientific Lecture and Course Notes in Chemistry* World Scientific Publishing Co.: Singapore, 1994.

[114] Barnett, R.; Reynolds, P.; W.A Lester, J. *J. Comput. Phys.* 1991 *96*(2), 258 – 276.

[115] Partridge, H.; Schwenke, D. W. *J. Chem. Phys.* 1997 *106*(11), 4618–4639.

[116] McCoy, A. B. Diffusion Monte Carlo Approaches for Studying Systems that Undergo Large Amplitude Vibrational Motions. In *Recent Advances in Quantum Monte Carlo Methods*, Vol. 3; Anderson, J. B.; Rothstein, S. M., Eds.; ACS Symposium Series 953: 2006.

[117] DiRisio, R. J.; Lu, F.; McCoy, A. B. *J. Phys. Chem. A* 2021 *125*(26), 5849–5859.

[118] Lu, F. *et al. J. Phys. Chem. A* 2022 *126*(25), 4013–4024.

[119] Colbert, D. T.; Miller, W. H. *J. Chem. Phys.* 1992 *96*, 1982–1991.

[120] Huang, X.; Carter, S.; Bowman, J. M. *J. Chem. Phys.* 2003 *118*, 5431–5441.

[121] Wilson, E. B.; Decius, J. C.; Cross, P. C. *Molecular Vibrations;* Dover: New York, 1955.

[122] McCoy, A. B.; Dzugan, L. C.; DiRisio, R. J.; Madison, L. R. *Faraday Discuss.* 2018 *212*, 443–466.

[123] McCoy, A. B.; Huang, X.; Carter, S.; Landeweer, M. Y.; Bowman, J. M. *J. Chem. Phys.* 2005 *122*, 061101.

[124] Schreiner, P. R.; Kim, S. J.; Schaefer, H. F., I.; Schleyer, P. v. R. M. *J. Chem. Phys.* 1993 *99*, 3716–3720.

[125] Mallory, J. D.; Mandelshtam, V. A. *J. Chem. Phys.* 2015 *143*(14), 144303.

[126] Schmiedt, H.; Jensen, P.; Schlemmer, S. *J. Chem. Phys.* 2016 *145*(7), 074301.

[127] Schmiedt, H.; Jensen, P.; Schlemmer, S. *Chem. Phys. Lett.* 2017 *672*, 34 – 46.

[128] Eigen, M. *Angew. Chem., Int. Ed. Engl.* 1964 *3*(1), 1–19.

[129] Zundel, G. *Adv. Chem. Phys.* 1999 pages 1–217.

[130] Bakker, H.; Nienhuys, H.-K. *Science* 2002 *297*(5581), 587–590.

[131] Xantheas, S. S. *Nature* 2009 *457*(7230), 673–674.

[132] Xantheas, S. S. *Science* 2016 *354*(6316), 1101–1101.

[133] Thämer, M.; De Marco, L.; Ramasesha, K.; Mandal, A.; Tokmakoff, A. *Science* 2015 *350*(6256), 78–82.

[134] Yeh, L.; Okumura, M.; Myers, J.; Price, J.; Lee, Y. *J. Chem. Phys.* 1989 *91*(12), 7319–7330.

[135] Jiang, J.-C.; Wang, Y.-S.; Chang, H.-C.; Lin, S. H.; Lee, Y. T.; Niedner-Schatteburg, G.; Chang, H.-C. *J. Am. Chem. Soc.* 2000 *122*(7), 1398–1410.

[136] Asmis, K. R.; Pivonka, N. L.; Santambrogio, G.; Brümmer, M.; Kaposta, C.; Neumark, D. M.; Wöste, L. *Science* 2003 *299*(5611), 1375–1377.

[137] Miyazaki, M.; Fujii, A.; Ebata, T.; Mikami, N. *Science* 2004 *304*(5674), 1134–1137.

[138] Shin, J.-W.; Hammer, N.; Diken, E.; Johnson, M.; Walters, R.; Jaeger, T.; Duncan, M.; Christie, R.; Jordan, K. D. *Science* 2004 *304*(5674), 1137–1140.

[139] Lin, C.-K.; Wu, C.-C.; Wang, Y.-S.; Lee, Y. T.; Chang, H.-C.; Kuo, J.-L.; Klein, M. L. *Phys. Chem. Chem. Phys.* 2005 *7*(5), 938–944.

[140] Headrick, J. M.; Diken, E. G.; Walters, R. S.; Hammer, N. I.; Christie, R. A.; Cui, J.; Myshakin, E. M.; Duncan, M. A.; Johnson, M. A.; Jordan, K. D. *Science* 2005 *308*(5729), 1765–1769.

[141] Mizuse, K.; Fujii, A. *J. Phys. Chem. Lett.* 2011 *2*(17), 2130–2134.

[142] Fujii, A.; Mizuse, K. *Int. Rev. Phys. Chem.* 2013 *32*(2), 266–307.

[143] Fournier, J. A.; Johnson, C. J.; Wolke, C. T.; Weddle, G. H.; Wolk, A. B.; Johnson, M. A. *Science* 2014 *344*(6187), 1009–1012.

[144] Fournier, J. A.; Wolke, C. T.; Johnson, M. A.; Odbadrakh, T. T.; Jordan, K. D.; Kathmann, S. M.; Xantheas, S. S. *J. Phys. Chem. A* 2015 *119*(36), 9425–9440.

[145] Khuu, T.; Rana, A.; Edington, S. C.; Yang, N.; McCoy, A. B.; Johnson, M. A. *J. Am. Soc. Mass Spectrom.* 2023 *34*(4), 737–744.

[146] Mella, M.; Kuo, J.-L.; Clary, D. C.; Klein, M. L. *Phys. Chem. Chem. Phys.* 2005 *7*, 2324–2332.

[147] Scuseria, G. E.; Janssen, C. L.; Schaefer III, H. F. *J. Chem. Phys.* 1988 *89*(12), 7382–7387.

[148] Dunning Jr, T. H. *J. Chem. Phys.* 1989 *90*(2), 1007–1023.

[149] de Lara-Castells, M.; Krems, R.; Buchachenko, A. A.; Delgado-Barrio, G.; Villarreal, P. *J. Chem. Phys.* 2001 *115*(22), 10438–10449.

[150] Purvis, G. D.; Bartlett, R. J. *J. Chem. Phys.* 1982 *76*(4), 1910–1918.

[151] Head-Gordon, M.; Pople, J. A.; Frisch, M. J. *Chem. Phys. Lett.* 1988 *153*(6), 503–506.

[152] Frisch, M. J.; Head-Gordon, M.; Pople, J. A. *Chem. Phys. Lett.* 1990 *166*(3), 275–280.

[153] Head-Gordon, M.; Head-Gordon, T. *Chem. Phys. Lett.* 1994 *220*(1), 122–128.

[154] Frisch, M. J. *et al.* "Gaussian 16 Revision C.01", 2016 Gaussian Inc. Wallingford CT.

[155] DiRisio, R. J.; Finney, J. M.; McCoy, A. B. *WIREs Computational Molecular Science* 2022 *12*(6), e1615.

[156] Martin, J. M. L.; Lee, T. J.; Taylor, P. R.; François, J. P. *J. Chem. Phys.* 1995 *103*(7), 2589–2602.

[157] Franke, P. R.; Stanton, J. F.; Douberly, G. E. *J. Phys. Chem. A* 2021 *125*(6), 1301–1324.

[158] Yang, Q.; Mendolicchio, M.; Barone, V.; Bloino, J. *Front. Astron. Space Sci.* 2021 *8*, 665232.

[159] Pérez, C.; Muckle, M. T.; Zaleski, D. P.; Seifert, N. A.; Temelso, B.; Shields, G. C.; Kisiel, Z.; Pate, B. H. *Science* 2012 *336*(6083), 897–901.

[160] Evangelisti, L.; Perez, C.; Lobsiger, S.; Seifert, N. A.; Zaleski, D. P.; Pate, B.; Kisiel, Z.; Temelso, B.; Shields, G. C. Deuterated water hexamer observed by chirped-pulse rotational spectroscopy. In *69th International Symposium on Molecular Spectroscopy*; : Urbana, IL, 2014.

[161] Hammer, N. I.; Diken, E. G.; Roscioli, J. R.; Johnson, M. A.; Myshakin, E. M.; Jordan, K. D.; McCoy, A. B.; Huang, X.; Bowman, J. M.; Carter, S. *J. Chem. Phys.* 2005 *122*(24), 244301.

[162] Nielsen, H. H. *Phys. Rev.* 1941 *60*(11), 794–810.

[163] Barone, V. *J. Chem. Phys.* 2004 *120*(7), 3059–3065.

[164] McCoy, A. B.; Guasco, T. L.; Leavitt, C. M.; Olesen, S. G.; Johnson, M. A. *Phys. Chem. Chem. Phys.* 2012 *14*, 7205–7214.

[165] Scott, A. P.; Radom, L. *J. Chem. Phys.* 1996 *100*(41), 16502–16513.

[166] Alecu, I. M.; Zheng, J.; Zhao, Y.; Truhlar, D. G. *J. Chem. Theory. Comput.* 2010 *6*(9), 2872–2887 PMID: 26616087.

[167] Samala, N. R.; Agmon, N. *Chem. Phys.* 2018 *514*, 164–175 Energy and Entropy of Change: From Elementary Processes to Biology.

[168] Vendrell, O.; Gatti, F.; Lauvergnat, D.; Meyer, H.-D. *J. Chem. Phys.* 2007 *127*(18), 184302.

[169] Vendrell, O.; Gatti, F.; Meyer, H.-D. *Agew. Chem. Int. Ed.* 2009 *48*, 352–355.

[170] Schröder, M.; Gatti, F.; Lauvergnat, D.; Meyer, H.-D.; Vendrell, O. *Nature Communications* 2022 *13*(1), 6170.

[171] Yang, N.; Duong, C. H.; Kelleher, P. J.; McCoy, A. B.; Johnson, M. A. *Science* 2019 *364*(6437), 275–278.

[172] Fetherolf, J. H.; Pavošević, F.; Tao, Z.; Hammes-Schiffer, S. *J. Phys. Chem. Lett.* 2022 *13*(24), 5563–5570 PMID: 35696537.

[173] Eigen, M. *Angew. Chem., Int. Ed. Engl.* 1964 *3*(1), 1–19.

[174] Zundel, G. *Adv. Chem. Phys.* 1999 *111*, 1–217.

[175] Fournier, J. A.; Wolke, C. T.; Johnson, C. J.; Johnson, M. A.; Heine, N.; Gewinner, S.; Schöllkopf, W.; Esser, T. K.; Fagiani, M. R.; Knorke, H.; Asmis, K. R. *Proceedings of the National Academy of Sciences* 2014 *111*(51), 18132–18137.

[176] Buchanan, E. G.; Dean, J. C.; Zwier, T. S.; Sibert, E. L. *J. Chem. Phys.* 2013 *138*(6), 064308/1–11.

[177] Tabor, D. P.; Hewett, D. M.; Bocklitz, S.; Korn, J. A.; Tomaine, A. J.; Ghosh, A. K.; Zwier, T. S.; Sibert, E. L. *J. Chem. Phys.* 2016 *144*, 224310.

[178] Yu, Q.; Carpenter, W. B.; Lewis, N. H. C.; Tokmakoff, A.; Bowman, J. M. *J. Phys. Chem. B* 2019 *123*(33), 7214–7224.

[179] Anderson, J. R. *The Architecture of Cognition;* Erlbaum: Mahwah, NJ, Repr. originally publ. 1983 ed.; 1996.

[180] LaBerge, D.; Samuels, S. J. *Cogn. Psychol.* 1974 *6*(2), 293–323.

[181] LaBerge, D. *The Brain's Art of Mindfulness;* Harvard University Press: Cambridge, MA and London, England, 1995.

[182] Cziksentmihalyi, M. *Flow – The Psychology of optimal experience;* Harper and Row: New York, 1st ed.; 1990.

[183] Toledo, S.; Dubas, J. M. *J. Chem. Educ.* 2016 *93*(1), 64–69.

[184] Toledo, S.; Dubas, J. M. *J. Chem. Educ.* 2017 *94*(8), 1043–1050.

[185] States, N.; Stone, E.; Cole, R. *Educ. Sci.* 2023 *13*(2), 192.

[186] Garrett-Roe, S. An implementation of mastery-based grading based on Marzano's Taxonomy in large-enrollment general chemistry. In *Biennial Conference on Chemical Education*; West Lafayette, IN, 2022.

[187] Teodorescu, R. E.; Bennhold, C.; Feldman, G.; Medsker, L. *Phys. Rev. ST Phys. Educ. Res.* 2013 *9*(1), 010103.

[188] Dubas, J. M.; Toledo, S. A. *Int. Rev. Econ. Educ.* 2016 *21*, 12–20.

[189] Wolcott, M. D.; Lobczowski, N. G. *Curr. Pharm. Teach. Learn.* 2021 *13*(2), 181–188.

[190] Hennink, M. M.; Kaiser, B. N.; Marconi, V. C. *Qual Health Res* 2017 *27*(4), 591–608.