Electronic Structure and Vibrational Spectroscopy in Hydrogen Bonded Clusters: Application to Water

Long H. Nguyen

A thesis submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN APPLIED CHEMICAL SCIENCE AND TECHNOLOGY

University of Washington

2026

Reading Committee:

Sotiris S. Xantheas, Chair

Thom H. Dunning, Jr.

Program Authorized to Offer Degree:
Chemistry

University of Washington

Abstract

Electronic Structure and Vibrational Spectroscopy in Hydrogen Bonded Clusters: Application to Water

Long H. Nguyen

Chair of the Supervisory Committee:
Dr. Sotiris S. Xantheas
Department of Chemistry

Hydrogen bonding is undoubtedly one of the most important interactions in nature that affects molecular structure, charge distribution, energetics, and encodes spectroscopic signatures in vibrational spectra. However, achieving theoretical descriptions that are simultaneously accurate, chemically interpretable, and computationally efficient remains challenging. This thesis addresses some of these issues in hydrogen bonded clusters, with particular emphasis on water as a prototypical system whose structural and spectroscopic behavior is highly sensitive to its local environment.

The first part of this thesis explores how the vibrational properties of a water molecule change as it becomes part of increasingly large hydrogen bonded networks with increasing solvation shells (first, second, and third), simulating environments that mimic liquid water. Using advanced ab initio electronic structure methods, such as second order Møller-Plesset perturbation (MP2) and Coupled Cluster with Singles and Doubles plus a perturbative treatment of connected Triple replacements [CCSD(T)] theories, and comparing them with the results of classical interaction potentials, the analysis focuses on both the harmonic and anharmonic frequencies of the bending and stretching O–H vibrations. The H–O–H bending mode quickly shifts to higher frequencies (blue shift) with increasing solvation, rising from

1596 cm^{-1} in the monomer to the 1653–1664 cm^{-1} in the first solvation shell, and stays within the observed infrared range for liquid water as more shells are added. In contrast, the OH stretching frequencies experience significant red shifts upon solvation, which depend strongly on the structure of the local hydrogen bond network. Second shell tetrahedral networks show large red shifts due to cooperative hydrogen bonding, while third shell networks produce shifts that are typically found in bulk water. The bending vibration is mainly influenced by the immediate (nearest neighbor) hydrogen bond geometry and stabilizes quickly in tetrahedral environments, whereas the stretching modes reflect the broader distribution of hydrogen bond strengths that emerge in larger (beyond nearest neighbor) solvation shells. Overall, these findings deliver a systematic microscopic picture of how waters vibrational spectrum transitions from an isolated molecule to a bulk-like environment, highlighting the importance of including an environment-dependent 1-body term in classical many-body polarizable potentials for water.

The reliability of quantum mechanical methodologies is essential for modeling hydrogen bonding interactions. The valence Complete Active Space Self-Consistent Field (vCASSCF or vCAS) method is evaluated for its ability to reproduce the spectroscopic signatures associated with the formation of a hydrogen bond in hydrogen the bonded dimers of $(H_2O)_2$, $(HF)_2$, and HF–H_2O. The vCAS method predicts donor X–H bond contractions and corresponding vibrational blue shifts relative to the isolated monomers, which contradict both experimental observations and the results obtained with higher-level methods. Analysis of the active orbitals indicates that the failure of the vCAS method originates from an incorrect assignment of the orbitals to the active space, with orbitals associated with the oxygen lone pairs being included in the active space rather than orbitals associated with the O–H bonds. This leaves the donor bond without the leftright correlation required to describe bond elongation upon hydrogen bond formation. This has significant implications for subsequent treatments of dynamical correlation based on a reference vCAS wavefunction. In contrast, Hartree-Fock

(HF), MP2 and CCSD(T) provide a physically consistent description of hydrogen bonding for these systems, correctly reproducing the donor X–H bond elongation and the associated vibrational red shifts.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

GLOSSARY

ab initio: A quantum chemical approach in which molecular properties are computed directly from the electronic Schrödinger equation without empirical fitting to the property of interest.

active space: The subset of molecular orbitals and electrons treated explicitly in a multiconfigurational calculation such as CASSCF.

anharmonicity: Deviation of a molecular vibrational potential from the ideal quadratic form of the harmonic oscillator.

Born–Oppenheimer approximation: The approximation in which electronic and nuclear motions are separated by treating the nuclei as fixed while solving the electronic problem.

CASSCF: Complete active space self-consistent field. A multiconfigurational method in which all configurations arising from a chosen active space are included and both orbitals and configuration coefficients are optimized.

CBS limit: Complete Basis Set limit. The value approached as the one-electron basis set is systematically enlarged toward completeness.

CCSD(T): Coupled-cluster theory with singles and doubles plus a perturbative treatment of connected triples; commonly regarded as the benchmark or "gold standard" of modern quantum chemistry.

CHELPG: Charges from electrostatic potentials using a grid-based method. A scheme for deriving atomic point charges by fitting to the molecular electrostatic potential.

CM-DVR: Colbert–Miller discrete variable representation. A sinc-DVR approach that provides an efficient grid-based representation of vibrational Hamiltonians.

correlation energy: The difference between the exact nonrelativistic electronic energy and the Hartree–Fock energy in the complete basis limit.

DMS: Dipole Moment Surface. The molecular dipole moment represented as a function of nuclear geometry.

DVR: Discrete Variable Representation. A variational grid-based method for solving the vibrational Schrödinger equation.

dynamical correlation: Electron correlation arising from the instantaneous avoidance of electrons beyond the mean-field Hartree–Fock description.

Fermi resonance: Coupling between a fundamental vibration and a nearby overtone or combination band, often important in the vibrational spectroscopy of water.

Hartree–Fock (HF): A mean-field electronic structure method in which the wavefunction is approximated by a single Slater determinant.

harmonic approximation: The approximation in which the potential energy surface is expanded only to second order about the equilibrium geometry.

hydrogen bond: A directional intermolecular interaction involving a hydrogen atom bonded to one atom and interacting attractively with an acceptor site on another atom or molecule.

IR: Infrared.

local mode: A vibrational description in which motion is associated primarily with a specific bond or internal coordinate rather than a fully delocalized normal mode.

many-body expansion (MBE): Decomposition of a property of an N-molecule system into 1-body (monomers), 2-body (dimers), 3-body (trimers), and higher-body (n-mers) contributions.

MCSCF: Multi Configurational Self-Consistent Field. A class of methods that optimize both the orbitals and the configuration interaction coefficients simultaneously.

MP2: Second-order Møller–Plesset perturbation theory. A post-Hartree–Fock method that provides the leading perturbative correction for dynamical correlation.

normal mode: An independent vibrational coordinate obtained by diagonalizing the mass-weighted Hessian or, equivalently, the Wilson **GF** matrix problem.

PES: Potential Energy Surface. The electronic energy represented as a function of nuclear coordinates.

Pipek–Mezey localization: A unitary orbital localization procedure that produces orbitals resembling chemically intuitive bonds and lone pairs.

PO–DVR: Potential-Optimized Discrete Variable Representation. A DVR approach in which the grid is adapted to the shape of a reference potential.

polarizable potential: A classical interaction model in which the molecular charge distribution can respond to the local environment, typically through inducible dipoles.

PS PES: Partridge–Schwenke Potential Energy Surface. A near-spectroscopic potential for the gas phase water monomer.

QTAIM: Quantum Theory of Atoms In Molecules. A topological analysis of electron density used to define atoms, bonds, and related properties.

RHF: Restricted Hartree–Fock. A closed-shell Hartree–Fock method in which each spatial orbital is doubly occupied.

solvation shell: A shell of surrounding molecules around a central molecule, often classified as first, second, or higher shells according to proximity and coordination.

static correlation: Electron correlation arising when more than one electronic configuration is needed for a qualitatively correct zeroth-order description.

TTM: Thole-Type Model. A family of polarizable water potentials based on inducible dipoles and Thole damping.

vCAS: Valence Complete Active Space. A CASSCF active space constructed from the valence orbitals of the constituent atoms.

VPT2: Second-order Vibrational Perturbation Theory. A perturbative treatment of anharmonic vibrational effects based on cubic and quartic force constants.

1-body term: In the many-body expansion, the intramolecular distortion energy of an isolated monomer as a function of its internal coordinates.

2-body term: In the many-body expansion, the pairwise interaction energy between two monomers after subtraction of the corresponding 1-body contributions.

3-body term: In the many-body expansion, the nonadditive interaction energy among three monomers after subtraction of all 1-body distortion energies and all pairwise 2-body contributions.

ACKNOWLEDGMENTS

I would like to thank Daniel Schofield and Karisa Pierce from Seattle Pacific University for their guidance during my undergraduate years, especially throughout the COVID period. You both played a major role in assisting me to grow into the student I have become, and I will always appreciate having had you as my advisors during such a strange and difficult time. I would also like to thank Thom Dunning Jr. for guiding me early in my academic path, even before I had taken quantum chemistry, and for his continued mentorship during my masters program.

Special thanks go to the Xantheas Group for the collegial environment they created. Garrett and Max, thank you for answering ALL my questions, teaching me, and helping me correct my mistakes. Anna, having had you as my classmate, has made my first year at UW bearable. Madisen, thank you for your kindness, your care, and for always holding the group together. I would also like to thank the McCoy Group for having me in their group meetings and for always paying close attention to the smallest details of my projects.

Thank you to Sharon Durfy, my MSc. advisor, for your constant encouragement throughout this process. Your support enabled me to focus on science and achieve my goals.

Finally, I would like to express my deepest gratitude to my advisor, Sotiris Xantheas. At a pivotal moment in my academic journey, you gave me a tremendous opportunity to continue pursuing my passion for research. I look forward to continuing my training under your mentorship as I begin the next stage in the group.

DEDICATION

to Minzy, who provided invaluable emotional support and never let me take myself too seriously. Thank you for your endless cups of coffee, your tolerance for my erratic schedule, and for always reminding me that there is life outside the laboratory.

to my sister, my parents, especially my grandparents, for their unconditional love, for always believing in my potential, and for the countless sacrifices you made to allow me to pursue my education.

Chapter 1

INTRODUCTION

Although water is a small molecule, its hydrogen bonded network has far-reaching implications on chemistry, biology, and the physical environment. Its hydrogen bonded network controls condensed phase structure, thermodynamics, and dynamics. The IUPAC definition of a hydrogen bond emphasizes a directional interaction involving a hydrogen atom bound to a more electronegative atom and interacting with an acceptor site, supported by characteristic spectroscopic and structural signatures [1]. Because the O–H stretch is especially sensitive to local electrostatics and the hydrogen bond geometry, infrared (IR) spectroscopy has long served as a direct probe of different hydrogen bond environments in water and related systems.[2]

In the gas phase, an isolated H_2O molecule exhibits three fundamental vibrational modes: the symmetric stretch (ν_1), the bend (ν_2), and the antisymmetric stretch (ν_3).[3, 4] Upon the formation of a hydrogen bond, the donor O–H bond is typically weakened and lengthened relative to the free monomer, leading to a red shift and substantial broadening of the O–H stretching band. A useful structural interpretation of this trend is provided by *Badger's rule*, which correlates the stretching force constant k_e with the equilibrium bond length r_e for a wide class of bonds.[5] In its classic form,

$$k_e \, (r_e - d_{ij})^3 \approx A, \tag{1.1}$$

where A is an empirical constant and d_{ij} depends on the position of the bonded atoms (rows) in the periodic table.[5] Together with the harmonic relationship $\omega_e \approx \sqrt{k_e/\mu_{\mathrm{red}}}$, Eq. (1.1) rationalizes why hydrogen bond-induced O–H elongation generally lowers the stretching frequency. More recent studies have extended this structural interpretation beyond the

original empirical form of Badger's rule. Boyer *et al.* showed that, for water clusters, ion–water clusters, and liquid water, hydrogen-bond-induced elongation of the O–H bond is nearly linearly correlated with the corresponding red shift of the O–H stretching frequency.[6] In particular, they found an approximately constant ratio of -19 ± 1 cm^{-1} per 0.001 Å increase in the O–H bond length, providing a quantitative link between local hydrogen bond structure and vibrational frequency shifts.

These structural–spectroscopic relationships are clearly reflected in condensed phase water. At ambient conditions, liquid water exhibits a broad O–H stretching absorption centered in the mid-3000 cm^{-1} region, together with the H–O–H bending band near $\sim$1640 cm^{-1}; recommended optical constants and absorption coefficients have been tabulated with quantified uncertainties.[7] In ice I_h, the O–H stretching region shifts further to lower frequency and develops more structured features associated with the ordered hydrogen bonded network. Early and still influential measurements and assignments for ice I_h/I_c remain standard references for interpreting these spectral signatures.[8, 9] Beyond static hydrogen bond distributions, condensed phase spectra are also shaped by vibrational coupling and anharmonicity. In particular, Fermi resonance between the bending overtone and stretching fundamentals can measurably influence the O–H stretching lineshape and shoulders in both clusters and liquid water.[4]

These spectroscopic observations motivate a central methodological theme of this thesis: accurate simulation of water vibrational spectra requires a realistic potential energy surface (PES) for nuclear motion, a correspondingly accurate dipole moment surface (DMS) for intensities, and an anharmonic treatment beyond the simplest harmonic model.[2, 10] In the most rigorous formulation, the PES is identified with the electronic energy obtained within the Born–Oppenheimer approximation and is therefore an *ab initio* surface. In practice, however, evaluating the electronic energy directly at every geometry is rarely feasible for systems requiring extensive configurational sampling. This limitation has motivated a hierarchy of approximate potential energy functions, ranging from simple empirical force fields to highly accurate many-body representations fitted to large electronic structure datasets.

For water, the quality of this potential energy function is especially important because the balance between intramolecular distortion, hydrogen bonding, many-body polarization, and anharmonic vibrational coupling is unusually delicate. A model may reproduce selected bulk observables reasonably well and yet fail for gas phase clusters, interfacial environments, or vibrational spectroscopy. It is therefore useful to distinguish among the major classes of water potentials according to how they approximate the underlying Born–Oppenheimer surface.

The simplest and most widely used potentials for atomistic simulation are molecular mechanics (MM) force fields. In these models, electronic degrees of freedom are not treated explicitly. Instead, the energy is represented by analytic functions of the nuclear coordinates, with parameters chosen to reproduce selected macroscopic structural, energetic, or thermodynamic properties. The detailed intramolecular functional form is strongly model dependent: rigid water models eliminate intramolecular terms by fixing the monomer geometry, whereas flexible models include bond-stretching, angle-bending, and, when needed, torsional terms. For this reason, the common structure of these models is most clearly illustrated through their non-bonded interactions rather than through a single generic force field expression.

One of the most widely used nonbonded terms is the Lennard–Jones (12-6) potential,[11]

$$V_{\mathrm{LJ}}(r_{ij}) = 4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right], \tag{1.2}$$

where r_{ij} is the intermolecular separation between sites i and j, ϵ_{ij} sets the well depth, and σ_{ij} defines the effective collision diameter. The repulsive r^{-12} term empirically represents short-range Pauli repulsion, whereas the attractive r^{-6} term mimics London dispersion. In most classical water models, this van der Waals term is combined with Coulomb interactions between fixed partial charges located on either atomic or off-atomic sites.

The appeal of such models lies in their computational efficiency. By replacing the explicit electronic problem with parameterized analytic functions, they enable simulations of large systems and long time scales that would be inaccessible to direct *ab initio* methods. This efficiency, however, comes at the cost of physical realism. Traditional rigid fixed-charge models describe intermolecular interactions through effective pairwise terms and therefore cannot

respond explicitly to changes in the local electronic environment. For water, where hydrogen bonding strongly modifies molecular dipoles, bond lengths, and vibrational frequencies, this limitation can become severe.

Earlier and still widely used water force fields assume that the total interaction energy can be decomposed into pairwise-additive contributions between water molecules. Representative examples include SPC,[12] SPC/E,[13] TIP3P and TIP4P,[14] later refinements such as TIP4P/2005[15] and TIP4P/Ice,[16] TIP5P,[17] and the more recent OPC models.[18] In these models, each water molecule is represented as a rigid body with a small number of interaction sites carrying partial charges and, typically, a Lennard–Jones center fixed to the molecular geometry. The total energy is then obtained as the sum of effective two-body interactions over all molecular pairs. By construction, genuine many-body contributions are absent and are instead folded into the fitted two-body parameters.[19]

This strategy yields extremely efficient models and can reproduce selected liquid state observables surprisingly well. Nevertheless, their transferability across different environments is inherently limited. Hydrogen bonding changes the charge distribution, dipole moment, bond lengths, and vibrational frequencies of each molecule in an environment-dependent manner. A rigid pairwise model can only mimic these effects indirectly through fitted parameters. As a result, properties that span gas phase clusters, liquid water, interfaces, supercooled phases, and spectroscopy are difficult to reproduce simultaneously within a single pairwise-additive framework.

Achieving predictive accuracy across different thermodynamic conditions and chemical environments requires molecular models that properly capture the many-body character of intermolecular interactions. In the many-body expansion (MBE) formalism,[19] the total potential energy of a system composed of N monomers is expressed as a sum of one-body, two-body, three-body, and higher-order terms:

$$V_N = \sum_i V^{(1\mathrm{B})}(i) + \sum_{i<j} V^{(2\mathrm{B})}(i,j) + \sum_{i<j<k} V^{(3\mathrm{B})}(i,j,k) + \cdots , \tag{1.3}$$

In this framework, the one-body term describes the distortion energy of each monomer,

while higher-order terms describe intermolecular interactions of increasing complexity. For water, the many-body expansion is not merely a convenient formal decomposition, it mirrors the underlying physics. Hydrogen bonding changes the internal structure of each monomer, modifies the dipole moment, induces polarization, and generates nonadditive interaction energies that propagate through the network.

For water, the intramolecular one-body term is especially important because it governs the monomer geometry, harmonic frequencies, anharmonicity, and the response of the O–H stretches and H–O–H bend to hydrogen bonding. In the many-body expansion, this one-body term represents the intramolecular PES of an isolated water monomer, while higher-body terms encode intermolecular interactions of increasing order. For the gas-phase water monomer, the one-body term is known to near-spectroscopic accuracy from the Partridge–Schwenke PES and its associated DMS.[20, 21]

The PS PES and DMS were obtained by fitting high level electronic structure results, together with experimental information for H_2O, to analytic functional forms that reproduce the rovibrational spectrum of the monomer with near-spectroscopic accuracy. As a result, the PS representation yields highly accurate equilibrium geometries, vibrational fundamentals, and overtones for gas phase water and has become the *de facto* standard one-body term in many modern water potentials.[22, 23, 24, 25] This role is particularly important for spectroscopy: a potential may reproduce cluster binding energies reasonably well and still yield inaccurate vibrational frequencies if the monomer PES is not described correctly. Conversely, an accurate one-body surface alone is insufficient unless the intermolecular terms also capture hydrogen bonding, polarization, and many-body effects.

The realization of the limitations of pairwise additive models motivated the development of advanced force fields in which the molecular charge distribution responds explicitly to the local environment. One important class comprises of polarizable force fields, which introduce inducible dipoles or higher multipoles on each water molecule and thereby generate many-body interaction energy through mutual polarization. A particularly influential family is the set of Thole-type models (TTM),[26, 27, 22, 23, 24, 28] which were designed to approximate the

Born–Oppenheimer PES while remaining efficient enough for condensed phase simulations that include nuclear quantum dynamical effects.[29, 30]

A common formulation employs inducible point dipoles located on atoms or auxiliary sites. If two polarizable sites interact too closely, however, the dipole–dipole interaction becomes singular at short range. This pathology is removed using the Thole damping,[31] in which the dipole interaction tensor is smeared at short separations. In schematic form, the damped interaction tensor may be written as

$$\mathbf{T}_{ij} = \frac{f_e}{r_{ij}^3}\mathbf{I} - \frac{3f_t}{r_{ij}^5}\mathbf{r}_{ij}\mathbf{r}_{ij}^{\mathrm{T}}, \tag{1.4}$$

where f_e and f_t are damping functions and $\mathbf{r}_{ij}$ is the vector connecting sites i and j. This modification preserves the long-range dipolar interaction while eliminating the unphysical short-range divergence.

The TTM family developed by Xantheas and co-workers progressively refined the treatment of water's intramolecular flexibility and intermolecular induction. TTM2-F[22] employs a flexible monomer with geometry-dependent partial charges taken from the spectroscopically accurate Partridge–Schwenke monomer representation,[20] together with three atom-centered inducible dipoles coupled self-consistently through Thole damping. TTM2.1-F[23] revisited the parameterization to remove short-range pathologies encountered in condensed-phase simulations by improving the molecular dipoles in bulk-like environments. TTM3-F[24] shifted the emphasis toward vibrational spectroscopy by collapsing the distributed polarizability to a single inducible dipole at an off-nuclear M-site along the H–O–H bisector and introducing an effective monomer DMS, leading to an improved description of infrared band positions and intensities for clusters and liquid water. TTM4-F[28] further refined the polarizability surface and short-range damping, improving dielectric, structural, and spectroscopic properties. The Thole-type induction scheme has since been adopted as the many-body polarization backbone in several modern first-principles many-body water potentials.[19]

A complementary strategy is to include explicit many-body interaction terms fitted directly to *ab initio* calculations on small clusters. One of the most influential developments

in this direction is the use of permutationally invariant polynomials (PIPs), introduced and extensively developed by Bowman and coworkers.[32] In this approach, the fitted functional form is constructed so that the energy is invariant under permutation of equivalent atoms, as required by molecular symmetry. This makes it possible to fit high-dimensional PESs while preserving exact translational, rotational, and permutational symmetries.

Early accurate water potentials constructed in this spirit include WHBB[33] and CC-pol,[34] which use explicit low-order many-body fits to *ab initio* dimer and trimer data. Two of the most advanced potentials in this category are MB-pol[35, 10] and q-AQUA.[36] MB-pol combines high-fidelity two-body and three-body terms fitted to coupled-cluster reference data with a many-body polarization term derived from the TTM framework, which describes long-range two- and three-body interactions as well as all higher-order contributions. The result is a potential that attains near-chemical accuracy for interaction energies and reproduces structural, thermodynamic, and spectroscopic properties of water from gas-phase clusters to the liquid. q-AQUA extends this philosophy by including explicit many-body polynomial expansions through the four-body term, reflecting the fact that certain properties, especially forces, can benefit from contributions beyond three-body.[36] This additional flexibility can improve transferability across cluster sizes and thermodynamic state points, although at greater complexity and computational expense.

The hierarchy of potential energy functions described above spans a wide range of accuracy and computational cost. Classical molecular mechanics force fields are extremely efficient and remain indispensable for large scale simulations, but their rigid, pairwise additive character limits their reliability for spectroscopy and for quantum mechanical aspects of hydrogen bonded systems. Advanced polarizable force fields improve the description of induction and environmental response, yet still approximate short range quantum effects. High accuracy many-body potentials, especially those built from accurate one-body surfaces such as Partridge–Schwenke and explicit low order many-body terms, offer a route to near-*ab initio* accuracy across clusters, liquid water, and ice, albeit at greater computational cost and significantly more elaborate parameterization.

In this thesis, these different classes of potentials are discussed not simply as alternative simulation tools, but as different approximations to the same underlying Born–Oppenheimer surface. Their comparison is particularly revealing for water, where the interplay among monomer flexibility, hydrogen bonding, many-body polarization, and anharmonic vibrational structure is directly reflected in the predicted spectra. By constructing hierarchical cluster networks that mimic progressively larger solvation shells around a central water molecule, we examine how the vibrational bands and dipole moment surface of that molecule evolve with increasing solvation. This framework also allows us to re-evaluate the requirements placed on the one-body term in many-body-expansion-based classical potentials, especially when gas phase monomer accuracy is transferred to strongly hydrogen bonded and heterogeneous environments.

Chapter 2

THEORETICAL BACKGROUND

Quantum mechanics provides the most fundamental theoretical framework currently available for describing matter at microscopic length scales. At this level, molecular structure, chemical bonding, energetics, and spectroscopy arise from the interactions among electrons and nuclei. For hydrogen bonded systems, these interactions determine not only the strength and geometry of the hydrogen bond, but also the changes in charge distribution, dipole moment, and vibrational frequencies that accompany hydrogen bond formation. The total energy of a molecule can be quantitatively determined from the molecular Hamiltonian and obtained by solving the Schrödinger equation. The main focus of this thesis is the equilibrium state of the molecule, i.e., the observables are the eigenvalues of the Hamiltonian, which are constant over time. The time-independent Schrödinger equation is

$$\hat{H}\psi = E\psi, \tag{2.1}$$

where $\hat{H}$ is the Hamiltonian operator, ψ is the wavefunction, and E is the total energy of the system.

2.1 The Born–Oppenheimer electronic Hamiltonian and many-electron wavefunctions

In Cartesian coordinates, the full non-relativistic molecular Hamiltonian in SI units is written as

$$\begin{aligned}\hat{H} = &-\sum_{m=1}^{N_e}\frac{\hbar^2}{2m_e}\nabla_m^2 - \sum_{A=1}^{N_n}\frac{\hbar^2}{2M_A}\nabla_A^2 \\ &- \frac{1}{4\pi\varepsilon_0}\sum_{m=1}^{N_e}\sum_{A=1}^{N_n}\frac{Z_A e^2}{r_{mA}} + \frac{1}{4\pi\varepsilon_0}\sum_{m<n}^{N_e}\frac{e^2}{r_{mn}} + \frac{1}{4\pi\varepsilon_0}\sum_{A<B}^{N_n}\frac{Z_A Z_B e^2}{R_{AB}}.\end{aligned} \tag{2.2}$$

Here, N_e and N_n are the numbers of electrons and nuclei, respectively; $\mathbf{r}_m$ is the Cartesian position vector of electron m, and $\mathbf{R}_A$ is the Cartesian position vector of nucleus A. The quantities

$$r_{mA} = |\mathbf{r}_m - \mathbf{R}_A|, \qquad r_{mn} = |\mathbf{r}_m - \mathbf{r}_n|, \qquad R_{AB} = |\mathbf{R}_A - \mathbf{R}_B|$$

are the electron–nucleus, electron–electron, and nucleus–nucleus distances, respectively. The symbols m_e and M_A denote the electron mass and the mass of nucleus A; Z_A is the atomic number of nucleus A, so that its charge is $+Z_A e$; e is the elementary charge; $\hbar$ is the reduced Planck constant; and ε_0 is the vacuum permittivity. The operators ∇_m^2 and ∇_A^2 are Laplacians with respect to the electronic coordinate $\mathbf{r}_m$ and nuclear coordinate $\mathbf{R}_A$, respectively.

The first two terms in Eq. (2.2) are the electronic and nuclear kinetic energy operators, respectively. The remaining three terms describe the Coulomb attraction between electrons and nuclei, the Coulomb repulsion between electrons, and the Coulomb repulsion between nuclei. In atomic units, $\hbar = m_e = e = 4\pi\varepsilon_0 = 1$, and all masses are expressed relative to the electron mass. Under the Born–Oppenheimer approximation, the nuclei are held fixed at a geometry $\mathbf{R}$, so the nuclear kinetic energy operator is omitted from the electronic problem. The electronic Hamiltonian is therefore

$$\hat{H}_e(\mathbf{R}) = -\sum_{m=1}^{N_e} \frac{1}{2}\nabla_m^2 - \sum_{m=1}^{N_e}\sum_{A=1}^{N_n} \frac{Z_A}{r_{mA}} + \sum_{m<n}^{N_e} \frac{1}{r_{mn}} + \sum_{A<B}^{N_n} \frac{Z_A Z_B}{R_{AB}}. \tag{2.3}$$

It is often convenient to write Eq. (2.3) in the reduced form

$$\hat{H}_e(\mathbf{R}) = \sum_{m=1}^{N_e} \hat{h}(m) + \sum_{m<n}^{N_e} \frac{1}{r_{mn}} + V_{NN}(\mathbf{R}), \tag{2.4}$$

where m and n label electrons, and the one-electron operator is

$$\hat{h}(m) = -\frac{1}{2}\nabla_m^2 - \sum_{A=1}^{N_n} \frac{Z_A}{r_{mA}}. \tag{2.5}$$

The nuclear repulsion term is

$$V_{NN}(\mathbf{R}) = \sum_{A<B}^{N_n} \frac{Z_A Z_B}{R_{AB}}. \tag{2.6}$$

Because $V_{NN}(\mathbf{R})$ is constant for fixed nuclear geometry, therefore the reduced electronic Hamiltonian is defined as

$$\hat{H}_e^{\mathrm{red}}(\mathbf{R}) = \sum_{m=1}^{N_e} \hat{h}(m) + \sum_{m<n}^{N_e} \frac{1}{r_{mn}}, \tag{2.7}$$

and then adds $V_{NN}(\mathbf{R})$ to obtain the total Born–Oppenheimer electronic energy.

The associated electronic Schrödinger equation is

$$\hat{H}_e(\mathbf{R})\, \psi_{e,s}(1, 2, \ldots, N_e; \mathbf{R}) = E_{e,s}(\mathbf{R})\, \psi_{e,s}(1, 2, \ldots, N_e; \mathbf{R}), \tag{2.8}$$

where $\psi_{e,s}$ is the electronic wavefunction for electronic state s at fixed nuclear geometry $\mathbf{R}$. Here $1, 2, \ldots, N_e$ denote the combined spatial and spin coordinates of the electrons; explicitly,

$$m \equiv (\mathbf{r}_m, \sigma_m),$$

where $\mathbf{r}_m$ is the spatial coordinate of electron m and σ_m is its spin coordinate. The electronic eigenvalue $E_{e,s}(\mathbf{R})$ defines an adiabatic potential energy surface on which the nuclei move.

In the Born–Oppenheimer approximation,[37] the total molecular wavefunction Ψ_{tot} is approximated as a product of an electronic wavefunction and a nuclear wavefunction,

$$\Psi_{\mathrm{tot}}(1, 2, \ldots, N_e, \mathbf{R}) \approx \psi_{e,s}(1, 2, \ldots, N_e; \mathbf{R})\, \chi_{n,s\nu}(\mathbf{R}), \tag{2.9}$$

where $\psi_{e,s}$ is the many-electron wavefunction for electronic state s, and $\chi_{n,s\nu}$ is the nuclear wavefunction associated with vibrational-rotational state ν on electronic surface s.

The corresponding nuclear motion equation on electronic surface s is

$$\left[\hat{T}_n + E_{e,s}(\mathbf{R})\right] \chi_{n,s\nu}(\mathbf{R}) = E_{s\nu} \chi_{n,s\nu}(\mathbf{R}), \tag{2.10}$$

where $\hat{T}_n$ is the nuclear kinetic energy operator, $E_{e,s}(\mathbf{R})$ is the Born–Oppenheimer potential energy surface, and $E_{s\nu}$ is the total energy of nuclear state ν on electronic surface s.

Because electrons are indistinguishable particles obeying Fermi–Dirac statistics (fermions), the exact electronic wavefunction must be antisymmetric with respect to exchange of any two electrons. For a two-electron system,

$$\psi_e(1, 2) = -\psi_e(2, 1), \tag{2.11}$$

and, more generally, interchange of any two electron labels changes the sign of the electronic wavefunction.[38, 39] For N_e electrons,

$$\Psi_I(1,2,\ldots,N_e) = \frac{1}{\sqrt{N_e!}} \begin{vmatrix} \psi_{p_1}(1) & \psi_{p_2}(1) & \cdots & \psi_{p_{N_e}}(1) \\ \psi_{p_1}(2) & \psi_{p_2}(2) & \cdots & \psi_{p_{N_e}}(2) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_{p_1}(N_e) & \psi_{p_2}(N_e) & \cdots & \psi_{p_{N_e}}(N_e) \end{vmatrix}, \tag{2.12}$$

where Ψ_I is a Slater determinant corresponding to configuration I, and $\{\psi_p\}$ are one-electron spin-orbitals. The labels $p_1, p_2, \ldots, p_{N_e}$ specify which spin-orbitals are occupied in determinant I.

A single Slater determinant is only one possible antisymmetric wavefunction. In general, the exact many-electron wavefunction may be formally written as a linear combination of Slater determinants,

$$\psi_e(1,2,\ldots,N_e;\mathbf{R}) = \sum_I C_I(\mathbf{R})\,\Psi_I(1,2,\ldots,N_e;\mathbf{R}), \tag{2.13}$$

where Ψ_I denotes a Slater determinant and $C_I(\mathbf{R})$ is its corresponding expansion coefficient. This is the general form of a many-electron wavefunction, and practical electronic structure methods differ in how this expansion is approximated.

2.2 The Hartree–Fock approximation

The Hartree–Fock (HF) approximation assumes that the electronic wavefunction can be represented by a single Slater determinant and determines the optimal one-electron orbitals by minimizing the expectation value of the electronic Hamiltonian subject to orthonormality constraints. In this way, each electron moves in the average field created by all of the others. Historically, this mean-field description emerged from Hartree's self-consistent-field picture and Fock's incorporation of exchange through antisymmetrization.[40, 41]

Applying the variational principle to a single determinant wavefunction leads to the Hartree–Fock equations,

$$\hat{f}(1)\psi_p^{\mathrm{HF}}(1) = \varepsilon_p \psi_p^{\mathrm{HF}}(1), \tag{2.14}$$

where $\hat{f}$ is the Fock operator, ψ_p^{HF} is a canonical HF spin-orbital, ε_p is its corresponding orbital energy, and p labels a general spin-orbital. Here 1 denotes the combined spatial and spin coordinate of electron 1.

For a closed-shell restricted Hartree–Fock (RHF) treatment, it is convenient to work in terms of doubly occupied spatial orbitals. In that case, the Fock operator acting on a spatial orbital is

$$\hat{f}(\mathbf{r}_1) = \hat{h}(\mathbf{r}_1) + \sum_{j}^{\mathrm{occ}} \left[2\hat{J}_j(\mathbf{r}_1) - \hat{K}_j(\mathbf{r}_1) \right], \tag{2.15}$$

where the sum runs over occupied spatial orbitals. Here $\hat{h}(\mathbf{r}_1)$ is the one-electron core Hamiltonian, and $\hat{J}_j$ and $\hat{K}_j$ are the Coulomb and exchange operators associated with occupied spatial orbital ψ_j^{sp}, respectively. The Coulomb operator represents the classical average electron–electron repulsion, whereas the exchange operator is purely quantum mechanical and arises from the antisymmetry of the wavefunction.

The action of these operators on a spatial orbital ψ_p^{sp} is

$$\hat{J}_j(\mathbf{r}_1)\psi_p^{\mathrm{sp}}(\mathbf{r}_1) = \left[\int \frac{|\psi_j^{\mathrm{sp}}(\mathbf{r}_2)|^2}{r_{12}}\, d\mathbf{r}_2 \right] \psi_p^{\mathrm{sp}}(\mathbf{r}_1), \tag{2.16}$$

and

$$\hat{K}_j(\mathbf{r}_1)\psi_p^{\mathrm{sp}}(\mathbf{r}_1) = \left[\int \frac{\psi_j^{\mathrm{sp}*}(\mathbf{r}_2)\psi_p^{\mathrm{sp}}(\mathbf{r}_2)}{r_{12}}\, d\mathbf{r}_2 \right] \psi_j^{\mathrm{sp}}(\mathbf{r}_1), \tag{2.17}$$

where $r_{12} = |\mathbf{r}_1 - \mathbf{r}_2|$.

A key feature of Eq. (2.15) is that the Fock operator depends on the Hartree–Fock orbitals, while those orbitals are themselves solutions of the Hartree–Fock equations. The Hartree–Fock equations therefore cannot be solved directly and must instead be solved iteratively. This procedure is known as the self-consistent-field (SCF) method: one begins with an initial guess for the orbitals, constructs the corresponding Fock operator, solves for a set of new orbitals, and repeats the process until the input and output orbitals, or equivalently the density matrix and total energy, are converged.[40, 41]

In practical calculations, the Hartree–Fock equations are solved in a finite basis of atom-

centered functions. A spatial molecular orbital is expanded as

$$\psi_p^{\mathrm{sp}}(\mathbf{r}) = \sum_{\mu=1}^{\mathbb{N}} C_{\mu p}\chi_\mu(\mathbf{r}), \tag{2.18}$$

where $\psi_p^{\mathrm{sp}}(\mathbf{r})$ is spatial molecular orbital p, $\chi_\mu(\mathbf{r})$ is atomic basis function μ, $C_{\mu p}$ is a molecular orbital expansion coefficient, and $\mathbb{N}$ is the number of atomic basis functions.

In this atomic orbital basis, the matrix elements of the closed-shell RHF Fock operator are

$$F_{\mu\nu} = h_{\mu\nu} + \sum_{\lambda=1}^{\mathbb{N}}\sum_{\sigma=1}^{\mathbb{N}} P_{\lambda\sigma}\left[(\mu\nu|\lambda\sigma) - \frac{1}{2}(\mu\lambda|\nu\sigma)\right], \tag{2.19}$$

where $F_{\mu\nu}$ is a Fock-matrix element, $h_{\mu\nu}$ is a one-electron matrix element, $P_{\lambda\sigma}$ is an element of the density matrix, and $(\mu\nu|\lambda\sigma)$ is a two-electron repulsion integral.

The one-electron matrix element is

$$h_{\mu\nu} = \int \chi_\mu^*(\mathbf{r}_1)\,\hat{h}(\mathbf{r}_1)\,\chi_\nu(\mathbf{r}_1)\,d\mathbf{r}_1, \tag{2.20}$$

where

$$\hat{h}(\mathbf{r}_1) = -\frac{1}{2}\nabla_1^2 - \sum_{A=1}^{N_n}\frac{Z_A}{r_{1A}} \tag{2.21}$$

is the one-electron core Hamiltonian in atomic units. Here Z_A is the charge of nucleus A, $r_{1A} = |\mathbf{r}_1 - \mathbf{R}_A|$, and N_n is the number of nuclei.

The two-electron repulsion integral is

$$(\mu\nu|\lambda\sigma) = \iint \chi_\mu^*(\mathbf{r}_1)\chi_\nu(\mathbf{r}_1)\frac{1}{r_{12}}\chi_\lambda^*(\mathbf{r}_2)\chi_\sigma(\mathbf{r}_2)\,d\mathbf{r}_1\,d\mathbf{r}_2, \tag{2.22}$$

where $r_{12} = |\mathbf{r}_1 - \mathbf{r}_2|$. The first pair of indices, μ, ν, refers to functions evaluated at electron coordinate $\mathbf{r}_1$, while the second pair, λ, σ, refers to functions evaluated at electron coordinate $\mathbf{r}_2$.

For a closed-shell system, the density matrix is

$$P_{\lambda\sigma} = 2\sum_{i=1}^{N_e/2} C_{\lambda i}C_{\sigma i}^*, \tag{2.23}$$

where the factor of 2 accounts for double occupation of each spatial molecular orbital, N_e is the number of electrons, and the sum runs over the $N_e/2$ occupied spatial orbitals. For real molecular orbital coefficients, $C^*_{\sigma i} = C_{\sigma i}$.

The Hartree–Fock equations in this basis become the Roothaan–Hall matrix equations,

$$\mathbf{FC} = \mathbf{SC}\boldsymbol{\varepsilon}, \tag{2.24}$$

where $\mathbf{F}$ is the Fock matrix, $\mathbf{C}$ is the molecular-orbital coefficient matrix, $\mathbf{S}$ is the overlap matrix, and $\boldsymbol{\varepsilon}$ is the diagonal matrix of orbital energies.[42, 43]

The overlap matrix elements are

$$S_{\mu\nu} = \int \chi^*_\mu(\mathbf{r})\chi_\nu(\mathbf{r})\, d\mathbf{r}. \tag{2.25}$$

Because the Fock matrix depends on the density matrix, and the density matrix depends on the molecular orbital coefficients obtained from Eq. (2.24), the Hartree–Fock equations must be solved iteratively. This iterative procedure constitutes the self-consistent-field (SCF) method.[40, 41]

Once self-consistency is reached, the closed-shell Hartree–Fock energy is

$$E_{\mathrm{HF}} = V_{NN} + \frac{1}{2}\sum_{\mu=1}^{\mathbb{N}}\sum_{\nu=1}^{\mathbb{N}} P_{\mu\nu}\left(h_{\mu\nu} + F_{\mu\nu}\right), \tag{2.26}$$

where V_{NN} is the nuclear–nuclear repulsion energy.

The construction of the two-electron contribution to the Fock matrix requires four-index electron repulsion integrals over the atomic orbital basis. Therefore, the formal computational scaling of conventional Hartree–Fock theory is $\mathcal{O}(\mathbb{N}^4)$, where $\mathbb{N}$ is the number of atomic basis functions.[40, 41, 42]

The Hartree–Fock theory therefore provides a well-defined variational mean-field reference for molecular electronic structure. It incorporates exchange exactly within a single-determinant picture, but it neglects the instantaneous correlated motion of electrons beyond that of the mean field. For intermolecular interactions, hydrogen bonding, and vibrational spectroscopy, this missing electron correlation is often decisive, and its recovery motivates the post-Hartree–Fock methods discussed in the next section.

2.3 Electron Correlation and Correlated Methods

2.3.1 Electron correlation

The Hartree–Fock theory provides a variational mean-field description of the electronic structure, but it does not account for the instantaneous correlated motion of the electrons. In the Hartree–Fock picture, each electron moves in the average field of all the others, whereas in reality the electrons avoid one another more effectively than this mean-field description allows. The resulting missing stabilization is the *electron correlation energy*, conventionally defined as

$$E_{\mathrm{corr}} = E_{\mathrm{exact}}^{\mathrm{nonrelativistic}} - E_{\mathrm{HF}}^{\mathrm{CBS}}, \tag{2.27}$$

where $E_{\mathrm{exact}}^{\mathrm{nonrelativistic}}$ is the exact nonrelativistic Born–Oppenheimer electronic energy and $E_{\mathrm{HF}}^{\mathrm{CBS}}$ is the Hartree–Fock energy in the complete one-particle basis limit. Because Hartree–Fock is variational, E_{corr} is necessarily negative. In practical calculations, one often speaks analogously of the correlation energy recovered within a finite basis by comparing a correlated method with the corresponding Hartree–Fock result in that same basis.

Electron correlation comprises all many-electron effects missing from the single-determinant Hartree–Fock description. It is therefore useful to distinguish between *static* (or non-dynamical) correlation, which arises when more than one configuration has comparable importance, and *dynamical* correlation, which reflects the short-range instantaneous avoidance of electrons even when a single determinant remains a qualitatively good reference. In hydrogen bonded systems both aspects can matter: donor–acceptor polarization and weak charge transfer can alter the balance of configurations, while accurate binding energies, equilibrium structures, and vibrational shifts require recovery of the missing dynamical correlation.

2.3.2 Configuration Interaction Theory

A formally straightforward route toward incorporating electron correlation is configuration interaction (CI), in which the exact electronic wavefunction in a finite orbital basis is expanded

as a linear combination of Slater determinants or configuration state functions. Such a wavefunction is said to be *multi-configurational* (Eq. 2.13). If all determinants generated from a given orbital space are included, one obtains full CI, which is exact within that one-particle basis. In practice, however, the determinant space grows combinatorially and one is forced to truncate the expansion. Truncated CI models such as CISD, incorporating single and double excitations, recover part of the correlation energy, but they are not size-extensive, so their energetic errors do not scale properly with system size. This limitation motivates the use of more balanced multi-configurational methods for static correlation and more efficient perturbative or coupled-cluster approaches for dynamical correlation.[44]

If the physically important part of the wavefunction requires more than one determinant in the zeroth-order reference itself, then the system is said to have *multi-reference* character. Starting from a single Hartree–Fock reference determinant Ψ_0, the correlated electronic wavefunction can be organized in terms of excitations from the occupied orbitals into the virtual orbitals,

$$\psi_e = c_0\Psi_0 + \sum_{ia} c_i^a \Psi_i^a + \sum_{i<j}\sum_{a<b} c_{ij}^{ab}\Psi_{ij}^{ab} + \sum_{i<j<k}\sum_{a<b<c} c_{ijk}^{abc}\Psi_{ijk}^{abc} + \cdots , \tag{2.28}$$

where Ψ_i^a, Ψ_{ij}^{ab}, and Ψ_{ijk}^{abc} denote singly, doubly, and triply excited determinants, respectively. Here $i, j, k, \ldots$ label occupied orbitals, $a, b, c, \ldots$ label virtual orbitals, and c_0, c_i^a, c_{ij}^{ab}, and c_{ijk}^{abc} are the corresponding expansion coefficients. If one reference determinant dominates, this expansion is naturally viewed as a single-reference wavefunction. In contrast, when several reference determinants must be treated on equal footing, a more appropriate form is

$$\psi_e = \sum_{K\in\mathcal{R}} c_K\Psi_K + \sum_{L\notin\mathcal{R}} c_L\Psi_L, \tag{2.29}$$

where $\mathcal{R}$ denotes a chosen reference space of important determinants, K labels determinants inside the reference space, and L labels determinants outside the reference space. This is the conceptual foundation of multi-reference methods.

2.3.3 Complete Active Space Self–Consistent Field (CASSCF) method

When near-degeneracy of electronic states is important, multi-configurational self-consistent field (MCSCF) methods optimize both the orbitals and the configuration interaction coefficients simultaneously. Among these, the complete active space self-consistent field (CASSCF) method introduced by Ruedenberg, Roos and others remains the standard framework.[45, 46, 47, 48] In CASSCF, a chemically motivated active space is chosen and all configurations arising from distributing the active electrons among the active orbitals are included exactly. Orbitals outside the active space are kept doubly occupied (inactive) or empty (virtual). This provides a balanced treatment of static correlation and yields qualitatively reliable potential energy surfaces in cases where a single Hartree–Fock determinant is insufficient.

For hydrogen bonded dimers, however, the choice of active space is not a technical detail but a decisive physical approximation. As shown in our recent study of the water, hydrogen fluoride, and HF–H_2O dimers, a valence-CASSCF (vCAS) treatment can fail even at the qualitative level for hallmark hydrogen-bond observables.[49] Specifically, the vCAS wavefunction can predict contraction of the donor X–H bond and an accompanying blue shift of its stretching frequency, rather than the expected elongation and red shift. The origin of this failure is the result of the optimization of the vCAS wavefunction (orbitals and CI coefficients), which yields a lower energy by using one of the weakly occupied orbitals to correlate one of the lone pairs in the dimer and not an X-H bond pair (lack of left-right correlation in the covalent X-H bond. The same calculations also overbind the dimers, a result that raises immediate concerns for any subsequent dynamical correlation treatment that takes the vCAS wavefunction as its reference.[50, 51, 52, 53] This cautionary example motivates the emphasis of the present thesis on methods that distinguish clearly between static correlation problems and the predominantly dynamical correlation that controls much of the energetics in hydrogen bonded clusters. It also explains why multi-configurational methods should not be used as black-box replacements for single reference correlation theory

in weakly bound systems. The vCAS failure in hydrogen bonded dimers and its consequences for using such wavefunctions as references for later dynamical correlation treatments are discussed in detail in the last part of this thesis.[46, 48]

2.3.4 Møller–Plesset perturbation theory

For systems that remain qualitatively single reference, the leading dynamical correlation correction is provided by Møller–Plesset perturbation theory.[54, 55, 56, 57, 58] In Møller–Plesset theory, the Hartree–Fock determinant is taken as the zeroth-order reference and the fluctuation potential—that is, the difference between the exact electron–electron interaction and its Hartree–Fock mean-field representation—is treated perturbatively.[54] The second-order correction defines MP2, which may be written in canonical spin-orbital form as

$$E_{\mathrm{MP2}} = \sum_{i<j}^{\mathrm{occ}} \sum_{a<b}^{\mathrm{virt}} \frac{|\langle ij||ab\rangle|^2}{\varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b}, \tag{2.30}$$

where i, j denote occupied spin-orbitals, a, b denote virtual spin-orbitals, and ε_p are Hartree–Fock orbital energies. The antisymmetrized two-electron integral is

$$\langle ij||ab\rangle = \langle ij|ab\rangle - \langle ij|ba\rangle, \tag{2.31}$$

where

$$\langle ij|ab\rangle = \iint \psi_i^*(1)\psi_j^*(2)\frac{1}{r_{12}}\psi_a(1)\psi_b(2)\, d1\, d2. \tag{2.32}$$

Here $d1$ and $d2$ denote integration over both spatial and spin coordinates of electrons 1 and 2, respectively. Because the sums in Eq. (2.30) are restricted to unique occupied and virtual pairs, no additional prefactor appears. The formal structure of Eq. (2.30) makes clear that the MP2 correlation energy arises from double excitations out of the Hartree–Fock determinant. Singles do not contribute at second order because of Brillouin's theorem.[59]

MP2 is especially relevant to the present work for two reasons. First, it typically recovers a substantial fraction of the dynamical correlation energy at relatively modest computational cost, making it practical for systematic studies of water clusters across several basis sets

and cluster sizes.[54, 60, 61, 57, 62, 63, 64] Second, for hydrogen-bonded water systems, MP2 often captures the dominant energetic changes associated with intermolecular binding, cooperativity, and polarization sufficiently well to serve as a quantitative baseline for many-body analyses.[65, 66, 67, 68, 69]

The strengths and limitations of MP2 must also be kept in view. Because it is perturbative, MP2 is not variational and can overestimate binding, particularly in systems with strong near-degeneracy or where the Hartree–Fock reference is qualitatively poor.[70, 71, 67] Higher-order Møller–Plesset corrections exist, but they do not always improve the description monotonically and the perturbation series is not guaranteed to converge for all systems.[70, 71] For water, nevertheless, MP2 remains an extremely useful method because it provides a transparent and affordable route to the correlation part of the interaction energy, while still being systematically improvable through basis-set enlargement and through comparison with coupled-cluster benchmarks.[62, 64, 67, 72]

2.3.5 Coupled cluster theory

Coupled cluster (CC) theory provides a higher-accuracy single-reference framework for recovering electron correlation. The coupled-cluster wavefunction is written as

$$\psi_{\mathrm{CC}} = e^{\hat{T}}\Psi_0, \qquad \hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3 + \cdots , \tag{2.33}$$

where Ψ_0 is usually the Hartree–Fock reference determinant and the cluster operator $\hat{T}$ generates connected excitations.[73, 74] The exponential ansatz is crucial because it yields a size-extensive energy when the cluster operator is truncated at a fixed excitation rank.

In practice, CCSD includes all connected single and double excitations, CCSDT adds the triple excitations, CCSDTQ the quadruple excitations and so on. A particularly popular and accurate approximation adds a perturbative treatment of connected triple excitations [CCSD(T)].[75, 76] In conventional implementations, the formal computational scaling is $\mathcal{O}(\mathbb{N}^6)$ for CCSD and $\mathcal{O}(\mathbb{N}^7)$ for CCSD(T), where $\mathbb{N}$ is the number of one-particle basis functions.[74] For small water clusters, CCSD(T) has become a benchmark level of theory

for equilibrium structures, interaction energies, and harmonic vibrational frequencies, and it therefore provides the natural high-level comparison for both MP2 and valence CASSCF in this thesis.[77, 72, 62, 63]

2.4 Basis sets

Practical electronic structure calculations rely on finite one-particle basis sets. Gaussian-type orbitals (GTOs) are widely used because molecular integrals over GTOs can be evaluated efficiently, making them suitable for molecular electronic structure calculations.[78] Molecular orbitals are expanded in this finite basis as described in Eq. (2.18); the accuracy of the calculation is therefore controlled not only by the electronic structure method, but also by the completeness of the chosen basis set.

Systematically improvable basis set families provide controlled routes toward basis set convergence. Split-valence Pople-style basis sets, such as 6-31G, were among the early widely used Gaussian basis families.[79] In correlated calculations, the correlation-consistent basis sets of Dunning and co-workers, cc-pVXZ and aug-cc-pVXZ (where X = D, T, Q, 5, and 6), are especially useful because they were designed to recover atomic correlation energies systematically as the cardinal number X increases.[80, 81, 82] In this thesis, calculations were carried out primarily with the augmented correlation-consistent basis sets because diffuse functions are important for describing the extended charge distributions, polarization, and intermolecular interactions present in hydrogen bonded water clusters.

For vibrational spectroscopy, basis set quality matters in two coupled ways. First, the equilibrium geometry and harmonic force constants depend on the electronic potential energy surface. Second, the higher derivatives of the energy required for anharmonic corrections can be more sensitive to the electronic structure level and basis set.[83, 84] Similarly, dipole moments and dipole derivatives governing IR intensities can require a flexible electronic description, especially in hydrogen-bonded systems where charge redistribution and polarization are prominent.[22, 24, 85]

Taken together, these methods define the correlation hierarchy that is relevant here:

CASSCF for systems that genuinely require a multi-configurational description of static correlation, MP2 for affordable recovery of dynamical correlation in primarily single reference systems, and CCSD(T) for benchmark-quality single reference results. This hierarchy is directly tied to the scientific questions of the thesis: why a nominally sophisticated valence CASSCF treatment can fail for hydrogen bonded dimers, how the vibrational structure of a solvated water monomer evolves under increasingly complex hydrogen bonding environments, and how lower cost correlated methods can be calibrated against coupled cluster quality benchmarks for water clusters.[77, 72, 62, 63]

2.5 Wavefunction analysis and molecular properties: orbitals, charges, and dipoles

The computed wavefunctions and electron densities can be analyzed to produce chemically interpretable descriptors. Orbital localization methods, such as the Pipek–Mezey scheme, apply unitary transformations within the occupied orbital space, and sometimes within the virtual orbital space, to yield orbitals that align more closely with chemically intuitive bonds and lone pairs.[86]

Atomic charges are not quantum mechanical observables because they depend on how the molecular electron density is partitioned. Nevertheless, charge models are indispensable in interpreting molecular electrostatics, hydrogen bonding, and force field development. Electrostatic-potential-derived charges, such as CHELPG charges, fit atom-centered point charges to reproduce the molecular electrostatic potential around a molecule.[87] Related electrostatic fitting approaches and screened charge models further illustrate both the utility and the non-uniqueness of atom-centered charge representations.[88] Topological analyses of the electron density, such as the quantum theory of atoms in molecules (QTAIM), define atoms and bonding features using the electron density and its gradient field.[89]

A molecular property of central importance for vibrational spectroscopy is the electric dipole moment. Within the electronic Born–Oppenheimer framework, the dipole moment

surface for electronic state s at fixed nuclear geometry $\mathbf{R}$ is

$$\boldsymbol{\mu}_s(\mathbf{R}) = \langle \psi_{e,s}(1,2,\ldots,N_e;\mathbf{R}) \,|\hat{\boldsymbol{\mu}}|\, \psi_{e,s}(1,2,\ldots,N_e;\mathbf{R}) \rangle \,, \tag{2.34}$$

where the electric dipole moment operator in atomic units is

$$\hat{\boldsymbol{\mu}} = \sum_{A=1}^{N_n} Z_A \mathbf{R}_A - \sum_{m=1}^{N_e} \mathbf{r}_m. \tag{2.35}$$

Here, Z_A and $\mathbf{R}_A$ are the charge and Cartesian position of nucleus A, respectively, and $\mathbf{r}_m$ is the spatial coordinate of electron m. The first term is the nuclear part of the dipole moment operator, while the second term is the electronic part. The molecular dipole moment for electronic state s at fixed nuclear geometry $\mathbf{R}$ is obtained by taking the expectation value of this operator with respect to the electronic wavefunction, as in Eq. (2.34).

Within the double-harmonic approximation, IR intensities are controlled by derivatives of the dipole moment with respect to normal-mode coordinates,

$$I_k \propto \left| \frac{\partial \boldsymbol{\mu}}{\partial Q_k} \right|^2 , \tag{2.36}$$

where Q_k is the normal coordinate for vibrational mode k.[83] Beyond this approximation, accurate dipole moment surfaces become essential because the dipole moment must be represented over the region of nuclear configuration space sampled by the vibrational wavefunction.[20, 90] For water and hydrogen bonded systems, polarization and many-body contributions can significantly affect molecular dipoles and dipole derivatives, motivating models that incorporate many-body induction and short-range corrections in a physically consistent manner.[22, 24, 91, 85]

2.6 The vibrational Hamiltonian and its representations

Once the electronic problem has been solved within the Born–Oppenheimer approximation, the remaining task is to describe the motion of the nuclei on the adiabatic potential energy surface $E_{e,s}(\mathbf{R})$ associated with electronic state s.[37] The corresponding nuclear Schrödinger equation is

$$\hat{H}_{n,s}\, \chi_{n,s\nu}(\mathbf{R}) = E_{s\nu}\, \chi_{n,s\nu}(\mathbf{R}), \tag{2.37}$$

where

$$\hat{H}_{n,s} = \hat{T}_n + E_{e,s}(\mathbf{R}) \tag{2.38}$$

is the Born–Oppenheimer nuclear Hamiltonian and

$$\hat{T}_n = -\sum_{A=1}^{N_n} \frac{1}{2M_A} \nabla_A^2 \tag{2.39}$$

is the nuclear kinetic-energy operator in atomic units. Here M_A is the mass of nucleus A, and ∇_A^2 is the Laplacian with respect to the Cartesian coordinate $\mathbf{R}_A$ of nucleus A.

For a molecule containing N_n nuclei, there are $3N_n$ Cartesian nuclear coordinates. Of these, only $3N_n - 6$ correspond to vibrational motion for a nonlinear molecule, while $3N_n - 5$ correspond to vibrational motion for a linear molecule. The remaining coordinates describe overall translation and rotation.[83]

2.6.1 Cartesian and mass-weighted Cartesian representations

A convenient starting point is to express the nuclear Hamiltonian in terms of Cartesian displacements from an equilibrium geometry $\mathbf{R}^{(0)}$. Let

$$x_i = R_i - R_i^{(0)}, \qquad i = 1, 2, \ldots, 3N_n, \tag{2.40}$$

where R_i denotes a Cartesian nuclear coordinate and $R_i^{(0)}$ is its value at equilibrium. It is then natural to introduce mass-weighted Cartesian coordinates

$$\xi_i = \sqrt{m_i}\, x_i, \qquad i = 1, 2, \ldots, 3N_n, \tag{2.41}$$

where m_i is the mass associated with coordinate x_i, i.e., the mass of the nucleus to which that Cartesian component belongs.

The kinetic energy operator in mass-weighted Cartesian coordinates has the simple form

$$\hat{T} = \frac{1}{2} \sum_{i=1}^{3N_n} \hat{p}_{\xi_i}^2, \tag{2.42}$$

where $\hat{p}_{\xi_i}$ is the momentum operator conjugate to the mass-weighted coordinate ξ_i. In coordinate representation,

$$\hat{p}_{\xi_i} = -i\frac{\partial}{\partial \xi_i}, \tag{2.43}$$

in atomic units. Therefore,

$$\hat{p}^2_{\xi_i} = -\frac{\partial^2}{\partial \xi_i^2}, \tag{2.44}$$

and Eq. (2.42) is equivalent to

$$\hat{T} = -\frac{1}{2}\sum_{i=1}^{3N_n}\frac{\partial^2}{\partial \xi_i^2}. \tag{2.45}$$

Thus, in mass-weighted Cartesian coordinates, the nuclear Hamiltonian becomes

$$\hat{H}_{n,s} = \frac{1}{2}\sum_{i=1}^{3N_n}\hat{p}^2_{\xi_i} + V_s(\boldsymbol{\xi}), \tag{2.46}$$

or equivalently,

$$\hat{H}_{n,s} = -\frac{1}{2}\sum_{i=1}^{3N_n}\frac{\partial^2}{\partial \xi_i^2} + V_s(\boldsymbol{\xi}). \tag{2.47}$$

where $V_s(\boldsymbol{\xi}) \equiv E_{e,s}(\mathbf{R}(\boldsymbol{\xi}))$. Near equilibrium, the potential may be expanded as

$$V_s(\boldsymbol{\xi}) = V_s(\mathbf{0}) + \frac{1}{2}\sum_{i,j=1}^{3N_n} F_{ij}\,\xi_i\xi_j + \frac{1}{3!}\sum_{i,j,k=1}^{3N_n} F_{ijk}\,\xi_i\xi_j\xi_k + \cdots, \tag{2.48}$$

where F_{ij}, F_{ijk}, ... are the second-, third-, and higher-order derivatives of the potential with respect to the mass-weighted Cartesian coordinates, evaluated at equilibrium. Because $\mathbf{R}^{(0)}$ is an equilibrium geometry, the first derivatives vanish.

2.6.2 Internal coordinate representation

Although Cartesian coordinates are formally complete, they are not always the most chemically transparent coordinates for vibrational motion. It is often advantageous to transform to a set of internal coordinates

$$\mathbf{s} = \{s_1, s_2, \ldots, s_f\}, \qquad f = \begin{cases} 3N_n - 6, & \text{nonlinear molecule,} \\ 3N_n - 5, & \text{linear molecule,} \end{cases} \tag{2.49}$$

where the s_a may be bond lengths, bond angles, dihedral angles, or other symmetry-adapted internal coordinates.[83]

In a general curvilinear internal-coordinate representation, the vibrational Hamiltonian may be written as

$$\hat{H}_{\mathrm{vib},s} = -\frac{1}{2}\sum_{a,b=1}^{f} g^{-1/2}\frac{\partial}{\partial s_a} g^{1/2} G^{ab}(\mathbf{s})\frac{\partial}{\partial s_b} + V_s(\mathbf{s}), \tag{2.50}$$

where $G^{ab}(\mathbf{s})$ are the elements of the kinetic-energy metric tensor (or G-matrix), and

$$g = \det \mathbf{G}(\mathbf{s}). \tag{2.51}$$

Equation (2.50) makes clear that, in internal coordinates, the kinetic energy operator is generally coupled and coordinate-dependent. This representation is particularly useful for describing large amplitude motion and for grid-based methods such as DVR, where chemically meaningful internal coordinates can provide a more natural description of the vibrational problem. Internal coordinates, therefore, provide a chemically motivated representation of the vibrational Hamiltonian, but they do not, by themselves, diagonalize the harmonic vibrational problem.

2.6.3 Normal-mode representation

For small amplitude vibrations about equilibrium, one can construct normal coordinates as special linear combinations of the coordinate displacements that diagonalize the harmonic Hamiltonian. These normal coordinates may be obtained from mass-weighted Cartesian displacements or, equivalently, from a Wilson GF treatment.[83] Let

$$\Delta s_a = s_a - s_a^{(0)} \tag{2.52}$$

be the displacement of internal coordinate s_a from its equilibrium value $s_a^{(0)}$. In the harmonic approximation, the potential is written as

$$V_s(\mathbf{s}) \approx V_s(\mathbf{s}^{(0)}) + \frac{1}{2}\sum_{a,b=1}^{f} f_{ab}\,\Delta s_a \Delta s_b, \tag{2.53}$$

where

$$f_{ab} = \left.\frac{\partial^2 V_s}{\partial s_a \partial s_b}\right|_{\mathbf{s}^{(0)}} \tag{2.54}$$

are the harmonic force constants in internal coordinates.

The normal modes are obtained by solving the Wilson GF eigenvalue problem,[83]

$$\mathbf{GFL} = \mathbf{L}\boldsymbol{\Lambda}, \tag{2.55}$$

where $\mathbf{F}$ is the matrix of harmonic force constants, $\mathbf{G}$ is the kinetic-energy matrix, $\mathbf{L}$ contains the normal-mode eigenvectors, and

$$\boldsymbol{\Lambda} = \mathrm{diag}(\omega_1^2, \omega_2^2, \ldots, \omega_f^2) \tag{2.56}$$

contains the squared harmonic angular frequencies. The normal coordinates Q_k are defined by the linear transformation

$$\Delta\mathbf{s} = \mathbf{LQ}, \tag{2.57}$$

where $\mathbf{Q} = (Q_1, Q_2, \ldots, Q_f)^{\mathrm{T}}$.

In terms of the normal coordinates, the harmonic vibrational Hamiltonian becomes a sum of independent one-dimensional harmonic oscillators,

$$\hat{H}_{\mathrm{vib},s}^{\mathrm{harm}} = \frac{1}{2}\sum_{k=1}^{f}\left(-\frac{\partial^2}{\partial Q_k^2} + \omega_k^2 Q_k^2\right). \tag{2.58}$$

If dimensionless normal coordinates

$$q_k = \sqrt{\omega_k}\, Q_k \tag{2.59}$$

are introduced, Eq. (2.58) may be written as

$$\hat{H}_{\mathrm{vib},s}^{\mathrm{harm}} = \sum_{k=1}^{f}\omega_k\left(-\frac{1}{2}\frac{\partial^2}{\partial q_k^2} + \frac{1}{2}q_k^2\right). \tag{2.60}$$

The normal mode representation is the natural starting point for harmonic vibrational analysis and perturbative anharmonic treatments such as VPT2.[84] By contrast, internal coordinate representations are often more appropriate for strongly anharmonic or large

amplitude motions and for variational grid-based approaches such as DVR. In the present work, both points of view are useful: normal modes provide the harmonic and perturbative reference picture, while internal coordinates offer a more chemically intuitive framework for describing anharmonic vibrational motion in water and hydrogen bonded systems.

2.6.4 Discrete Variable Representation

An alternative to the normal-mode expansion is to solve the vibrational Schrödinger equation variationally on a grid. One effective framework for doing so is the discrete variable representation (DVR), in which the wavefunction is represented in a basis that is localized at a set of quadrature points. The central idea is to begin from a finite basis $\{\varphi_\ell(y)\}$ for a chosen vibrational coordinate y, and to construct the matrix representation of the coordinate operator,

$$Y_{\ell\ell'} = \langle \varphi_\ell | y | \varphi_{\ell'} \rangle. \tag{2.61}$$

Here, y is a generic vibrational coordinate, $\varphi_\ell(y)$ is a finite-basis representation function, and ℓ, ℓ' label the primitive basis functions.

Diagonalization of the coordinate matrix $\mathbf{Y}$ yields a transformed basis in which the coordinate operator is diagonal. In that localized basis, operators that depend only on the coordinate are also diagonal, so that the potential energy matrix takes the simple form

$$V_{\alpha\beta} = V(y_\alpha)\delta_{\alpha\beta}, \tag{2.62}$$

where y_α is DVR grid point α, and $\delta_{\alpha\beta}$ is the Kronecker delta.

The principal advantage of DVR is that the potential energy operator is represented by point-wise evaluation on the grid, while the kinetic energy operator retains a structured analytic or semi-analytic form. In one dimension, this leads to an efficient and accurate representation of the vibrational Hamiltonian. In multiple dimensions, one commonly constructs a direct product basis,

$$\Theta_{\ell_1\ell_2\cdots\ell_d}(y_1, y_2, \ldots, y_d) = \varphi^{(1)}_{\ell_1}(y_1)\varphi^{(2)}_{\ell_2}(y_2)\cdots\varphi^{(d)}_{\ell_d}(y_d), \tag{2.63}$$

where d is the number of vibrational coordinates included explicitly in the DVR calculation. The multidimensional Hamiltonian can then be assembled from one-dimensional kinetic energy representations together with the potential evaluated on the corresponding direct product grid.

Among the many possible DVR constructions, the sinc DVR introduced by Colbert and Miller is particularly important because it provides a simple and general representation for one-dimensional kinetic energy operators on uniform grids.[92] For coordinates that are strongly anharmonic or confined to non-uniform regions of configuration space, it is often advantageous to employ a potential optimized DVR (PO-DVR),[93] in which the one-dimensional basis is adapted to the shape of a reference potential. Such approaches are especially valuable when a fully separable harmonic representation becomes inadequate and when one wishes to retain a variational treatment of large-amplitude motion, tunneling, or strong anharmonicity.

2.6.5 Second-Order Vibrational Perturbation Theory

A complementary route to anharmonic vibrational structure is provided by vibrational perturbation theory. Rather than solving the nuclear motion problem variationally on a grid, one treats the cubic and quartic terms in Eq. (2.48) as perturbations to the harmonic oscilator Hamiltonian. In second-order vibrational perturbation theory (VPT2), the energy levels are commonly written as[94, 84]

$$E(\mathbf{v}) = \sum_{k=1}^{f} \omega_k \left(v_k + \frac{1}{2} \right) + \sum_{k=1}^{f} \sum_{l=1}^{f} x_{kl} \left(v_k + \frac{1}{2} \right) \left(v_l + \frac{1}{2} \right), \tag{2.64}$$

where $\mathbf{v} = (v_1, v_2, \ldots, v_f)$ is the vector of vibrational quantum numbers, f is the number of vibrational degrees of freedom, ω_k is the harmonic frequency of mode k, and x_{kl} are the anharmonic constants derived from the cubic and quartic force constants. With this convention, the corresponding fundamental wavenumber for mode k is approximated as

$$\tilde{\nu}_k^{\text{fund}} \approx \omega_k + 2x_{kk} + \sum_{l \neq k}^{f} x_{kl}. \tag{2.65}$$

VPT2 is attractive because it preserves the intuitive language of normal modes while incorporating the leading anharmonic corrections in a computationally efficient manner. It is therefore widely used when the potential energy surface is reasonably well described by a quartic force field in the vicinity of equilibrium. Nevertheless, the method requires care in systems where accidental near-degeneracies or strong resonances are present. In water, for example, couplings between stretching modes and overtone or combination states can lead to Fermi resonances that perturb the simple perturbative picture. In such cases, a purely perturbative treatment may become unreliable, and a variational approach such as DVR can provide a more robust description.

2.7 Final remarks

The theoretical framework developed in this chapter connects the electronic structure problem directly to the vibrational observables analyzed in this thesis. Within the Born–Oppenheimer approximation, electronic structure methods define the potential energy surface and molecular properties at fixed nuclear geometries, while the nuclear motion Hamiltonian determines the vibrational energy levels supported by that surface. Errors in the electronic wavefunction, the treatment of electron correlation, or the basis set description can therefore propagate into predicted geometries, harmonic frequencies, anharmonic shifts, dipole moments, and infrared intensities.

This electronic–vibrational connection motivates the first part of the thesis, where the vibrational structure of a solvated water monomer is studied in increasingly complex hydrogen bonded environments. Harmonic normal modes provide the first description of how the bending and O–H stretching frequencies respond to solvation. VPT2 extends this picture by incorporating leading anharmonic corrections and mode couplings, while DVR provides a variational route to reduced dimensional anharmonic fundamentals along chemically important internal coordinates of the water molecule.

The same connection is central to the second part of the thesis, where the behavior of valence CASSCF for hydrogen bonded dimers is examined. Although CASSCF is designed to

describe static correlation through a multi-configurational active space, the quality of the resulting potential energy surface depends strongly on whether the active orbitals describe the physically relevant correlation. When the active space correlates the acceptor lone pair orbitals rather than the donor X–H bond, the predicted structural and vibrational response of the hydrogen bond can become qualitatively incorrect. Comparison with single reference correlated methods, such as MP2 and CCSD(T), therefore provides a useful diagnostic of whether the multi-configurational description is physically meaningful.

Together, the electronic structure and vibrational methods introduced in this chapter provide the foundation for the central questions of the thesis: how solvation modifies the intramolecular vibrations of water, how increasingly accurate treatments of the potential energy surface change the predicted spectroscopic signatures of hydrogen bonded water clusters, and how electron correlation controls the structure and vibrational response of the hydrogen bond.

Chapter 3

ENVIRONMENTAL EFFECTS IN THE VIBRATIONAL SHIFTS OF WATER: FROM THE MONOMER TO MODELS OF THE BULK

3.1 Introduction

Waters ubiquitous presence on Earth and its anomalously complex behavior have motivated extensive research into accurately modeling its intermolecular interactions. Despite its simple molecular formula, bulk H_2O exhibits remarkable emergent properties (e.g., density anomaly, multiple ice phases) that stem from its directional hydrogen-bonding network.[95] A single water molecule (monomer) in the gas phase has a dipole moment of $\sim$1.85 D, but in the condensed phase its effective dipole can increase by 3050% due to the polarizing influence of surrounding molecules.[19, 95, 96] Many-body effectscooperative polarization and hydrogen-bonding interactions involving three or more moleculesplay a key role in determining liquid waters structure and dynamics.[97] In this work we focus on bridging the gap between monomer-level properties and bulk water, with particular emphasis on how a molecules local environment modulates its intramolecular potential energy and dipole moment surfaces.

From a spectroscopic perspective, this environmental sensitivity is vividly encoded in the infrared (IR) spectrum. The fundamental vibrational frequencies for the isolated water monomer are at 1595 cm^{-1} (H–O–H bend), 3657 cm^{-1} (O–H symmetric stretch) and 3756 cm^{-1} (OH anti-symmetric stretch).[98] In the IR spectra of liquid water the HOH bending vibration appears as a relatively narrow band centered near 1640 cm^{-1}, blue-shifted by $\sim$ 45 cm^{-1} from the gas phase fundamental and broadened by only a few tens of wavenumbers. By contrast, the OH stretching band is broader and more intense. It consists of a broad peak in the 30003700 cm^{-1} range with a shoulder and a full width at half maximum (FWHM)

of $\sim$400 cm^{-1}, typically de-convoluted into substructures near 3200, 3400, and 3600 cm^{-1}, reflecting a wide distribution of local hydrogen-bond strengths and geometries in the liquid.[99, 100, 101, 91, 102, 103, 104] Strongly hydrogen-bonded OH groups contribute to the most red-shifted, low-frequency wing, whereas weakly bound or free OH groups absorb near the higher frequency shoulder. The vibrational energy relaxation and transfer processes as well as the frequency fluctuations of the OH stretching and HOH bending vibrations have been investigated in the context of ultrafast dynamics in liquid water.[105, 106] The fact that a single water molecules vibrational frequency can vary by hundreds of wavenumbers depending on its surroundings underscores that intramolecular properties such as the OH stretch frequency and the total molecular dipole moment are not fixed molecular constants in the liquid, but emergent, environment-dependent quantities. Any realistic interaction potential for water should therefore reproduce not only cluster and bulk energies, but also these environment-dependent vibrational signatures. In this study we construct hierarchical cluster networks mimicking the increasing solvation shells of a central water molecule and evaluate the convergence of its vibrational bands and features of its DMS with the degree of solvation. At the same time we re-evaluate the requirements for the 1-body term in the many body expansion-based classical potentials.

3.1.1 Inspiration behind the Solvated-Monomer Approach

Recent work by Herman and Xantheas using the TTM2.1-F potential provides a clear illustration of how strongly a water molecules dipole moment depends on its local environment.[19] By analyzing the distribution of monomer dipole moments in clusters ranging from $n = 5 - 102$, they showed that both the average dipole and its spread increase with cluster size (from $\sim$ 2.2 D in small clusters to nearly $\sim$ 2.82.9 D in large, bulk-like clusters). Tetrahedrally coordinated clusters mimicking the first and second solvation shell of water ($n = 5, 17$) as well as a fully coordinated water dimer ($n = 8$) displayed well-separated dipole moment populations in the different solvation shells, and molecules with higher coordination numbers systematically exhibited larger dipole moments than less-coordinated molecules. At the

same time, the fully relaxed, larger clusters (n = 53, 102) showed broader, more bulk-like distributions. This previous analysis highlighted the fact that the static enhanced dipole built into pairwise interaction potentials to account for the many-body effects (i.e. TIP4P) is not truly environment dependent, undermining their transferability across different environments and pointing toward the need for an explicitly environment-dependent, inducible molecular dipole.

A similar picture emerges from vibrational spectroscopy. The broad OH stretching band in the IR spectrum of liquid water, contrasted with the comparatively sharp HOH bending band, reflects the fact that a water molecules intramolecular force constants are modulated by its local hydrogen-bond environment.[91, 102, 103] In a previous study, Paesani, Xantheas, and Voth combined centroid molecular dynamics with the TTM3-F interaction potential to investigate the IR spectrum of HOD in water.[91] They found that a cluster of roughly 10 water molecules around an HOD reproduces the major features of the liquids OH stretch line shape, and a cluster of $\sim$30 molecules (about two solvation shells) yields near-quantitative agreement with the bulk spectrum. Beyond the second solvation shell, additional water molecules have a negligible influence on the HOD vibrational frequency. In their analysis, the OH stretching frequency was almost perfectly correlated with the local electric field experienced by that OH bond, which is dominated by the first- and second-shell neighbors. These results demonstrated that, within the accuracy of the TTM3-F interaction potential, the spectroscopic signature of a water monomer is primarily determined by its immediate environment, and that molecules beyond the second solvation shell contribute relatively little to the vibrational spectra.

Motivated by these previous findings, we attempt to provide an *ab initio*-based investigation of the effect of the local environment on the intramolecular vibrational frequencies. A truly transferable 1-body term for water must reproduce not only interaction energies but also geometries, dipole moments, and vibrational observables across environments. Our approach is facilitated by employing a hierarchy of clusters with increasing solvation shells to assess how the frequencies of a single water molecule evolve from the gas phase values to a bulk-like

environment. By analyzing the central molecules dipole moment via CHELPG-derived[107] charges and its vibrational fundamentals along selected internal coordinates, we obtain a first principles analog of the environment-dependent dipole and frequency trends. In a future study we will use these data to parametrize a field-responsive monomer model, such that its charge distribution (and associated dipole) and vibrational force constants continuously adjust with the external electric field and local coordination. Rather than treating many-body effects only through global force-field parameters, we examine a target water molecule within an ensemble and explicitly encode how its intramolecular potential responds to surrounding molecules. By design, the solvated monomer approach also remedies the asymptotic limitation of gas-phase 1-body surfaces: along a stretched OH coordinate in a hydrogen-bonded network, the resulting PES connects to the heterolytic OH^- + H_3O^+ ionic dissociation limit rather than to the homolytic $OH^\bullet$ + $H^\bullet$ radical asymptote of the isolated monomer.

3.2 Methodology

3.2.1 Construction of cluster geometries

The water cluster networks studied were $(H_2O)_n$ with n = 1, 2, 5, 8, 17, and 53 are shown in Fig. 3.1. Equilibrium structures of the monomer (C_{2v} symmetry) and dimer (C_s symmetry) were optimized at the second-order Møller–Plesset perturbation theory[61] (MP2) and the Coupled-Cluster Singles and Doubles with perturbative Triples [108, 109] [CCSD(T)] levels of theory with Dunnings augmented correlation-consistent basis sets, aug-cc-pVXZ (abbreviated aVXZ, where $X = D$, T, Q and 5).[110, 82]

For the larger clusters we designed the geometries that model the increase in the number of solvation shells around a central molecule, while maintaining C_2 symmetry for the solvated monomer and C_s symmetry for the solvated dimer throughout the geometry optimizations. The first solvation shell is modeled by a tetrahedrally coordinated monomer, $(H_2O)_5$ or *tet*-5, in which four nearest neighbors are hydrogen bonded to a central water in a (*ddaa*) arrangement [111, 112] (*d*: donor, *a*: acceptor), with all O–O–O angles fixed at 109.5°. This

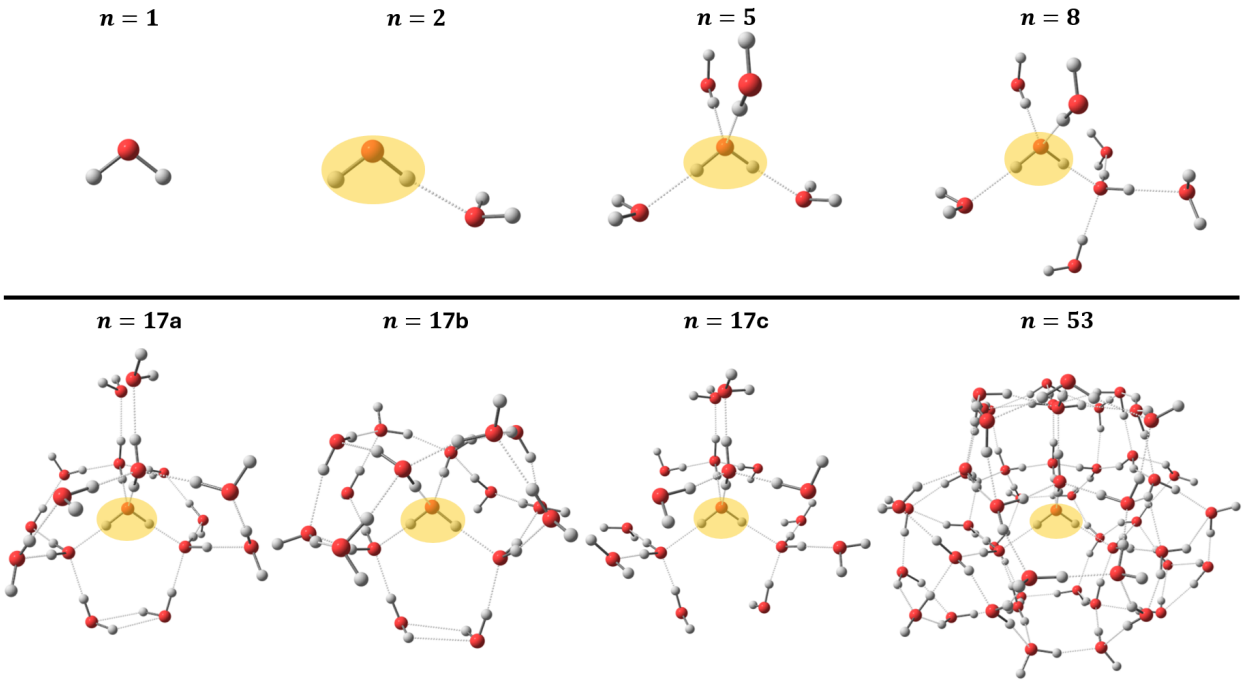

Figure 3.1: Cluster geometries of $(H_2O)_n$, where $n = 1$ (monomer), 2 (dimer), 5 (tetrahedrally coordinated monomer), 8 (tetrahedrally coordinated dimer), 17 (second solvation shell), 53 (third solvation shell), considered in this work. Structure 17a is constrained to maintain tetrahedral coordination throughout, whereas 17b retains tetrahedral coordination only in the first shell with a relaxed second shell. Structure 53 mimics up to the third solvation shell retaining the tetrahedrality up to the second solvation shell with a relaxed third shell. Structure 17c is made up from the two inner solvation shells cut out from the $(H_2O)_{53}$ cluster. The colored regions indicate the water molecule for which the DVR is performed.

geometrically constrained pentamer, which is not a minimum on the PES, was optimized at the MP2 with up to a quadruple zeta basis set. Subsequently, keeping the four peripheral monomers frozen at their MP2/aVQZ geometry, the central monomer of the pentamer was further optimized at the CCSD(T)/aVTZ level. Likewise, we constructed the tetrahedrally coordinated dimer, $(H_2O)_8$ or *tet*-8, to model the first solvation shell of the dimer: three water molecules (*a*, *d*, *d*) were tetrahedrally arranged about the donor monomer and three other (*a*, *a*, *d*) about the acceptor monomer of $(H_2O)_2$. This octamer, which also is not a minimum on the PES, was optimized at the MP2/aVXZ, X = D, T, levels. For the $(H_2O)_{17}$ cluster we considered three networks: one in which the monomer is embedded in two tetrahedrally coordinated solvation shells [$(H_2O)_{17a}$ or *tet-a*-17], another in which only the first solvation shell is constrained to be tetrahedral while the second shell is relaxed [$(H_2O)_{17b}$ or *tet-b*-17] and a third one cut out from the optimized $(H_2O)_{53}$ or *tet*-53 network having tetrahedral coordination in the first and second solvation shells but different O–O distances than the previous two [$(H_2O)_{17c}$ or *tet-c*-17]. These three networks were optimized under the geometric constraints at the MP2/aVXZ, X=D, T, levels. As noted earlier, all structures larger than the dimer do not correspond to minima, so a full unconstrained optimization of $(H_2O)_5$ will relax it to the ring minimum,[113] the optimization of $(H_2O)_8$ to the cube minimum,[72, 62] whereas the optimization of $(H_2O)_{17}$ to the interior structure, viz. a fully solvated water molecule inside a near spherical cluster.[63] Finally, $(H_2O)_{53}$ or *tet*-53 is intended to mimic the third solvation shell of the monomer. We enforced tetrahedrality up to the second shell and relaxed the 2 solvation shells under this constraint while fully relaxing the third solvation shell at the MP2/aVDZ level. The HF, MP2 and CCSD(T) energies and associated harmonic vibrational frequencies were obtained with the NWChem 7.2.3[114] and Gaussian 16[115] suites of electronic structure codes.

3.2.2 The Vibrational Hamiltonian and Discrete Variable Representation (DVR)

The vibrational energies were computed from the internal bond-angle coordinates Hamiltonian for a triatomic molecule with the volume element $dr_1 dr_2 d\theta$:

$$H^{3\mathrm{D}} \;=\; T \;+\; V' \;+\; V \tag{3.1}$$

where

$$\begin{aligned} 2T = {} & p_{r_1} G_{r_1 r_1} p_{r_1} + p_{r_2} G_{r_2 r_2} p_{r_2} + p_\theta G_{\theta\theta} p_\theta \\ & + p_{r_1} G_{r_1 r_2} p_{r_2} + p_{r_2} G_{r_1 r_2} p_{r_1} + p_{r_1} G_{r_1 \theta} p_\theta \\ & + p_\theta G_{r_1 \theta} p_{r_1} + p_{r_2} G_{r_2 \theta} p_\theta + p_\theta G_{r_2 \theta} p_{r_2} \end{aligned}$$

and

$$p_{r_1} = -i\hbar \frac{\partial}{\partial r_1}, \qquad p_{r_2} = -i\hbar \frac{\partial}{\partial r_2}, \qquad p_\theta = -i\hbar \frac{\partial}{\partial \theta},$$

the G-matrix elements are

$$\begin{aligned} G_{r_1 r_1} = G_{r_2 r_2} &= \frac{1}{m_O} + \frac{1}{m_H}, \\ G_{\theta\theta} &= \frac{1}{m_H r_1^2} + \frac{1}{m_H r_2^2} + \frac{1}{m_O}\left(\frac{1}{r_1^2} + \frac{1}{r_2^2} - \frac{2\cos\theta}{r_1 r_2}\right), \\ G_{r_1 r_2} &= \frac{\cos\theta}{m_O}, \qquad G_{r_1 \theta} = -\frac{\sin\theta}{m_O r_2}, \qquad G_{r_2 \theta} = -\frac{\sin\theta}{m_O r_1}, \end{aligned}$$

and

$$\begin{aligned} V' = {} & \frac{\hbar^2}{m_H r_1 r_2} cos\theta - \frac{\hbar^2}{m_O r_1 r_2} cos\theta - \frac{\hbar^2}{m_O r_2 r_1} cos\theta \\ & + \frac{\hbar^2}{2 m_O r_1 r_2} cos\theta - \frac{\hbar^2}{8} G_{\theta\theta}(2 + cot^2\theta) \end{aligned}$$

where r_1, r_2, and θ are the two O–H bond lengths and the H–O–H angle. V' is the mass-dependent term of the kinetic operator that originates from the transformation from Cartesian to internal coordinates. The potential V is the *ab initio* PES evaluated on the grid with the various methods [MP2, CCSD(T), classical potentials]. Atomic masses for ^{1}H and ^{16}O were used throughout. We adopted the analytical expressions of the G-matrix elements and V' tabulated by Frederick and Woywod.[116]

The Hamiltonian was set up to employ Colbert-Miller's discrete variable representation (DVR or CM-DVR),[117] a powerful method to probe anharmonicity by numerically solving the Schrödinger equation on a uniform sinc-DVR grid $\{q_i\}$ where $q_i \in \{r_1, r_2, \theta\}$ with equal spacing Δq. Since direct three-dimensional (3D) discretization on a uniform grid is prohibitively expensive, we begin with extracting the information from the one-dimensional (1D) cuts of the potential for each internal coordinate q_i, while holding all other internal coordinates fixed at their optimized values. The 1D Hamiltonian is

$$H^{\mathrm{1D}} = \frac{1}{2}\, p_J\, G_{JJ'}\, p_{J'} + V'(q_i) + V(q_i), \tag{3.2}$$

where the momentum matrix elements $p_{JJ'}$ in Eq. 3.2 were taken from Gardenier and coworkers, analogous to the CM-DVR procedure:[118]

$$p_{JJ'} \;=\; \begin{cases} 0, & J = J', \\ \dfrac{-\mathrm{i}\hbar}{\Delta q}\dfrac{(-1)^{J-J'}}{(J-J')}, & J \neq J', \end{cases} \tag{3.3}$$

where J and J' denote DVR grid-point indices. These 1D CM-DVR solutions provided the needed quadrature points/weights to deploy the potential-optimized DVR (PO–DVR),[93, 119] where 1D CM-DVR quadratures are converted into non-uniform 1D grids tailored to $V(q)$, a procedure that dramatically reduces the number of grid points without sacrificing accuracy. The resulting direct-product 3D PO–DVR grids ($n_{r_1} \times n_{r_2} \times n_\theta$) were used to access the convergence of the fundamentals of H_2O and $(H_2O)_2$ with respect to the size of the grid. This optimization is essential for the extension of this method to larger clusters, where tight CM-DVR grids are computationally prohibitive for the 1D and 3D problems.

The harmonic frequencies of $(H_2O)_n$, $n = 2$, 5, 8, 17, were computed for the $H_2O(D_2O)_{n-1}$ isotopologues, consisting of a central H_2O and perdeuterated coordinated molecules. This isotopic decoupling suppresses the intermolecular HH vibrational coupling so the reported modes are localized on the respective monomer in the cluster. For the solvated monomers, in addition to the isotopic decoupling protocol, we restricted the anharmonic treatment to the

targeted monomer with the DVR performed on the coordinates of the studied monomer, and selected only the modes associated with that monomer for calculations at the second-order vibrational perturbation theory (VPT2). For all 3D PO–DVR calculations the grid size was selected to converge the O–H stretches to < 2 cm^{-1} and the bend to < 0.5 cm^{-1} (*vide infra*). The fundamental vibrational energies (ν_i) obtained from the 3D DVR are coH_2O (C_{2v} symmetry)mpared with those from VPT2.

The scheme used to arrive at the best CCSD(T)/Complete Basis Set (CBS) estimates for the anharmonic frequencies taking into account the level of electron correlation and the size of the basis set used for the DVR calculations is described below. To estimate the VPT2 fundamentals (ν_i) at the CCSD(T) level of theory with a specific basis set, we added the MP2 anharmonic correction to the CCSD(T) harmonic frequencies (ω_i) as used in previous spectroscopic benchmarks.[120, 121]

$$\nu_{\mathrm{VPT2}}^{\mathrm{CCSD(T)}} \approx \omega_{\mathrm{harm}}^{\mathrm{CCSD(T)}} + \left(\nu_{\mathrm{VPT2}}^{\mathrm{MP2}} - \omega_{\mathrm{harm}}^{\mathrm{MP2}}\right) \tag{3.4}$$

where $\omega_{\mathrm{harm}}^{\mathrm{CCSD(T)}}$ and $\omega_{\mathrm{harm}}^{\mathrm{MP2}}$ are the harmonic frequencies at the CCSD(T) and MP2 levels, respectively, $\nu_{\mathrm{VPT2}}^{\mathrm{MP2}}$ is the MP2 VPT2 fundamental and $\nu_{\mathrm{VPT2}}^{\mathrm{CCSD(T)}}$ is the estimated CCSD(T) VPT2 fundamental. All VPT2 anharmonic frequencies were obtained with Gaussian 16.[115] We estimated basis set effects by applying a correction equal to the difference between the anharmonic frequency computed with the aVXZ basis set and the corresponding CBS value. Here, the CBS reference is taken as the monomer result obtained with the aV5Z basis set. The best estimate for the frequencies evaluated via DVR, which takes into account the above corrections, is:

$$\begin{aligned}\nu_{\mathrm{cluster}}^{\mathrm{CCSD(T)/CBS}} \approx\ & \nu_{\mathrm{cluster}}^{\mathrm{CCSD(T)/aVXZ}} \\ & + \left(\nu_{\mathrm{monomer}}^{\mathrm{CCSD(T)/aV5Z}} - \nu_{\mathrm{monomer}}^{\mathrm{CCSD(T)/aVXZ}}\right)\end{aligned} \tag{3.5}$$

where

$$\nu_{\text{cluster}}^{\text{CCSD(T)/aVDZ}} \approx \nu_{\text{cluster}}^{\text{MP2/aVDZ}} + \left(\nu_{tet-5}^{\text{CCSD(T)/aVDZ}} - \nu_{tet-5}^{\text{MP2/aVDZ}}\right) \tag{3.6}$$

3.3 Results & Discussion

3.3.1 The Gas Phase Water Monomer H_2O

We performed initial convergence tests for the gas phase H_2O as regards the level of theory, basis set and DVR grid size. Table 3.1 lists the equilibrium geometry (r_e, θ_e), the harmonic, VPT2, and DVR bending (ν_2), symmetric (ν_1), and antisymmetric stretching (ν_3) frequencies, along with the results with the PS PES for comparison.

Table 3.1: Equilibrium bond lengths r_e (Å), angles θ_e (°), harmonic (ω) and anharmonic vibrational frequencies ν_1, ν_2, and ν_3 (cm^{-1}) of H_2O at the MP2 and CCSD(T) levels of theory. Their values for the PS PES and experiment (Exp.) are listed for comparison.

Method		**MP2**				**CCSD(T)**				**PS**	**Exp.**
		aVDZ	aVTZ	aVQZ	aV5Z	aVDZ	aVTZ	aVQZ	aV5Z		
r_e (Å)		0.9659	0.9614	0.9589	0.9584	0.9665	0.9616	0.9590	0.9584	0.9578	0.9578[122]
θ_e (°)		103.87	104.11	104.27	104.34	103.96	104.19	104.38	104.44	104.51	104.48[122]
	ω	1622	1628	1632	1632	1638	1646	1650	1650		
ν_2	VPT2	1573	1578	1580	1579	1588	1595	1598	1597	1595	1595[98]
	DVR	1572	1576	1579	1578	1588	1594	1597	1596		
	ω	3803	3822	3840	3843	3787	3811	3831	3835		
ν_1	VPT2	3622	3654	3668	3670	3605	3643	3659	3662	3657	3657[98]
	DVR	3626	3656	3670	3672	3602	3638	3655	3658		
	ω	3938	3948	3966	3969	3905	3920	3941	3945		
ν_3	VPT2	3745	3768	3782	3784	3712	3740	3757	3760	3756	3756[98]
	DVR	3749	3770	3784	3788	3707	3734	3752	3756		
$2\nu_2$	DVR	3112	3117	3121	3119	3144	3152	3155	3155	3152	3152[98]

Both MP2 and CCSD(T) monotonically change the values of the internal coordinates, decreasing r_e and increasing θ_e with increasing size of the basis set. For a given basis set, MP2 predicts a slightly shorter bond length than CCSD(T) by 0.1–0.6 mÅ and a slightly smaller H–O–H angle by ~ 0.08–$0.11°$. With the aV5Z basis set the two theories produce practically indistinguishable values for r_e (0.9584 Å). The internal geometry at the CCSD(T) level of theory with the aVQZ and aV5Z basis sets is very close to that of the PS PES ($r_e^{\mathrm{PS}} = 0.9578$ and $\theta_e^{\mathrm{PS}} = 104.51°$). MP2/aVTZ is already quite accurate, with deviations of +3.6 mÅ in r_e (0.9614 Å) and $-0.40°$ in θ_e (104.11°), both being $< 0.4\%$ compared to the PS values, while MP2/aVDZ (0.9659 Å, 103.87°) remains within $< 1\%$ for both quantities.

Table 3.2: Convergence of gas phase H_2O DVR fundamentals (cm^{-1}) with respect to the 3D PO-DVR grid size $n_{r_1} \times n_{r_2} \times n_\theta$. Δ denotes the deviation from the highest product grid for each frequency at the specific level of theory.

n_i			DVR fundamentals (cm^{-1})					
n_{r_1}	n_{r_2}	n_θ	ν_2	Δ	ν_1	Δ	ν_3	Δ
MP2/aVDZ								
5	5	5	1571.6	0.1	3130.4	495.9	3554.1	194.7
6	6	6	1571.7	0.0	3114.8	511.9	3618.5	130.6
7	7	7	1571.7	0.0	3625.1	1.5	3670.4	78.8
8	8	8	1571.7	0.0	3626.3	0.3	3748.8	0.3
9	9	9	1571.7	0.0	3626.5	0.1	3749.1	0.1
10	10	10	1571.7	0.0	3626.6	0.0	3749.1	0.0
11	11	11	1571.5	0.2	3626.6	0.0	3749.1	0.0
12	12	12	1571.7	0.0	3626.6	0.0	3749.1	0.0
13	13	13	1571.7	0.0	3626.6	0.0	3749.1	0.0
15	15	15	1571.7	0.0	3626.6	0.0	3749.1	0.0

20	20	20	1571.7	0.0	3626.6	0.0	3749.1	0.0
25	25	25	1571.7	0.0	3626.6	0.0	3749.1	0.0
MP2/aVTZ								
5	5	5	1576.3	0.1	3189.9	228.9	3285.4	176.8
6	6	6	1576.4	0.0	3118.7	537.0	3648.3	121.7
7	7	7	1576.4	0.0	3654.4	1.3	3656.9	113.1
8	8	8	1576.4	0.0	3655.5	0.3	3769.7	0.3
9	9	9	1576.4	0.0	3655.7	0.1	3769.9	0.1
10	10	10	1576.4	0.0	3655.7	0.1	3769.9	0.1
11	11	11	1576.3	0.1	3655.7	0.0	3770.0	0.0
12	12	12	1576.4	0.0	3655.8	0.0	3770.0	0.0
13	13	13	1576.4	0.0	3655.8	0.0	3770.0	0.0
20	20	20	1576.4	0.0	3655.8	0.0	3770.0	0.0
MP2/aVQZ								
5	5	5	1580.6	0.2	3147.2	520.3	3586.1	195.4
6	6	6	1578.7	0.1	3122.1	547.8	3662.2	122.3
7	7	7	1578.7	0.0	3647.2	22.6	3668.5	116.1
8	8	8	1578.8	0.0	3669.5	0.3	3784.3	0.3
9	9	9	1578.8	0.0	3669.8	0.1	3784.5	0.1
10	10	10	1578.8	0.0	3669.8	0.0	3784.5	0.0
12	12	12	1578.8	0.0	3669.8	0.0	3784.6	0.0
13	13	13	1578.7	0.0	3669.8	0.0	3784.6	0.0
20	20	20	1578.8	0.0	3669.8	0.0	3784.6	0.0
25	25	25	1578.8	0.0	3669.8	0.0	3784.6	0.0
MP2/aV5Z								
7	7	7	1578.0	0.0	3645.9	26.4	3671.0	116.9

8	8	8	1578.0	0.0	3672.0	0.3	3787.6	0.3
10	10	10	1578.0	0.0	3672.3	0.0	3787.8	0.0
12	12	12	1578.0	0.0	3672.3	0.0	3787.8	0.0
CCSD(T)/aVDZ								
6	6	6	1587.8	0.0	3144.9	457.7	3592.9	114.3
7	7	7	1587.8	0.0	3600.8	1.8	3705.5	1.7
8	8	8	1587.8	0.0	3602.1	0.4	3706.8	0.4
9	9	9	1587.8	0.0	3602.5	0.1	3707.1	0.1
10	10	10	1587.8	0.0	3602.5	0.0	3707.1	0.0
15	15	15	1587.8	0.0	3602.6	0.0	3707.2	0.0
CCSD(T)/aVTZ								
6	6	6	1593.9	0.0	3152.0	486.7	3630.0	104.5
7	7	7	1594.0	0.0	3637.1	1.6	3710.2	24.2
8	8	8	1594.0	0.0	3638.3	0.4	3734.1	0.4
9	9	9	1594.0	0.0	3638.6	0.1	3734.4	0.1
10	10	10	1594.0	0.0	3638.6	0.0	3734.4	0.0
15	15	15	1594.0	0.0	3638.7	0.0	3734.5	0.0
CCSD(T)/aVQZ								
7	7	7	1596.5	0.0	3653.8	1.6	3701.7	50.5
8	8	8	1596.5	0.0	3655.1	0.4	3751.9	0.4
9	9	9	1596.5	0.0	3655.3	0.1	3752.1	0.1
10	10	10	1596.5	0.0	3655.4	0.0	3752.2	0.0
15	15	15	1596.5	0.0	3655.5	0.0	3752.2	0.0
CCSD(T)/aV5Z								

7	7	7	1596.2	0.0	3657.1	1.6	3699.2	57.1
8	8	8	1596.2	0.0	3658.4	0.4	3756.0	0.3
10	10	10	1596.2	0.0	3658.7	0.0	3756.3	0.0

To assess the numerical convergence of the DVR fundamentals, we constructed PO-DVR grids for each internal coordinate with varying numbers of points, and performed 3D DVR calculations on the associated direct product grids (Table 3.2). For very coarse grids ($n_i = 56$ points per coordinate), the bending fundamental is captured reasonably well, whereas the stretching fundamentals can deviate by several hundred cm^{-1} from their converged values, underscoring the challenge of representing the strongly anharmonic OH potentials with too few grid points. Increasing to $n_i = 7$ substantially reduces these errors bringing ν_1 within a few cm^{-1} and ν_3 to within ~ 100 cm^{-1} of the dense grid limits (Table S1). Further refinement to $n_i = 8$ reduces the remaining differences to ~ 0.4 cm^{-1} across all levels of theory and basis sets. The residual DVR discretization error at $n_i = 8$ is therefore negligible for all basis sets and correlation methods used. We accordingly adopt an $8 \times 8 \times 8$ PO–DVR grid for subsequent calculations of the fundamentals for the monomer, so that deviations from the PS reference primarily reflect differences in the underlying PES rather than DVR discretization artifacts.

Fig. 3.2 shows the variation of the harmonic (triangles), VPT2 (crosses) and DVR (circles) fundamentals for the various basis sets at the MP2 (left panel) and CCSD(T) (right panel) levels of theory. The harmonic ν_2 values are higher than the PS ones by 28-38 cm^{-1} at MP2 and 43-55 cm^{-1} at CCSD(T) as the basis is enlarged. As regards the stretches, the harmonic blue shifts are larger, amounting to $\Delta\nu_1 = 146$–186 cm^{-1} (MP2) and 130–178 cm^{-1} [CCSD(T)], and to $\Delta\nu_3 = 182$–213 cm^{-1} (MP2) and 149–189 cm^{-1} [CCSD(T)]. The inclusion of anharmonicity via either VPT2 or DVR brings the fundamentals closer to the experimental values as the basis set increases. At a fixed theory/basis level, the differences between VPT2 and DVR are < 2 cm^{-1} for ν_2 and typically $\sim$ 1–6 cm^{-1} for ν_1 and ν_3. At the CCSD(T) level, DVR converges to near-quantitative agreement with experiment as the basis is enlarged:

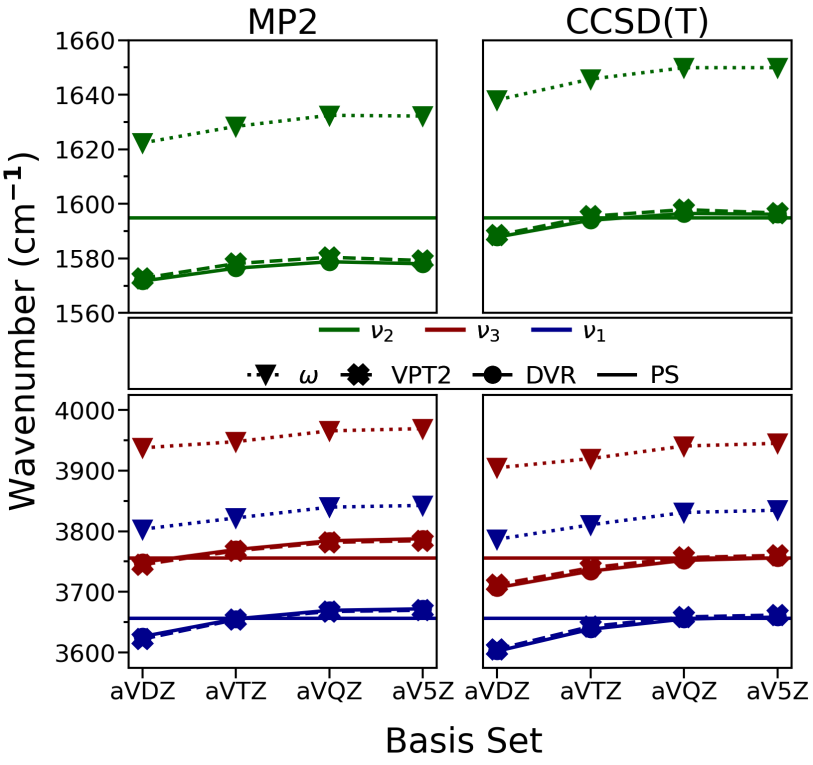

Figure 3.2: Harmonic (triangles), VPT2 (squares) and DVR (circles) vibrational frequencies (cm^{-1}) ν_1 (blue), ν_2 (green), ν_3 (red) of gas phase H_2O with the correlation consistent basis sets at the MP2 (left) and CCSD(T) (right) levels of theory. Horizontal lines denote the PartridgeSchwenke (PS) fundamentals, which are within <1 cm^{-1} from experiment.

with the aV5Z basis set $\nu_2 = 1596$, $\nu_1 = 3658$, and $\nu_3 = 3756$ cm^{-1}, values that are within 1, 1, and 0 cm^{-1} from experiment, respectively. By contrast, MP2 exhibits a systematic low/high bias in the bending/stretching fundamentals with large basis sets. Practically, these monomer benchmarks provide a consistent reference for quantifying solvation-induced frequency shifts in the cluster hierarchy, and they confirm that the remaining errors in the MP2-based monomer PES are small compared to the several-hundred-wavenumber shifts in the stretching frequencies induced by hydrogen bonding in the larger clusters.[113]

3.3.2 The Gas Phase Water Dimer $(H_2O)_2$

We extended the study to the water dimer to quantify how the presence of an intermolecular hydrogen bond (HB) perturbs the intramolecular structure and fundamentals of the donor monomer (indicated by a shaded area in Fig. 3.1) of $(H_2O)_2$. We use r^b and r^f to indicate the donor hydrogen bonded and free O–H bond lengths, respectively. The donor fundamentals analyzed are the bend (ν_2), the hydrogen bonded O–H $\nu(OH_b)$, and the free O–H $\nu(OH_f)$ stretches. We employed a reduced-dimensional DVR along r^b_{OH}, r^f_{OH}, θ, while holding the O–O stretch fixed, reflecting its weak coupling to the intramolecular modes. The bending coordinate is represented as a weighted combination of the two half-angles relative to the H–O–H bisector,

$$\theta_{\text{bend}} = 0.53\,\theta_{\frac{1}{2}}^{(b)} + 0.47\,\theta_{\frac{1}{2}}^{(f)},$$

where $\theta_{\frac{1}{2}}^{(b)}$ is the half-angle between the donor hydrogen bonded O–H vector and the H–O–H bisector, and $(\theta_{\frac{1}{2}}^{(f)})$ is the half angle between the free O–H vector and the H–O–H bisector. The coefficients were obtained by projecting the harmonic bending normal mode of the donor water onto two half-angle internal coordinates defined with respect to the fixed equilibrium H–O–H bisector. These weights describe the normal mode of the donor $(H_2O)_2$ bending vibration that reflects the slightly larger participation of the donating hydrogen in the bending motion at the dimer minimum. Table 3.3 lists the equilibrium geometry and frequencies of the dimer at the various levels of theory and basis sets.

Table 3.3: Geometrical parameters, harmonic (ω) and anharmonic (ν) vibrational frequencies (cm^{-1}) of the donor water in $(H_2O)_2$ at the MP2 and CCSD(T) levels of theory. r_e^b and r_e^f are the water donor hydrogen bonded and free O–H equilibrium bond lengths, respectively, θ_e is the donor equilibrium H–O–H angle, and r_{O-O}^d is the equilibrium O–O distance. The vibrationally averaged O–O distances $\tilde{r}_{O-O}^d$ are included to compare to experiment.[123]

Method		**MP2**				**CCSD(T)**				**CCSD(T)-F12b/CBS**	**Exp.**
		aVDZ	aVTZ	aVQZ	aV5Z	aVDZ	aVTZ	aVQZ	aV5Z	**correction**[124]	
r_e^b (Å)		0.9728	0.9686	0.9663		0.9723	0.9678	0.9653		0.9641	
r_e^f (Å)		0.9648	0.9604	0.9579		0.9656	0.9609	0.9582		0.9569	
θ_e (°)		104.27	104.49	104.64		104.28	104.52	104.69		104.85	
r_{O-O}^d (Å)		2.9166	2.9072	2.9026		2.9256	2.9140	2.9100		2.9092	
$\tilde{r}_{O-O}^d$ (Å)		2.9754	2.9645	2.9525		2.9844	2.9713	2.9599			2.976(1)[123]
ν_2	ω	1641	1649	1652		1659	1668	1671			
	ω^{est}		1647 (-2)	1651 (-1)	1651 ± 2		1666 (-2)	1670 (-1)	1670 ± 2		
	VPT2	1598	1601	1604		1615	1620	1623			1620[125]
	$VPT2^{est}$		1603 (2)	1605 (1)	1604 ± 2		1622 (2)	1624 (1)	1623 ± 2		
	DVR	1599	1605	1607		1615	1623	1624			
	DVR^{est}		1604 (-1)	1606 (-1)	1605 ± 1		1621 (-2)	1624 (0)	1623 ± 2		
$\nu(OH_b)$	ω	3705	3720	3734		3712	3731	3748			
	ω^{est}		3724 (4)	3742 (8)	3744 ± 8		3736 (5)	3756 (8)	3760 ± 8		
	VPT2	3532	3558	3566		3539	3569	3580			3601[125]
	$VPT2^{est}$		3564 (6)	3577 (11)	3579 ± 11		3576 (7)	3592 (12)	3595 ± 12		
	DVR	3501	3526	3535		3513	3544	3556			
	DVR^{est}		3530 (4)	3544 (9)	3547 ± 9		3549 (5)	3566 (10)	3569 ± 10		
$\nu(OH_f)$	ω	3904	3915	3933		3875	3891	3912			
	ω^{est}		3914 (-1)	3932 (-1)	3936 ± 1		3890 (-1)	3911 (-1)	3916 ± 1		
	VPT2	3714	3738	3751		3685	3714	3730			3730[125]
	$VPT2^{est}$		3737 (-1)	3751 (0)	3754 ± 1		3713 (-1)	3730 (0)	3734 ± 1		
	DVR	3717	3738	3753		3679	3707	3725			
	DVR^{est}		3738 (0)	3752 (-1)	3756 ± 1		3706 (-1)	3724 (-1)	3728 ± 1		
$2\nu_2$	VPT2	3163	3167	3171		3199	3205	3209			3194[126]
	DVR	3167	3175	3176		3198	3208	3210			

Table 3.4: Convergence of gas phase $(H_2O)_2$ DVR fundamentals (cm^{-1}) with respect to the 3D PO-DVR grid size $n_{r_1} \times n_{r_2} \times n_\theta$. Δ denotes the deviation from the highest product grid for each frequency at the specific level of theory. $\nu(OH_b)$ and $\nu(OH_f)$ denote the hydrogen bonded and free OH stretches of the donor molecule, respectively.

n_i			DVR fundamentals (cm^{-1})					
n_{r_1}	n_{r_2}	n_θ	ν_2	Δ	$\nu(OH_b)$	Δ	$\nu(OH_f)$	Δ
MP2/aVDZ								
5	5	5	1599.2	0.1	3172.3	328.6	3378.8	338.0
6	6	6	1599.1	0.0	3166.4	334.5	3485.6	231.2
7	7	7	1599.1	0.0	3497.7	3.2	3715.3	1.5
8	8	8	1599.1	0.0	3500.1	0.8	3716.4	0.3
9	9	9	1599.1	0.0	3500.7	0.2	3716.6	0.1
10	10	10	1599.1	0.0	3500.9	0.0	3716.8	0.0
15	15	15	1599.1	0.0	3500.9	0.0	3716.8	0.0
MP2/aVTZ								
5	5	5	1605.2	0.1	3179.6	347.4	3414.4	324.2
6	6	6	1605.3	0.0	3175.0	351.9	3512.8	225.9
7	7	7	1605.3	0.0	3524.0	3.0	3737.4	1.3
8	8	8	1605.3	0.0	3526.2	0.8	3738.3	0.3
9	9	9	1605.3	0.0	3526.8	0.2	3738.6	0.1
10	10	10	1605.3	0.0	3527.0	0.0	3738.7	0.0

The same PO-DVR convergence analysis was carried out for the gas phase water dimer (Table 3.4). As in the monomer case, we find that the bending fundamental ν_2 converges faster with the grid size. Even coarse grids with $n_i = 56$ already reproduce the converged

bend mode to better than 0.1 cm^{-1} at both the MP2/aVDZ and MP2/aVTZ levels. However, the hydrogen-bonded $\nu(OH_b)$ and free OH ($\nu(OH_f)$) stretches exhibit large errors on this size grids, with typical deviations of $\sim$ 200350 cm^{-1} from the converged grids. Increasing the grid to $n_i = 7$ removes most of this discrepancy; the stretching fundamentals now differ from the dense grid limits by only $\sim$ 13 cm^{-1} across basis sets at MP2. Further refinement to an $n_i = 8$ product grid reduces the residual errors below 1 cm^{-1} for all three modes, with no statistically meaningful changes upon extending to $n_i = 9$, $n_i = 10$ or $n_i = 15$ grid. This behavior mirrors the monomer convergence pattern, namely rapid convergence of the bend and slower but systematic convergence of the strongly anharmonic stretches, and justifies the use of an $n_i = 8$ PO–DVR grid for the dimer's subsequent calculations at higher levels of theory and larger basis sets, for which DVR discretization errors are safely below the other sources of uncertainty discussed in the main text.

When comparing to the gas phase monomer, we reference both donor stretches to the average of the monomer stretches, $\overline{\nu}_{OH}$, evaluated at the same level of theory, basis set, and vibrational treatment. Across the three vibrational treatments, the donor bend modes are blue-shifted by about 25 cm^{-1} relative to the monomer bend, and the $\nu(OH_f)$ fundamentals are mildly blue-shifted by about 30 cm^{-1} relative to the monomer average stretch $\overline{\nu}_{OH}$ at both the MP2 and CCSD(T) levels of theory. However, the $\nu(OH_b)$ modes are strongly red-shifted by roughly 150–170 cm^{-1} at the MP2 and 125–135 cm^{-1} at the CCSD(T) levels of theory.

The results of Table 3.3 indicate that anharmonic treatment of the dimer brings the results into near-quantitative agreement with experimental values from Ref. 125. The donor bend and $\nu(OH_f)$ stretch at the CCSD(T)/VPT2 level of theory converge to the experimental values within a few cm^{-1} and place the $\nu(OH_b)$ stretch within $\sim$ 0.5–1% of experiment as the cardinal number of the basis set increases. DVR yields a systematically larger red-shift for the $\nu(OH_b)$ stretch, remaining $\sim$ 40–50 cm^{-1} below experiment even at CCSD(T)/aVQZ, consistent with the stronger coupling of $\nu(OH_b)$ to other modes that are omitted in our reduced-dimensional DVR. Table 3.3 shows that VPT2 and DVR give a consistent description of the donor intramolecular fundamentals in $(H_2O)_2$ with the aVDZ through the aVQZ basis

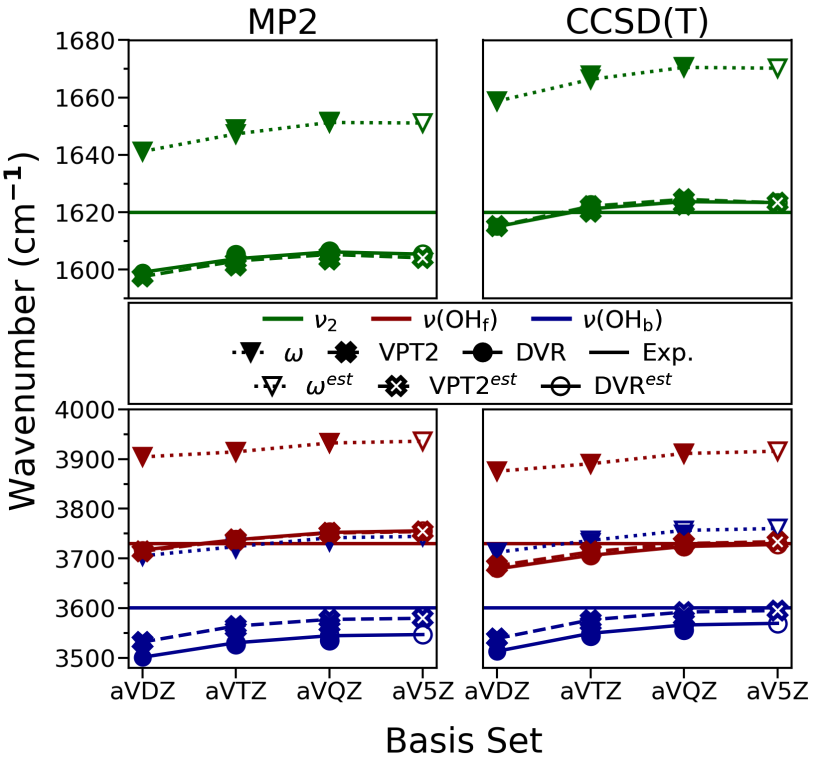

Figure 3.3: Harmonic (triangles), VPT2 (squares) and DVR (circles) vibrational frequencies (cm^{-1}) $\nu(OH_b)$(blue), ν_2 (green), and $\nu(OH_f)$ (red) of gas phase $(H_2O)_2$ with the correlation consistent basis sets at the MP2 (left) and CCSD(T) (right) levels of theory. Horizontal lines trace the experimental values (Exp.).[125]

sets. For the ν_2 and $\nu(OH_f)$ fundamentals, the two anharmonic treatments track each other closely, remaining within a few cm^{-1} for each basis set and reaching sub-percent deviations by the aVTZ basis set. We also highlight that the estimated frequencies (^{est}) in Table 3.3) from Eq. 3.5 are in excellent agreement with the calculated results.

The agreement between the VPT2 and DVR results is exemplary. At the MP2 level, ν_2 deviates from experiment by -1.4% to -1.0% (VPT2) and -1.3% to -0.8% (DVR) as the basis increases from aVDZ to aVQZ, while $\nu(OH_f)$ stays near the experimental values and changes sign smoothly with basis from -0.4% (aVDZ) to $+0.6\%$ (aVQZ) for both the VPT2

and DVR anharmonic treatments. The agreement between VPT2 and DVR is retained at the CCSD(T) level while the comparison to experiment improves: ν_2 lies within -0.3% to $+0.2\%$ for the two anharmonic treatments, while $\nu(\mathrm{OH_f})$ improves from -1.2% (VPT2) and -1.4% (DVR) at aVDZ to essentially 0.0% (VPT2) and -0.2% (DVR) at aVQZ. The only persistent difference between the two vibrational treatments is the hydrogen bonded stretch $\nu(\mathrm{OH_b})$, where DVR is systematically more red-shifted than VPT2 by $\sim$ 24–32 $\mathrm{cm^{-1}}$, i.e., an additional $\sim$ 0.7–0.9% shift relative to experiment. This corresponds to MP2 deviations of -1.9% to -1.0% (VPT2) versus -2.8% to -1.8% (DVR), and CCSD(T) deviations of -1.7% to -0.6% (VPT2) versus -2.4% to -1.3% (DVR), consistent with the enhanced sensitivity of $\nu(\mathrm{OH_b})$ to intermolecular coupling effects in the dimer. A complementary comparison is the bend overtone $2\nu_2$, where VPT2 and DVR predicted the values that are close together and in good agreement with the experimental values from Bouteiller, Tremblay, and Perchard.[126]

3.3.3 The First Solvation Shell of the Monomer: Tetrahedrally Coordinated Pentamer $(\mathbf{H_2O})_5$

A tetrahedral (*ddaa*) arrangement around the central H_2O, resulting in a tetrahedrally coordinated pentamer $(H_2O)_5$ or *tet*-5, imposes angular constraints and a cooperative local field from four nearest neighbors. The surrounding water molecules donate and accept hydrogen bonds in a way that is more isotropic than the dimer, so the central monomer experiences stronger many-body induction and charge redistribution. The results of Table 3.5 confirm the expected structural signatures manifested by longer OH bonds and a larger H–O–H angle, accompanied by a blue shift of the bend and red-shifts of both stretches with respect to the monomer and dimer values.

At MP2/aVDZ the O–H bond (r_e) of the central water molecule increases from 0.9659 Å (monomer) to 0.9764 Å (*tet*-5), while the H-O-H angle (θ_e) increases from 103.87° to 105.00°. These shifts correspond to $\Delta r_e = +10.5$ mÅ and $\Delta\theta_e = +1.13°$ relative to the monomer with the same basis set. Therefore, increasing the coordination from one hydrogen bond in the dimer to four nearest neighbors in *tet*-5 amplifies the intramolecular imprint. Furthermore, these

Table 3.5: Geometric parameters, harmonic (ω) and anharmonic (ν) vibrational frequencies (cm^{-1}) of the tetrahedrally coordinated monomer in $(H_2O)_5$ at the MP2 and CCSD(T) levels of theory. The estimated values (^{est}) from Eq. 3.5 and Eq. 3.6 are also listed with their difference from the calculated values in parentheses.

		MP2				**CCSD(T)**				
		aVDZ	aVTZ	aVQZ	aV5Z	aVDZ	aVTZ	aVQZ	aV5Z	**Exp.**
r_e (Å)		0.9764	0.9730	0.9707		0.9755	0.9716			0.970±0.005,[127] 0.983±0.008[128]
θ_e (°)		105.00	105.39	105.49		105.03	105.53			106.1±0.9,[127] 104.1±1.9[128]
r^d_{O-O} (Å)		2.8318	2.8213	2.8204		2.8429	2.8204			2.80±0.001[129]
ν_2	ω	1670	1675	1677		1683	1693			
	ω^{est}		1677 (2)	1681 (4)	1680 ± 4		1690 (-3)	1694 ± 4	1694 ± 4	
	VPT2	1642	1645			1654	1663			
	VPT2^{est}		1647 (3)	1650 ± 3	1648 ± 3		1661 (-2)	1663 ± 3	1662 ± 3	
	DVR	1633	1636	1636		1647	1650			
	DVR^{est}		1638 (2)	1640 (4)	1639 ± 4		1653 (3)	1656 ± 3	1655 ± 3	
ν_1	ω	3603	3616	3629		3628	3645			
	ω^{est}		3622 (6)	3639 (10)	3642 ± 10		3652 (7)	3672 ± 10	3676 ± 10	
	VPT2	3409	3441			3434	3470			
	VPT2^{est}		3442 (1)	3455 ± 1	3457 ± 1		3472 (2)	3487 ± 2	3490 ± 2	
	DVR	3394	3417	3426		3427	3456			
	DVR^{est}		3423 (6)	3438 (12)	3440 ± 12		3463 (7)	3480 ± 12	3483 ± 12	
ν_3	ω	3687	3690	3704		3700	3704			
	ω^{est}		3697 (7)	3715 (11)	3719 ± 11		3715 (11)	3736 ± 11	3740 ± 11	
	VPT2	3479	3497			3492	3511			
	VPT2^{est}		3502 (5)	3516 ± 5	3519 ± 5		3520 (9)	3537 ± 9	3540 ± 9	
	DVR	3447	3461	3472		3468	3488			
	DVR^{est}		3468 (8)	3482 (10)	3485 ± 10		3496 (8)	3513 ± 10	3517 ± 10	
$2\nu_2$	VPT2	3154	3154			3180	3189			
	DVR	3230	3233	3233		3258	3261			

structural parameters are close to the experimental liquid water values ($R_{\mathrm{OH}} = 0.970 \pm 0.005$ Å, $\theta_{\mathrm{HOH}} = 106.1 \pm 0.9°$[127] and $R_{\mathrm{OH}} = 0.983 \pm 0.008$ Å, $\theta_{\mathrm{HOH}} = 104.1 \pm 1.9°$[128]) as well as with the experimental nearest-neighbor O$\cdots$O distance of approximately 2.8 Å, obtained from x-ray diffraction.[129]

Table 3.6: Convergence of tetrahedrally coordinated monomer, $(H_2O)_5$, DVR fundamentals (cm^{-1}) with respect to the 3D PO-DVR grid size $n_{r_1} \times n_{r_2} \times n_\theta$. Δ denotes the deviation from the highest product grid for each frequency at the specific level of theory.

n_i			DVR fundamentals (cm^{-1})					
n_{r_1}	n_{r_2}	n_θ	ν_2	Δ	ν_1	Δ	ν_3	Δ
MP2/aVDZ								
5	5	5	1633.5	0.5	3171.6	224.2	3260.9	187.4
6	6	6	1633.1	0.1	3234.5	161.3	3371.6	76.7
7	7	7	1633.0	0.0	3389.8	6.0	3442.0	6.3
8	8	5	1632.7	0.3	3394.4	1.4	3446.6	1.7
8	8	6	1633.0	0.0	3394.4	1.4	3446.6	1.7
8	8	7	1633.0	0.0	3394.2	1.6	3446.6	1.7
8	8	8	1633.0	0.0	3394.2	1.6	3446.6	1.7
9	9	5	1632.7	0.3	3395.6	0.2	3447.8	0.5
9	9	6	1632.9	0.1	3395.5	0.3	3447.8	0.5
9	9	7	1632.9	0.1	3395.4	0.4	3447.8	0.5
9	9	9	1633.0	0.0	3395.4	0.4	3447.8	0.5
10	10	10	1633.0	0.0	3395.6	0.2	3448.0	0.3
12	12	12	1633.0	0.0	3395.8	0.0	3448.2	0.1
14	14	14	1633.0	0.0	3395.8	0.0	3448.3	0.0
MP2/aVTZ								

5	5	5	1636.3	0.2	3189.9	228.9	3285.4	176.8
6	6	6	1636.3	0.1	3239.6	179.2	3395.3	66.9
7	7	7	1636.2	0.0	3412.9	5.9	3456.2	6.0
8	8	8	1636.2	0.0	3417.1	1.7	3460.5	1.8
9	9	9	1636.1	0.0	3418.2	0.6	3461.7	0.6
10	10	10	1636.1	0.0	3418.5	0.3	3462.0	0.3
12	12	12	1636.1	0.0	3418.6	0.2	3462.1	0.2
15	15	15	1636.2	0.0	3418.8	0.0	3462.3	0.0
CCSD(T)/aVDZ								
7	7	7	1647.0	0.0	3424.1	3.2	3465.4	3.2
8	8	8	1647.0	0.0	3426.7	0.5	3468.1	0.5
9	9	9	1647.0	0.0	3427.2	0.1	3468.6	0.1
10	10	10	1647.0	0.0	3427.3	0.0	3468.7	0.0

The investigation of the convergence of the vibrational frequencies of the central molecule in *tet*-5 with respect to the DVR grid (Table 3.6) reaffirms the previously reported trend: the bend is rapidly converged with respect to grid density, whereas the strongly anharmonic stretches demand substantially denser product grids. Table 3.6 shows that the discrepancy due to the grid size for $(H_2O)_5$ is modestly larger than that for the isolated monomer and dimer. At MP2 with the aVDZ and aVTZ basis sets the $n_i = 8$ grid produces stretches that differ from the dense grid limits by roughly 1.0–2.0 cm^{-1}, whereas these differences are reduced to below 0.5 cm^{-1} with the $n_i = 9$ grid. This indicates that the tetrahedral environment slightly increases the sensitivity of the stretch fundamentals to DVR discretization. The residuals with respect to the converged grid decrease at the CCSD(T) level. At CCSD(T)/aVDZ the stretches are within ~ 3.0 cm^{-1} with a $n_i = 7$ grid and within ~ 0.5 cm^{-1} with a $n_i = 8$ grid from the converged results. For this system we also tested non-cubic grids, in an attempt to reduce the number of grid points for the angle coordinate while maintaining accuracy

and convergence (see Table 3.6). The $8 \times 5 \times 8$ (5 points in the angle coordinate) grid reproduces the ν_3 fundamental obtained with the cubic $n_i = 8$ grid (ν_2 and ν_1 are within 0.3 cm^{-1}). Likewise the ν_2, ν_1, and ν_3 with the $9 \times 9 \times 5$ grid are within 0.3, 0.2, and 0.0 cm^{-1}, respectively, from the fundamentals obtained with the cubic $n_i = 9$ grid. These results suggest that the tetrahedrally coordinated monomer's fundamentals are converged with a $9 \times 5 \times 9$ grid, whereas the $8 \times 5 \times 8$ grid provides a reliable choice balancing accuracy and computational cost. The DVR fundamentals reported in Table 3.5 are obtained from the cubic $n_i = 8$ grid to ensure consistent comparison with the previous results for the gas phase monomer and dimer.

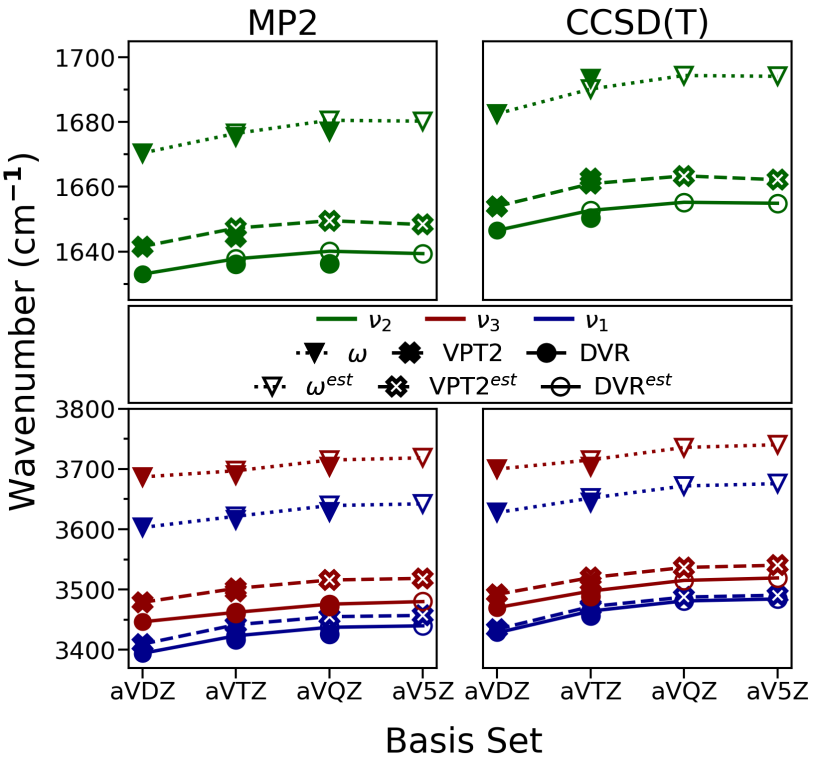

Figure 3.4: Harmonic (triangles), VPT2 (squares) and DVR (circles) vibrational frequencies (cm^{-1}) ν_1(blue), ν_2 (green), and ν_3 (red) of the central water molecule in the tetrahedrally coordinated $(H_2O)_5$ with the correlation consistent basis sets at the MP2 (left) and CCSD(T) (right) levels of theory. Filled/open symbols correspond to the calculated/estimated values.

In *tet*-5 the bend of the central monomer is blue-shifted, whereas the stretches are red-shifted relative to the gas-phase monomer. At VPT2/aVTZ the bend shifts are $\Delta\nu_2 = 66$ cm^{-1} (MP2) and 67 cm^{-1} [CCSD(T)], while the stretch shifts $\Delta\nu_1/\Delta\nu_3$ are $-213/-271$ cm^{-1} at MP2, and $-173/-229$ cm^{-1} at CCSD(T). The reduced-dimensional DVR approach produces slightly larger shifts in which $\Delta\nu_2/\Delta\nu_1/\Delta\nu_3$ are $+60/-238/-309$ cm^{-1} at MP2/aVTZ, and $+56/-182/-246$ cm^{-1} at CCSD(T)/aVTZ level. A notable feature of the *tet*-5 O–H stretches is their relatively small splitting, which remains on the order of 50–80 cm^{-1} irrespective of the method employed.

The VPT2 and DVR results remain in quantitative agreement on the overall solvation imprint, while exhibiting a systematic method separation. For *tet*-5 VPT2 consistently produces slightly larger fundamentals (smaller red shift) than DVR for the central molecule with differences that are small on the scale of the total solvation-induced shifts but reproducible across basis sets and levels of electronic structure theory. At the MP2 level, the fundamentals with VPT2 exceed those from DVR by 8–9 cm^{-1} for the bend, by 15–25 cm^{-1} for ν_1, and by 30–40 cm^{-1} for ν_3 over the basis set range from aVDZ to aVTZ. The same trends are observed at the CCSD(T) level with the differences between VPT2 and DVR being 7–12 cm^{-1} for ν_2, 7–14 cm^{-1} for ν_1, and 20–25 cm^{-1} for ν_3. The persistence (and mild growth) of this difference between VPT2 and DVR with increasing environmental strength supports the view that the stretches are increasingly sensitive to couplings with the intermolecular coordinates that are not explicitly represented in the reduced-dimensional DVR manifold, whereas the normal mode framework underlying VPT2 can retain part of this influence even when anharmonicity is applied only to selected modes.

The estimated values (^{est}) for the fundamentals listed in Table 3.5 are obtained by transferring basis set increments from the gas phase monomer to the embedded monomer reference values with the aVDZ basis set for the same electronic structure level and/or vibrational treatment. For MP2 this procedure yields uniformly slightly higher frequencies than the explicit *tet*-5 calculations. For the harmonic and DVR bend, the estimated values exceed the calculated ones by 1–4 cm^{-1}, while the harmonic stretch estimates are higher

by 6–11 cm^{-1}. The VPT2 estimates remain very close to the calculated values, differing by 1–5 cm^{-1}. For CCSD(T)/aVTZ the bend is reproduced to within a few cm^{-1}, with the harmonic estimate slightly lower by about 3 cm^{-1}, while the stretch estimates are modestly higher by 7–11 cm^{-1} (harmonic), 2–9 cm^{-1} (VPT2), and 7–8 cm^{-1} (DVR). Overall, the monomer-derived estimate from Eq. 3.5 provides a practical uncertainty range for the embedded monomer that is small compared to the induced absolute shifts due to solvation. Given this consistent performance (errors of only a few cm^{-1} for ν_2 and at most on the order of 10 cm^{-1} for the stretches) we adopt the same monomer-based transfer scheme to estimate the fundamentals with the larger basis sets for the larger clusters, where explicit CCSD(T) anharmonic calculations are prohibitively expensive.

3.3.4 The First Solvation Shell of the Dimer: Tetrahedrally Coordinated Octamer $(\mathbf{H_2O})_\mathbf{8}$

The tetrahedrally coordinated dimer embedded in the $(H_2O)_8$ cluster (*tet*-8) mimics the first solvation shell of the gas phase dimer. Analogously, we denote the $r_e^{b_1}$ and $r_e^{b_2}$ of the donor of the solvated dimer in a similar manner with the r_e^b and r_e^f in the donor of the isolated dimer. This is the first structure in our set where the O–O separation ($r_{\mathrm{O-O}}^{d_1}$) associated with the O–H_{b_1} coordinate is about 2.70 Å (see Table 3.7), closely matching the experimental O–O distance of $\sim$ 2.75 Å in ice Ih at 60 K,[130, 131, 132, 133] indicative of a substantially stronger hydrogen bond (ice-like) than in $(H_2O)_2$ and *tet*-5. In contrast, the O–O distance ($r_{\mathrm{O-O}}^{d_2}$) along the O–H_{b_2} coordinate is about 2.85 Å, suggesting a similar hydrogen bond (liquid-like) to *tet*-5. Table 3.7 summarizes the intramolecular structures and vibrational fundamentals of the donor embedded dimer in *tet*-8 as the bend ν_2, the O–H stretches $\nu(OH_{b_1})$ and $\nu(OH_{b_2})$, corresponding to $r_e^{b_1}$ and $r_e^{b_2}$.

At the MP2/aVDZ level the embedded dimer's donor exhibits a substantially larger intramolecular distortion than the gas phase dimer donor. Relative to the donor O–H in $(H_2O)_2$, $r_e^{b_1}$ increases to 0.9851 Å ($\Delta r_e = +12.3$ mÅ), $r_e^{b_2}$ increases to 0.9749 Å ($\Delta r_e = +10.1$ mÅ), and the donor angle opens up to 105.61° ($\Delta\theta_e = +1.35°$). At the MP2/aVTZ level, the same trend holds with slightly larger bond-length changes. These additional distortions

Table 3.7: Geometric parameters, harmonic (ω) and anharmonic (ν) vibrational frequencies (cm^{-1}) of the donor molecule of the tetrahedrally coordinated dimer in $(H_2O)_8$ at the MP2 and CCSD(T) levels. The $r^{d_1}_{O-O}$ and $r^{d_2}_{O-O}$ are the O–O distances along a O–H bond $r^{b_1}_e$ (corresponding to the r^b_e of the donor of gas phase dimer) and along the $r^{b_2}_e$ (corresponding to the "free" O–H bond of the gas phase dimer's donor molecule) embedded in *tet*-8. $\nu(OH_{b_1})$ and $\nu(OH_{b_2})$ are the two hydrogen bonded O–H stretching frequencies of the embedded dimer's donor molecule that correspond to the hydrogen bonded and free stretching frequencies in the gas phase (non-coordinated) dimer. The estimated values (^{est}) from Eq. 3.5 and Eq. 3.6 are also listed with their difference from the calculated values in parentheses.

Method		**MP2**				**CCSD(T)**			
		aVDZ	aVTZ	aVQZ	aV5Z	aVDZ	aVTZ	aVQZ	aV5Z
$r^{b_1}_e$ (Å)		0.9851	0.9823						
$r^{b_2}_e$ (Å)		0.9749	0.9714						
θ_e (°)		105.61	105.99						
$r^{d_1}_{O-O}$ (Å)		2.6921	2.6818						
$r^{d_2}_{O-O}$ (Å)		2.8521	2.8443						
ν_2	ω	1671	1681						
	ω^{est}		1677 (-4)	1681 ± 4	1681 ± 4	1683 ± 4	1691 ± 4	1695 ± 4	1695 ± 4
	VPT2	1662	1677						
	VPT2^{est}		1667 (-10)	1670 ± 10	1669 ± 10	1675 ± 10	1682 ± 10	1685 ± 10	1684 ± 10
	DVR	1640	1646						
	DVR^{est}		1644 (-2)	1647 ± 2	1646 ± 2	1653 ± 2	1660 ± 2	1662 ± 2	1662 ± 2
$\nu(OH_{b_1})$	ω	3449	3451						
	ω^{est}		3468 (17)	3486 ± 17	3489 ± 17	3474 ± 17	3498 ± 17	3518 ± 17	3522 ± 17
	VPT2	3246	3261						
	VPT2^{est}		3278 (17)	3291 ± 17	3293 ± 17	3270 ± 17	3308 ± 17	3324 ± 17	3327 ± 17
	DVR	3160	3156						
	DVR^{est}		3190 (34)	3204 ± 34	3206 ± 34	3193 ± 34	3229 ± 34	3246 ± 34	3249 ± 34
$\nu(OH_{b_2})$	ω	3677	3686						
	ω^{est}		3687 (1)	3705 ± 1	3709 ± 1	3690 ± 1	3705 ± 1	3726 ± 1	3731 ± 1
	VPT2	3472	3497						
	VPT2^{est}		3495 (-2)	3509 ± 2	3512 ± 2	3485 ± 2	3513 ± 2	3530 ± 2	3534 ± 2
	DVR	3461	3475						
	DVR^{est}		3482 (7)	3496 ± 7	3500 ± 7	3482 ± 7	3510 ± 7	3527 ± 7	3532 ± 7
$2\nu_2$	VPT2	3141	3152						
	DVR	3270	3278						

upon tetrahedrally solvating the dimer are comparable to the ones in *tet*-5, where the changes in the central (solvated) monomer relative to the gas phase monomer are $\Delta r_e = +10.5$ mÅ and $\Delta\theta_e = +1.13°$ at MP2/aVDZ, while also producing a comparably large elongation of the nominally "free" O–H bond, that is consistent with a stronger and more anisotropic local hydrogen bond environment.

The bending coordinate of *tet*-8 is represented as a weighted of the two half angles, $\theta_{\frac{1}{2}}^{(b_1)}$ and $\theta_{\frac{1}{2}}^{(b_2)}$, with a slightly different contribution than the gas phase (non-coordinated) dimer:

$$\theta_{\text{bend}} = 0.52\,\theta_{\frac{1}{2}}^{(b_1)} + 0.48\,\theta_{\frac{1}{2}}^{(b_2)}.$$

Fig. 3.5 shows the pronounced difference between the $\nu(\text{OH}_{\text{b}_1})$ (blue) and $\nu(\text{OH}_{\text{b}_2})$ (red) fundamentals, indicating two different hydrogen bond environments consistent with the different structural parameters $r_{\text{O–O}}^{d_1}$ and $r_{\text{O–O}}^{d_2}$. Relative to the gas phase dimer donor, the embedded dimer in *tet*-8 is associated with a clear incremental solvation imprint. The VPT2 results predict a large stiffening of the bend, with $\Delta\nu_2 = +64$ cm^{-1} at aVDZ and $+76$ cm^{-1} at aVTZ. This is accompanied by additional redshifts in the stretching modes, namely $\Delta\nu(\text{OH}_{\text{b}_1}) = -286$ (aVDZ) and -296 (aVTZ) cm^{-1} relative to $\nu(\text{OH}_{\text{b}})$ of the dimer, and $\Delta\nu(\text{OH}_{\text{b}_2}) = -185$ (aVDZ) and -241 (aVTZ) cm^{-1} relative to $\nu(\text{OH}_{\text{f}})$ of the dimer. In contrast, DVR produces a more modest incremental bend shift, with $\Delta\nu_2 = 40$ cm^{-1} at aVDZ and 41 cm^{-1} at aVTZ, while the stretching modes remain strongly softened, with $\Delta\nu(\text{OH}_{\text{b}_1}) = -341$ and -370 cm^{-1}, and $\Delta\nu(\text{OH}_{\text{b}_2}) = -256$ and -263 cm^{-1}, relative to $\nu(\text{OH}_{\text{b}})$ and $\nu(\text{OH}_{\text{f}})$ of the dimer, respectively. In a similar manner to *tet*-5, the first solvation shell of the gas phase monomer and dimer induces a blue shift in the bend of about 60 cm^{-1} and a red shift in the O–H stretches of about 200 to 300 cm^{-1} with an exceptional ice-like environment where the stronger hydrogen bond enhances the red-shift up to around 400 cm^{-1} (*vide infra*).

For this network, the VPT2 and DVR results remain quantitatively consistent. Relative to DVR, VPT2 places ν_2 higher by 23 and 31 cm^{-1} at MP2/aVDZ and MP2/aVTZ, respectively, and overestimates the $\Delta\nu(\text{OH}_{\text{b}_1})$ by 85 and 105 cm^{-1} and by 12 and 22 cm^{-1} for $\Delta\nu(\text{OH}_{\text{b}_2})$.

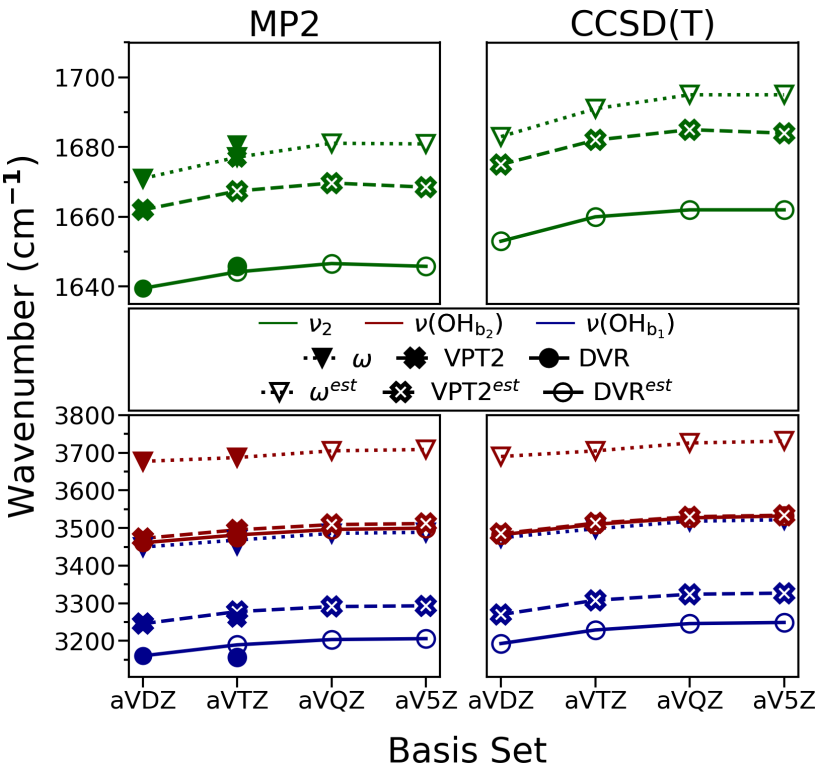

Figure 3.5: Harmonic (triangles), VPT2 (squares) and DVR (circles) vibrational frequencies (cm^{-1}) $\nu(OH_{b_1})$(blue), ν_2 (green), and $\nu(OH_{b_2})$ (red) of the donor water molecule in the tetrahedrally coordinated dimer within the $(H_2O)_8$ cluster with the correlation consistent basis sets at the MP2 (left) and CCSD(T) (right) levels of theory. Filled/open symbols correspond to the calculated/estimated values.

Since $\Delta\nu(OH_{b_1})$ associates with a strong hydrogen bond where the perturbation is large, a large difference between VPT2 and DVR is expected even with our best estimate from Eq. 3.5 and Eq. 3.6 for this mode. However, these offsets are small compared with the solvation-induced softening of the dimer donor, which amounts to several hundred cm^{-1}. Additionally, the comparison between the calculated and estimated MP2/aVTZ values in Table 3.7 shows that transferring gas phase monomer / basis set increments remains a useful approximation for the singly solvated dimer, albeit with variable, mode-dependent accuracy.

3.3.5 The Second Solvation Shell of the Monomer: Tetrahedrally Coordinated Heptadecamer $(\mathbf{H_2O})_{17}$

To probe how the second solvation shell further influences the intramolecular structure and fundamental frequencies of an embedded water molecule, we considered three second solvation shell motifs in the $(H_2O)_{17}$ cluster shown in Fig. 3.1 and described in the Methodology section. Similar to the previous case for *tet*-5, we analyze the central monomer's bend (ν_2), symmetric stretch (ν_1), and antisymmetric stretch (ν_3), using isotopic decoupling to localize the normal modes on the embedded H_2O. The corresponding structural and vibrational parameters are summarized in Table 3.8.

The effect of the second solvation shell is to amplify the intramolecular distortion of the solvated monomer relative to both the gas phase monomer and the singly solvated water in *tet*-5. In the *tet-a*-17 network the central monomer's O–H bonds are substantially elongated and the H–O–H angle opens up further. At the MP2/aVDZ level the O–H bond (r_e) of the central molecule in *tet-a*-17 increases to 0.9950 Å ($\Delta r_e = +29.1$ mÅ relative to the gas-phase monomer and $\Delta r_e = +18.6$ mÅ relative to *tet*-5). Likewise, the H-O-H angle (θ_e) increases to 106.19° ($\Delta\theta_e = +2.32°$ relative to H_2O and $\Delta\theta_e = +1.19°$ relative to *tet*-5). The same trend holds at the MP2/aVTZ level. Interestingly, the r^d_{O-O} of *tet-a*-17 at MP2/aVDZ and MP2/aVTZ levels of 2.6752 and 2.6628 Å, respectively. These large geometric distortions indicate that the electrostatic field of the fully tetrahedral second shell and hydrogen bond cooperativity *substantially* polarize the embedded water beyond the effect of the first shell tetrahedral motif. By contrast, relaxing the second shell in *tet-b*-17 produces an apparent structural response with the O–H bonds still lengthening, but the H–O–H angle remains comparatively close to the gas phase value and is *smaller* than in the central water molecule of *tet*-5. Moreover, relaxing the second shell resulting in a longer r^d_{O-O} than *tet-a*-17 by 30.0 mÅ (MP2/aVDZ) and 42.0 mÅ (MP2/aVTZ). Thus, the first-shell tetrahedral coordination alone is sufficient to drive the large O–H elongation, but the angular response is highly sensitive to the tetrahedrality of the second solvation shell. This result signifies the effect of the fleeting

hydrogen bonding network on the sampling of bend angles and correspondingly the bending vibrational frequencies, as evidenced by the deviations in ν_2 for *tet-a*-17 and *tet-b*-17 relative to our best estimates from Eq. 3.5 reported in Table 3.8.

In crystalline Ih ice each water molecule participates in four hydrogen bonds arranged in a nearly tetrahedral network.[111, 134] The O–H and O–O distances of the *tet-a*-17 and *tet-b*-17 networks place them close to the experimental and theoretical results for ice Ih, for which the covalent O–H bond length is near 0.985±0.05 Å[131, 132] (1.000–1.008 Å by Kuhs and Lehmann[130]) and the nearest-neighbor O–O distance is near 2.75 .[130, 131, 132, 133] The intramolecular H–O–H angle is typically centered near 106° in simulations and spectroscopic analyses of ice and liquid water.[135, 131] This increase reflects the influence of the tetrahedral hydrogen bond network, which pulls the two OH bonds apart relative to the isolated molecule. In this study, *tet-a*-17 exhibits a central H–O–H angle of about 106° (see Table 3.8), placing it essentially as the characteristic bend angle observed for tetrahedrally coordinated water molecules in ice-like environments. The close agreement with the value for ice suggests that enforcing tetrahedral coordination through the first and second solvation shells in *tet-a*-17 reproduces the local angular constraints of the ice lattice remarkably well and captures the essential local geometry of an ice-like coordination environment around the central molecule. These comparisons suggest that the *tet*-17 structural motifs are best viewed as probing the crossover from liquid-like to ice-like local order, with *tet-a*-17 lying at the most strongly ordered end of that spectrum.

Table 3.8: Geometric parameters, harmonic (ω) and anharmonic (ν) vibrational frequencies (cm^{-1}) at MP2 and CCSD(T) of the tetrahedrally coordinated central H_2O in the three $(H_2O)_{17}$ motifs, viz. the fully tetrahedral two shell $(H_2O)_{17a}$ (*tet-a*-17), the tetrahedral first, relaxed second shell $(H_2O)_{17b}$ (*tet-b*-17), and the the fully tetrahedral two shell network taken out of $(H_2O)_{53}$ (*tet-c*-17). r^d_{O-O} is the donor O–O distance along a donating hydrogen bond of the embedded monomer. The estimated values (^{est}) of *tet-a*-17, *tet-b*17, and *tet-c*17 from Eq. 3.5 and Eq. 3.6 are also listed with their difference from the calculated values in parentheses.

		MP2				**CCSD(T)**			
		aVDZ	aVTZ	aVQZ	aV5Z	aVDZ	aVTZ	aVQZ	aV5Z
					$(\mathbf{H_2O})_{\mathbf{17a}}$				
r_e (Å)		0.9950	0.9926						
θ_e (°)		106.19	106.60						
r^d_{O-O} (Å)		2.6752	2.6628						
ν_2	ω	1673	1692						
	ω^{est}		1680 (-12)	1684 ± 12	1683 ± 12	1686 ± 12	1693 ± 12	1697 ± 12	1697 ± 12
	DVR	1646	1661						
	DVR^{est}		1650 (-11)	1653 ± 11	1652 ± 11	1660 ± 11	1666 ± 11	1668 ± 11	1668 ± 11
ν_1	ω	3263	3268						
	ω^{est}		3282 (14)	3299 ± 14	3302 ± 14	3288 ± 14	3312 ± 14	3332 ± 14	3336 ± 14
	DVR	2934	2930						
	DVR^{est}		2963 (33)	2977 ± 33	2979 ± 33	2966 ± 33	3002 ± 33	3019 ± 33	3022 ± 33
ν_3	ω	3292	3284						
	ω^{est}		3302 (18)	3320 ± 18	3323 ± 18	3305 ± 18	3320 ± 18	3340 ± 18	3345 ± 18
	DVR	2936	2942						
	DVR^{est}		2957 (15)	2971 ± 15	2975 ± 15	2957 ± 15	2985 ± 15	3002 ± 15	3007 ± 15
$2\nu_2$	DVR	3284	3308						

					$(\mathbf{H_2O})_{\mathbf{17b}}$				
	r_e (Å)	0.9924	0.9891						
	θ_e (°)	104.68	104.94						
	$r^d_{\mathrm{O-O}}$ (Å)	2.7053	2.7044						
ν_2	ω	1691	1701						
	ω^{est}		1697 (-4)	1701 ± 4	1701 ± 4	1703 ± 4	1711 ± 4	1715 ± 4	1715 ± 4
	DVR	1661	1668						
	DVR^{est}		1665 (-3)	1668 ± 3	1667 ± 3	1675 ± 3	1681 ± 3	1683 ± 3	1683 ± 3
ν_1	ω	3313	3334						
	ω^{est}		3331 (-3)	3349 ± 3	3352 ± 3	3338 ± 3	3361 ± 3	3382 ± 3	3386 ± 3
	DVR	3003	3030						
	DVR^{est}		3032 (2)	3046 ± 2	3049 ± 2	3035 ± 2	3072 ± 2	3088 ± 2	3092 ± 2
ν_3	ω	3345	3358						
	ω^{est}		3355 (-3)	3373 ± 3	3376 ± 3	3358 ± 3	3373 ± 3	3393 ± 3	3398 ± 3
	DVR	3010	3033						
	DVR^{est}		3031 (-2)	3046 ± 2	3049 ± 2	3032 ± 2	3059 ± 2	3077 ± 2	3081 ± 2
$2\nu_2$	DVR	3313	3325						
					$(\mathbf{H_2O})_{\mathbf{17c}}$				
	r_e (Å)	0.9841							
	θ_e (°)	105.55							
	$r^d_{\mathrm{O-O}}$ (Å)	2.9297							
ν_2	DVR	1638							
	DVR^{est}		1643 ± 7	1645 ± 7	1644 ± 7	1652 ± 7	1658 ± 7	1661 ± 7	1660 ± 7
ν_1	DVR	3188							
	DVR^{est}		3217 ± 33	3231 ± 33	3233 ± 33	3220 ± 33	3256 ± 33	3273 ± 33	3276 ± 33
ν_3	DVR	3280							
	DVR^{est}		3301 ± 11	3315 ± 11	3318 ± 11	3301 ± 11	3328 ± 11	3346 ± 11	3350 ± 11
$2\nu_2$	DVR	3292							

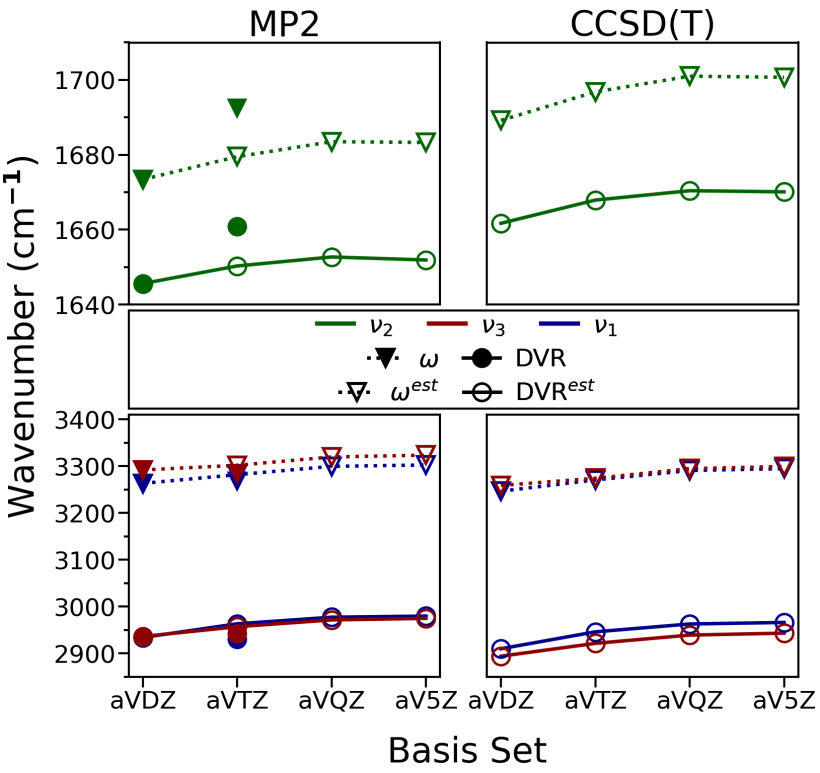

Figure 3.6: Harmonic (triangles), VPT2 (squares) and DVR (circles) vibrational frequencies (cm^{-1}) ν_1(blue), ν_2 (green), and ν_3 (red) of the central H_2O inside the fully tetrahedrally coordinated two solvation shells in *tet-a*-17 at the MP2 (left) and CCSD(T) (right) levels of theory. Filled/open symbols correspond to the calculated/estimated values.

As regards the vibrational frequencies, the second shell primarily affects the O–H stretching fundamentals, producing very large red-shifts into the 3000 cm^{-1} region, while the bend stiffens only moderately and continues its monotonic blue shift relative to the gas phase monomer. In particular, the near-degeneracy of the two stretches is observed at both the harmonic and DVR treatments. For *tet-a*-17 at the MP2/aVDZ level, the harmonic bend increases to 1673 cm^{-1}, corresponding to $\Delta\nu_2 = 51$ cm^{-1} with respect to the gas phase monomer and to a much smaller shift of $\Delta\nu_2 = 3$ cm^{-1} with respect to *tet*-5; the corresponding shifts at the MP2/aVTZ level amount to $\Delta\nu_2 = 64$ cm^{-1} (gas phase monomer), and 17 cm^{-1} (*tet*-5).

The reduced dimensional DVR predicts a larger blue shift for the bend: at MP2/aVDZ ν_2 increases to 1646 cm^{-1} ($\Delta\nu_2 = 74$ cm^{-1} from the monomer), and 13 cm^{-1} from *tet*-5) whereas at MP2/aVTZ it increases to 1661 cm^{-1} ($\Delta\nu_2 = 84$ cm^{-1} from the monomer and 25 cm^{-1} from *tet*-5). Therefore, while the bend's blue shift increases monotonically with additional solvation shells, the incremental step from the first to the second shell is modest compared with the much larger changes seen in the stretches. For the O–H stretches the second shell introduces a much larger incremental effect of the order of several hundred wavenumbers relative to both the monomer and *tet*-5. MP2/aVDZ produces $\Delta\nu_1/\Delta\nu_3$ harmonic shifts of $-540/-646$ cm^{-1} relative to H_2O and $-340/-395$ cm^{-1} relative to *tet-5*. Even stronger red shifts are observed with DVR, which yields $\Delta\nu_1/\Delta\nu_3 = -693/-813$ cm^{-1} from the monomer and $-461/-511$ cm^{-1} from *tet*-5. The same pattern persists at the MP2/aVTZ level, for which the harmonic red shifts relative to the monomer are $-554/-664$ cm^{-1}, while the DVR red shifts increase to $-730/-831$ cm^{-1}.

The comparison between *tet-a*-17 and *tet-b*-17 isolates the role of the ordering/packing in the second solvation shell. Although both clusters contain the same number of water molecules, relaxing the second shell in *tet-b*-17 systematically reduces the magnitude of stretch softening, consistent with its shorter O-H bond and longer donor $r^d_{\mathrm{O-O}}$ distance relative to *tet-a*-17. At the MP2/aVDZ level, the harmonic stretches are higher than those in *tet-a*-17 by $+50/+53$ cm^{-1}, whereas the corresponding anharmonic DVR stretches are higher by $+69/+74$ cm^{-1}. The difference of the stretches at the MP2/aVTZ level are moderately larger, viz. $\Delta\nu_1/\Delta\nu_3$ for *tet-b*-17 relative to *tet-a*-17 amount to $+66/+74$ cm^{-1} (harmonic) and $+100/+91$ cm^{-1} (DVR).

Importantly, the trend in the bend is different. Despite the weaker stretch softening, *tet-b*-17 exhibits a *larger* bend blue harmonic shift of 18 cm^{-1} at the MP2/aVDZ (8 cm^{-1} at the MP2/aVTZ) level relative to *tet-a*-17. The corresponding blue shifts at the DVR level are 15 cm^{-1} (MP2/aVDZ) and 7 cm^{-1} (MP2/aVTZ). This apparent oddity reflects a partial decoupling between the physical drivers of the bend and stretch manifolds: the stretch frequencies track hydrogen bond strength and local field along the O–H direction

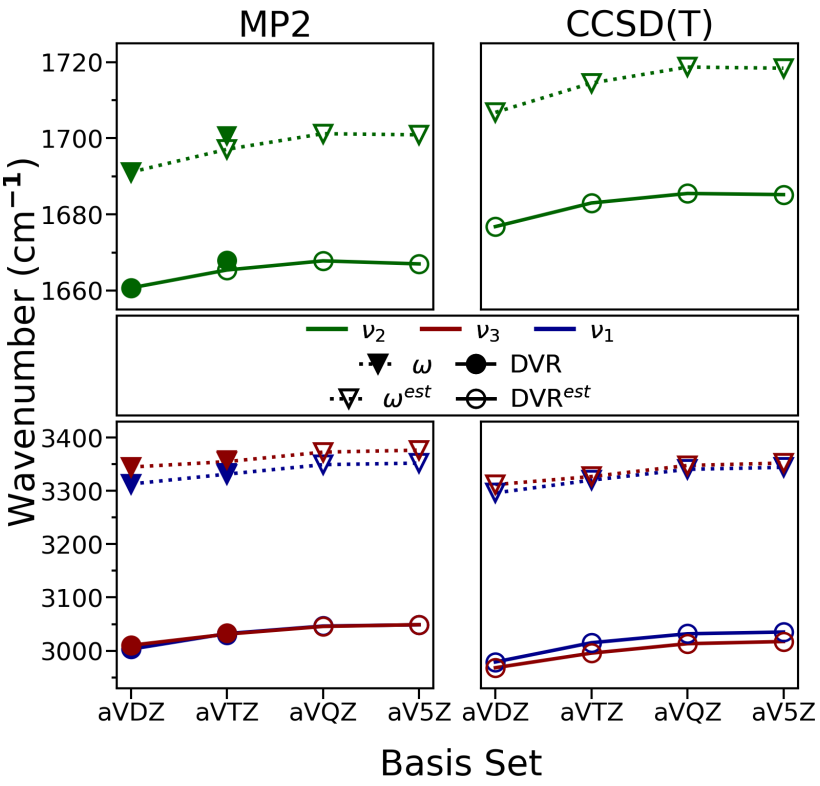

Figure 3.7: Harmonic (triangles), VPT2 (squares) and DVR (circles) vibrational frequencies (cm^{-1}) ν_1(blue), ν_2 (green), and ν_3 (red) of the central H_2O inside the tetrahedrally coordinated (first shell) / relaxed (second shell) in *tet-b*-17 at the MP2 (left) and CCSD(T) (right) levels of theory. Filled/open symbols correspond to the calculated/estimated values.

(well captured by r^d_{O-O}), whereas the bend is more sensitive to angular confinement and collective steric/electrostatic constraints imposed by the first shell tetrahedral framework. Thus, relaxing the second shell can weaken the donor-directed field (less stretch red shift) while still maintaining (or even enhancing) the angular constraints that stiffen the bend.

The observed large red shifts support the interpretation in the context of ice Ih rather than simply as deviations from the liquid water behavior. In crystalline ice Ih the H–O–H bending band lies near 1650 cm^{-1},[136, 137, 138] while the O–H stretching manifold is strongly structured and collective, with the IR spectrum dominated by a band centered

at 3220 cm^{-1}, a shoulder near 3150 cm^{-1} and another shoulder at 3380 cm^{-1}.[137, 138]. The Raman spectrum has O–H stretching frequencies appearing at 3083 cm^{-1}, 3310 cm^{-1}, 3320 cm^{-1}.[138, 139, 140, 141, 142] In this sense, the ν_1 and ν_3 DVR fundamentals of *tet-a*-17 and *tet-b*-17, which fall in the $\sim$ 2930–3035 cm^{-1} region, are more reminiscent of strongly hydrogen bonded ice-like local environments rather than of ambient liquid water.

An additional spectral hallmark of these two networks mimicking the second solvation shell is that the H–O–H bending overtone, $2\nu_2$, is pulled directly into the O–H stretching manifold. In *tet-a*-17 and *tet-b*-17 the DVR $2\nu_2$ lies near 3300 cm^{-1}, meaning that it is no longer spectroscopically isolated as it lies inside the broad condensed phase O–H stretch envelope. As the hydrogen bond network red-shifts, the stretch fundamentals move downward towards the 3000–3400 cm^{-1} region becoming near degenerate with the $2\nu_2$ overtone. This enables a strong Fermi resonance mixing, prominently between the nominal symmetric stretch ν_1 and the $2\nu_2$ overtone. In liquid water this near-degeneracy allows the overtone to borrow intensity from bright stretching transitions and to contribute to the observed O–H stretch band shape rather than appearing as a weak overtone feature.[143, 144, 145] In the present network hierarchy, the emergence of $2\nu_2$ within the stretch window provides a direct microscopic analog of the condensed phase assignment problem: the vibrational eigenstates in the 3200–3400 cm^{-1} region acquire mixed bend–stretch character, so interpreting the stretch band purely in terms of local O–H fundamentals becomes increasingly challenging and incomplete.[143]

The estimated frequencies from Eq. 3.5 and Eq. 3.6 are listed in Table 3.8. The MP2/aVTZ estimates for *tet-b*-17 track the calculated values to within just 4 cm^{-1}, indicating that the monomer-derived estimates remain a practical uncertainty range even for a two solvation shell network. For the more ordered *tet-a*-17 network, our best estimates show larger deviations, amounting to 11–12 cm^{-1} smaller for the bend and 15–33 cm^{-1} larger for the stretches, compared to the calculated values. This is consistent with basis set convergence becoming increasingly environment-dependent as the hydrogen bond network strengthens. Nevertheless the uncertainty in the estimated fundamentals remains minimal compared to the pronounced

solvation-induced red-shifts observed for the stretching modes, often spanning several hundred cm^{-1}. As a result, the overall qualitative conclusion stands firm as the second solvation shell plays a pivotal role in driving O–H stretching frequencies closer to those found in bulk environments, while at the same time it perpetuates the steady blue shift of the bending mode relative to the gas phase monomer.

3.3.6 The Third Solvation Shell of the Monomer: $(\mathbf{H_2O})_{53}$

Table 3.9: Geometric parameters and DVR anharmonic (ν) vibrational frequencies (cm^{-1}) of a H_2O molecule embedded in three solvation shells in the $(H_2O)_{53}$ network with tetrahedrality enforced through the second shell and a relaxed third shell. The DVR fundamentals were computed on a $9 \times 9 \times 5$ grid at the MP2/aVDZ level. The estimated values (^{est}) with the larger basis sets and at the CCSD(T) level of theory and their uncertainties (see text) are listed.

		MP2				**CCSD(T)**			
		aVDZ	aVTZ	aVQZ	aV5Z	aVDZ	aVTZ	aVQZ	aV5Z
r_e (Å)		0.9841							
θ_e (°)		105.55							
r^d_{O-O} (Å)		2.9297							
ν_2	DVR	1638							
	DVR^{est}		1642 ± 7	1645 ± 7	1644 ± 7	1652 ± 7	1658 ± 7	1660 ± 7	1660 ± 7
ν_1	DVR	3187							
	DVR^{est}		3216 ± 20	3230 ± 20	3233 ± 20	3220 ± 20	3256 ± 20	3273 ± 20	3276 ± 20
ν_3	DVR	3287							
	DVR^{est}		3307 ± 11	3322 ± 11	3325 ± 11	3308 ± 11	3335 ± 11	3353 ± 11	3357 ± 11
$2\nu_2$	DVR	3292							

The $(H_2O)_{53}$ cluster (*tet*-53) mimicks the third solvation shell of a water molecule in the adopted network hierarchy. Consistent with the construction of the smaller tetrahedral motifs, tetrahedral coordination was enforced through the second shell while allowing the third shell to relax, producing a more bulk-like outer environment around an embedded monomer that

retains a locally ice-like two solvation shell network (see Fig. 3.1). The optimized structural parameters, the calculated DVR fundamentals of the central monomer at the MP2/aVDZ level, and their estimated values with larger basis sets and higher correlation levels of theory [CCSD(T)] are listed in Table 3.9.

The central monomer in *tet*-53 remains substantially distorted relative to the gas phase monomer but, importantly, the distortion pattern differs from the smaller, more rigid tetrahedral networks *tet-a*-17 and *tet-b*-17. At MP2/aVDZ, the central molecule's O–H bond length (r_e) is 0.9841 Å and the H–O–H angle (θ_e) is 105.55°, corresponding to shifts of $\Delta r_e = +18.2$ mÅ and $\Delta\theta_e = +1.68°$ relative to the gas phase monomer. However, when comparing to the two solvation shell networks, the central bond is *shorter* ($\Delta r_e = -10.9$ mÅ vs. *tet-a*-17 and -8.3 mÅ vs. *tet-b*-17), while the angle falls between the two ($\Delta\theta_e = -0.64°$ vs. *tet-a*-17 and $+0.87°$ vs. *tet-b*-17). This reduced intramolecular distortion is consistent with the markedly longer donor hydrogen bond separation: the characteristic donor r^d_{OO} increases from 2.6752 Å in *tet-a*-17 and 2.7053 Å in *tet-b*-17 to 2.9297 Å in *tet*-53 (MP2/aVDZ), indicating a weaker direct hydrogen bond constraint at the central molecule once the outer network is allowed to relax.

Relative to the gas phase monomer, *tet*-53 exhibits the expected qualitative fingerprint of solvation. Quantitatively, *tet*-53 shifts the bend by $\Delta\nu_2 = +66$ cm^{-1} while softening the symmetric and antisymmetric stretches by $\Delta\nu_1 = -439$ cm^{-1} and $\Delta\nu_3 = -462$ cm^{-1} compared to the MP2/aVDZ monomer DVR values. Comparing *tet*-53 to the previous two solvation shell *tet-a*-17 and *tet-b*-17 motifs, highlights a key distinction between extending tetrahedral constraints and adding more distant solvent degrees of freedom. The $(H_2O)_{17}$ networks already produce strongly red-shifted stretches in the ~ 3000 cm^{-1} range, whereas *tet*-53 yields higher stretch fundamentals ($\nu_1 = 3187$ cm^{-1} and $\nu_3 = 3287$ cm^{-1}), consistent with a weaker direct donor hydrogen bond in the relaxed third solvation shell environment as evidenced by the longer $r^d_{\mathrm{O-O}}$ of 2.9297 Å. To put it another way, when a tightly constrained tetrahedral network is enforced – specifically through *tet-a*-17 – the embedded monomer is pushed toward pronounced hydrogen bonding, resulting in stretching frequencies reminiscent

of those found in ice. By permitting further relaxation in the outer shell, *tet*-53 alleviates some of the most intense short-range restrictions. At the same time, it enhances overall polarization and maintains a substantial net red shift compared to the gas phase monomer. This nuanced interplay highlights how modifications in the network structure can profoundly influence both its bonding environment and associated vibrational shifts..

Finally, in order to isolate whether the third solvation shell relaxation materially affects the embedded monomer fundamentals, we repeated the DVR calculation on a truncated model consisting only of the first two solvation shells (*tet-c*-17) extracted from the *tet*-53 geometry. The DVR results of this two solvation shell network in the *tet*-53 are within 7 cm^{-1} from the ones from the full three solvation shell *tet*-53 results. This small difference indicates that the fundamental frequencies of the embedded water monomer are essentially converged at the second tetrahedrally coordinated solvation shell, a fact that is consistent with the results previously obtained with classical potentials.[91, 146] The comparison with *tet-c*-17 further emphasizes that the vibrational shifts are not determined by cluster size alone. Although *tet-c*-17 retains two nominal solvation shells, its donor O–O distance is much longer than in *tet-a*-17 and *tet-b*-17, moving it away from the compressed ice-like structural regime. Correspondingly, its stretching fundamentals shift upward toward the higher frequency portion of the ice/liquid overlap region, illustrating that long-range tetrahedral order by itself is not sufficient but the local donor geometry, especially the coupled evolution of r_{OH} and $r^{d}_{\mathrm{O-O}}$, is what controls the magnitude of the red shift in the O–H stretches.

3.3.7 Convergence Toward Bulk-like Properties

The best estimated CCSD(T)/CBS fundamentals, $\nu_{\mathrm{cluster}}^{\mathrm{CCSD(T)/CBS}}$, for gas phase and the embedded water monomer for the various networks $n = 1, 2, 5, 8, 17$, and 53 are summarized in Table 3.10 and Fig. 3.8. The shaded regions indicate the approximate full-width-at-half-maximum (FWHM) ranges of the corresponding liquid water absorption bands, while the dashed lines mark the experimental band centers, viz. the H–O–H bending mode near 1642 cm^{-1} and the broad O–H stretch envelope centered near 3405 cm^{-1}.[148, 149, 143] The

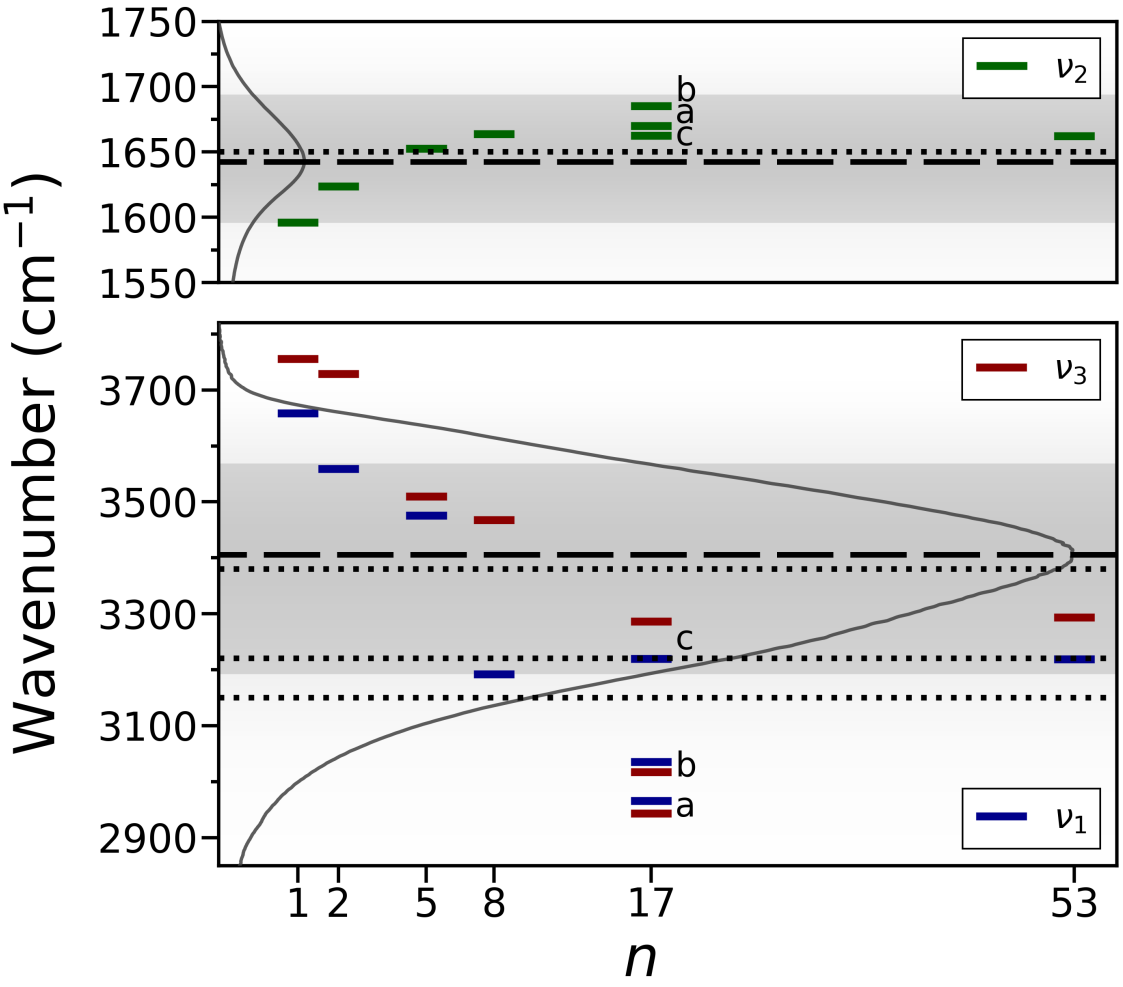

Figure 3.8: Best estimates $\nu_{\text{cluster}}^{\text{CCSD(T)/CBS}}$ for the vibrational fundamentals of $(H_2O)_n$, n = 1, 2, 5, 8, 17a, 17b, 17c, 53. The shaded regions indicate the full-width-at-half-maximum (FWHM); the dashed lines mark the band centers of the IR bands of liquid water in the bending (1642±49 cm^{-1}) and O–H stretching (3405±188 cm^{-1}) regions at 25°C;[147] the dotted lines trace the experimental peaks of the IR spectra of Ice I at 100 K,[137, 138] viz. the bending at 1650 cm^{-1} and the three main features of the O–H stretching band at 3150 cm^{-1} (shoulder), 3220 cm^{-1} (main band), and 3380 cm^{-1} (shoulder).

corresponding main features of the ice Ih IR spectra[137, 138] are indicated with dotted lines.

The uncertainties in the estimated frequencies were determined empirically from the observed errors of the additive estimation scheme wherever direct comparisons were available. For each mode and method, the uncertainty was taken as the largest absolute deviation between estimated and explicit values at the lower basis levels, providing a conservative measure of the error due to transferability. When no direct comparison was possible, the uncertainty was transferred from the closest calibrated motif. The uncertainty of both *tet*-*c*-17 and *tet*-53 were taken as the root mean squared of the errors from *tet*-5, *tet*-*a*-17, and *tet*-*b*-17. The reported ± values therefore represent transferred systematic uncertainties of the estimation procedure that accounts for the coarseness of the 1D DVR potential grid, theory, and basis set.

The best estimated fundamentals show a clear and mode-dependent convergence toward the condensed phase spectral regions with increasing solvation. Fig. 3.8 shows that the bending frequency enters the experimental liquid-water FWHM window (1642±42 cm^{-1}) already at $n = 2$ and remains within the bulk band for all larger networks, while from *tet*-5 onward it is also consistent with the ice-Ih bending region around 1650 cm^{-1}. This rapid convergence is consistent with the bending mode being primarily sensitive to the local, nearest-neighbor hydrogen-bond environment, which is already largely established with the inclusion of the first solvation shell.[149]

In contrast, both O–H stretching fundamentals undergo substantially larger red shifts with increasing network size and move within the broad liquid-water O–H stretching envelope. The most dramatic change occurs for the networks mimicking the second solvation shell. The ν_1/ν_3 are strongly red-shifted to 3022/3007 cm^{-1} (*tet*-*a*-17) and 3092/3081 cm^{-1} (*tet*-*b*-17). These values are, in some cases, outside the bulk water FWHM range (see Fig. 3.8) but nonetheless within the spectral envelope of the broad O-H stretch. These values lie much closer to the main ice-Ih stretching region, bounded by the lower shoulder near 3150 cm^{-1}, the principal feature at 3220 cm^{-1}, and the upper shoulder near 3380 cm^{-1}. For the larger but more relaxed *tet*-*c*-17 and *tet*-53 networks, the stretching frequencies shift back upward to

Table 3.10: Best estimates $\nu_{\text{cluster}}^{\text{CCSD(T)/CBS}}$ from Eq. 3.5, Eq. 3.6 and DVR results from classical potentials for the anharmonic (ν) vibrational frequencies (cm^{-1}) of the $(H_2O)_n$ networks. The experimental results for the liquid are indicated as the peak of the vibrational band and the full-width-at-half-maximum (FWHM). The prominent features of the IR spectra for ice Ih, including approximate values for the shoulders (sh.), are also indicated.

		H_2O	$(H_2O)_2$	*tet*-5	*tet*-8	*tet-a*-17	*tet-b*-17	*tet-c*-17	*tet*-53	Exp. (liquid)	Exp. (ice)
	$\nu_{\text{cluster}}^{\text{CCSD(T)/CBS}}$	1596	1624±2	1655±3	1662±2	1668±11	1683±3	1660±7	1660±7		
	TTM2.1-F	1594	1623	1616	1625	1631	1628	1632	1612		
ν_2	TTM3-F	1594	1614	1577	1582	1573	1579	1569	1573	1642±49[147]	1650±30[137]
	MB-pol	1594	1618	1638	1653	1670	1685	1650	1656		
	q-AQUA	1594	1622	1652	1656	1662	1668	1668	1660		
	$\nu_{\text{cluster}}^{\text{CCSD(T)/CBS}}$	3658	3569±10	3483±7	3249±34	3022±33	3092±2	3276±20	3276±20		
	TTM2.1-F	3656	3587	3595	3594	3591	3591	3557	3574		
ν_1	TTM3-F	3656	3413	3227	3191	2988	3057	3161	3181		
	MB-pol	3656	3556	3467	3218	3095	3123	3373	3395		3150 (sh.)[137, 138]
	q-AQUA	3656	3556	3492	3252	3148	3253	3433	3424		
	$\nu_{\text{cluster}}^{\text{CCSD(T)/CBS}}$	3756	3725±1	3517±8	3532±7	3007±15	3081±2	3350±11	3357±11	3405±188[147]	3220[137, 138]
	TTM2.1-F	3755	3740	3660	3659	3635	3637	3611	3636		3380 (sh.)[137, 138]
ν_3	TTM3-F	3755	3741	3295	3284	3082	3161	3202	3233		
	MB-pol	3755	3728	3492	3498	3110	3138	3377	3404		
	q-AQUA	3755	3725	3507	3537	3159	3262	3436	3446		

3276/3350 and 3276/3357 cm^{-1}, respectively, while remaining on the red side (lower values) of the liquid band center (3405 cm^{-1}).

This non-monotonic behavior suggests that intermediate size tetrahedral networks can over stabilize unusually strong local hydrogen bonding around the embedded monomer, leading to larger red shifts than those characteristic of either bulk liquid water or typical ice Ih local environments, whereas larger networks recover additional structural flexibility and heterogeneity that partially relax the local environment toward more bulk- and ice-like distributions.[148, 91, 150, 143] It should be noted that the direct comparison with the absorption bands of bulk water should be interpreted qualitatively because the experimentally measured O–H stretching envelope reflects inhomogeneous broadening and vibrational coupling effects (including stretch–bend Fermi resonance contributions), whereas the present values correspond to discrete fundamentals of a specific embedded structural motif.[143, 149]

Table 3.11: Comparison between the full dimensional (with deuteration scheme) harmonic frequency ω and the approximate **FG** harmonic frequency ω' of the central monomer in $(H_2O)_n$ where n = 1, 5, 17a, 17b, and 53.

			H_2O	*tet*-5	*tet-a*-17	*tet-b*-17	*tet*-53
ν_2	MP2/aVDZ	ω	1622.3	1670.4	1673.4	1691.1	
		ω'	1622.3 (0.0)	1679.4 (6.0)	1686.4 (13.0)	1704.5 (13.4)	1682.8
	CCSD(T)/aVDZ	ω	1638.1	1682.5			
		ω'	1634.5 (-3.6)	1687.5 (5.0)			
ν_1	MP2/aVDZ	ω	3803.3	3602.9	3263.0	3312.8	
		ω'	3803.3 (0.0)	3603.5 (0.6)	3256.0 (-7.0)	3307.6 (-5.2)	3469.9
	CCSD(T)/aVDZ	ω	3786.7	3627.7			
		ω'	3786.8 (0.1)	3626.0 (-1.7)			
ν_3	MP2/aVDZ	ω	3937.5	3686.9	3291.6	3344.6	
		ω'	3937.5 (0.0)	3684.5 (-2.4)	3276.5 (-15.1)	3332.3 (-12.3)	3549.0
	CCSD(T)/aVDZ	ω	3904.6	3699.9			
		ω'	3904.8 (0.2)	3695.5 (-4.4)			

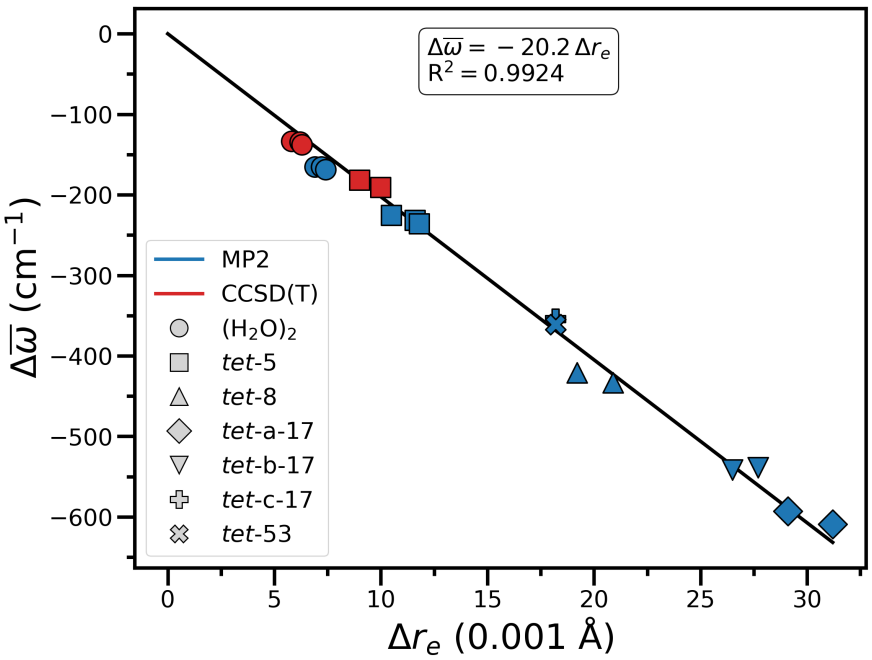

Figure 3.9: Badger's rule correlation between $\Delta\overline{\omega}$ and Δr_e of the solvated monomer in $(H_2O)_n$. The $\Delta\overline{\omega}$ are the shifts with respect to the average of the harmonic frequency of the symmetric and antisymmetric stretch of the gas phase monomer; for $(H_2O)_2$ and $(H_2O)_8$ the shift is with respect to $\nu(OH_b)$. Likewise, Δr_e are the shifts with respect to the equilibrium bond length of the gas phase monomer.

A complementary view of the convergence toward bulk-like behavior is provided by the correlation between the O–H bond elongation and the harmonic stretching red shift of the embedded monomer, a structural-spectroscopic hallmark originally reported by Badger and Bauer[151] and later extended to hydrogen bonded systems.[113] Previous work by Boyer *et al.* quantitatively determined that hydrogen bonding induces a shift of 19±1 cm^{-1} in the O–H stretching frequency per 0.001 Å elongation of the O–H bond length.[152] We performed the Wilson **GF** (or sometimes referred to as **FG**) method[153, 83] in the internal coordinates (r_1, r_2, θ) as an alternate method to approximate the harmonic vibrational frequencies (ω') of *tet*-53. The force-constant matrix **F** was constructed using central finite differences on a cubic

grid centered at the minimum. The **G** maxtrix was evaluated at the equilibrium geometry, analogous to the G matrix element introduced in Eq. 3.1. The displacement around the minimum is $\delta r = 0.005$ bohr and $\delta\theta = \delta r/r_e$. The ω' results for the $n = 1$, 5, 17a, 17b, and 53 networks, reported in Table 3.11 for comparison with the available values for ω, produce a global fit of $\Delta\overline{\omega} = -20.2\,\Delta r_e$ ($R^2 = 0.9924$), which lies within the error bar of the previously reported correlation (see Fig. 3.9 where we used $\nu(\mathrm{OH_b})$ for the dimer and *tet*-8. Although the present correlation is based on our selected local minimum structures with central H_2O surrounded by perdeuterated D_2O molecules, rather than the global minimum structures in that earlier study, the fitted slope is within the error bar of their reported value. The fact that both MP2 and CCSD(T) data collapse onto essentially the same line suggests that this structural–spectral correlation is robust and largely independent of the level of electronic structure theory, even though MP2 generally predicts slightly larger bond elongations and correspondingly larger red-shifts than CCSD(T).[77]

3.3.8 Results with classical potentials

The methodology employed to determine the anharmonic frequencies of solvated water within diverse cluster networks using electronic structure approaches – including geometry optimization with C_2 (C_s for solvated dimer) and tetrahedral constraints, PO–DVR, and grid techniques – was likewise utilized with the TTM2.1-F, TTM-3F, MB-pol and q-AQUA classical interaction potentials. The DVR fundamentals obtained with the $10 \times 10 \times 10$ size 3D PO–DVR grid with the four classical potentials, listed in Table 3.10, clearly demonstrate that the predicted frequencies are strongly controlled by the choice of the 1-body intramolecular PES and DMS. All four potentials reproduce the monomer fundamentals essentially as expected, but their behavior begins to diverge once the water molecule is embedded in a hydrogen bond network. In particular, the table shows that the three modes respond differently within the same model, indicating that agreement for one mode is not sufficient to guarantee a balanced description of the intramolecular potential in a condensed or cluster environment.

With respect to the $\nu_{\text{cluster}}^{\text{CCSD(T)/CBS}}$ best estimates, MB-pol has a mean-absolute-error of 31 cm^{-1}, followed by q-AQUA (49 cm^{-1}), TTM3-F (92 cm^{-1}), and TTM2.1-F (183 cm^{-1}). For the bending fundamental ν_2, q-AQUA and MB-pol perform best, with MAEs of 5 and 7 cm^{-1}, respectively, and both remain close to the condensed-phase bending region, near the experimental values 1642±49 cm^{-1} for liquid water and 1650±30 cm^{-1} for ice Ih.[147, 137] In contrast, TTM2.1-F systematically underestimates ν_2, while TTM3-F underestimates it even more strongly. For the stretches, the picture is more nuanced. TTM2.1-F performs poorly, with MAEs of 266 and 251 cm^{-1} for ν_1 and ν_3, respectively, because it systematically underestimates the hydrogen-bond-induced red-shifts and keeps both stretching fundamentals much too high across the solvated clusters. These values are also too high relative to the main ice-Ih stretching features at 3150 (sh.), 3220, and 3380 (sh.) cm^{-1},[137, 138] and they remain on the high-frequency side of the broad liquid stretching band centered at 3405±188 cm^{-1}.[147] TTM3-F is substantially better for the stretches, especially in the strongly solvated and ice-like motifs, where it captures the large red-shifts in better agreement with our best estimates than TTM2.1-F; its MAEs are 94 cm^{-1} for ν_1 and 114 cm^{-1} for ν_3. However, this improved stretch behavior comes at the expense of the bend that is too low, and not uniform for all clusters. These differences can be attributed to the use of the original, gas phase PS (TTM2.1-F) versus an effective, liquid-like DMS (TTM3-F) for the 1-body term.[154]

MB-pol and q-AQUA provide a more balanced overall description. MB-pol is associated with MAEs of 47.8 and 37 cm^{-1} for ν_1 and ν_3, respectively, and performs well for the smaller and intermediate networks such as *tet*-5 and *tet*-8. Q-AQUA also describes the bends well, however it produces somewhat larger errors than those of MB-pol, with MAEs of 77 cm^{-1} for ν_1 and 66 cm^{-1} for ν_3. Both MB-pol and q-AQUA nonetheless overestimate the stretches for the most strongly solvated clusters, especially *tet-a*-17 and *tet-b*-17, and neither reproduces their near degeneracy. Relative to experiment for liquid water and ice, both models place the stretches for the larger cluster in the broad condensed phase envelope,[147, 137, 138]. The previous results indicate that even a highly accurate gas phase 1-body term is not fully transferable to an embedded water molecule in a heterogeneous hydrogen bond network. To

this end, a more realistic description should be based on an environment-dependent 1-body term, so that the intramolecular potential of a solvated monomer can respond explicitly to its local solvation environment (*vide infra*).

3.3.9 O–H Bond Dissociation Channels

As previously pointed out by Fanourgakis and Xantheas during the development of the TTM3-F potential,[154] the PS PES, being a gas phase monomer PES, dissociates homolytically along a stretched OH coordinate to $OH^{\bullet}$ + $H^{\bullet}$ radicals, viz. to neutral fragments. This means that upon stretching the O–H bond the charge on the hydrogen atom goes to zero ($q_{\mathrm{H}} \rightarrow 0$). However, in a fully connected hydrogen bonded condensed phase environment, the appropriate dissociation channel is the heterolytic one, occurring via proton transfer and leading to the OH^- and H_3O^+ ions, which are further stabilized by the surrounding solvent; in other words, for the solution-based PES $q_{\mathrm{H}} \rightarrow +1$ as the O–H bond elongates. A 1-body term based purely on the gas phase (i.e., the PS) PES, cannot therefore capture the correct long-range behavior of highly stretched OH bonds in solution or in clusters, precisely in the regime that is most important for describing autoionization, proton transfer, and extreme hydrogen bond distortions. This realization motivated the development of an *effective* monomer 1-body term used in the TTM3-F potential, in which the charge on the hydrogen atom increased linearly upon elongation of the O–H bond of a solvated water. In the TTM3-F monomer DMS the effective charges $\tilde{q}^i$, where $i = \mathrm{H}_1$, H_2, O, are related to the charges q^i of the PS DMS according to[154]

$$\begin{aligned} \tilde{q}^{H_1} &= q^{H_1} + d_r \left(r_{OH_1} - r_e\right) + d_\theta \left(\theta - \theta_e\right), \\ \tilde{q}^{H_2} &= q^{H_2} + d_r \left(r_{OH_2} - r_e\right) + d_\theta \left(\theta - \theta_e\right), \\ \tilde{q}^{O} &= -\tilde{q}^{H_1} - \tilde{q}^{H_2} \end{aligned} \tag{3.7}$$

where r_e and θ_e are the equilibrium monomer bond length and bend angle, and $d_r = 0.5$ e/Å, $d_\theta = 0.012$ e/rad were *ad hoc* parameters. The PESs of the model networks presented in this

study offer the opportunity to revisit this issue, seeking a more quantitative description of the 1-body term of water in a condensed-like environment.

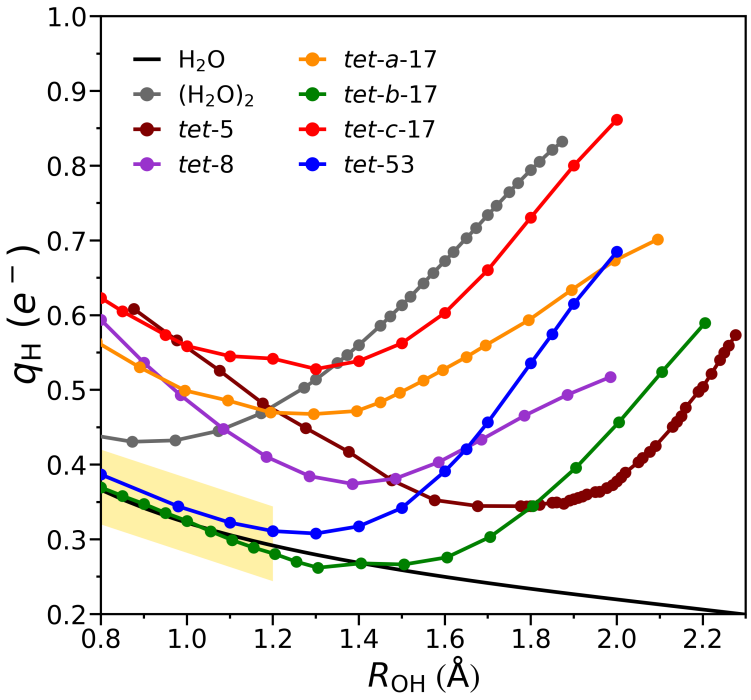

Figure 3.10: Modified CHELPG charge on hydrogen, q_H, as a function of the O–H bond length, R_{OH}, for $(H_2O)_n$, n = 1, 2, 5, 8, 17a, 17b, 17c, 53. The second R_{OH} and the angle H–O–H are fixed at their equilibrium values. The shaded area indicates the range of O–H stretches sampled during a molecular dynamics simulation of liquid water.

Fig. 3.10 shows the computed CHELPG[155] charge of one of the hydrogen atoms of the central water molecule (q_H) upon elongation of the O–H bond at the MP2/aVDZ level for the studied hydrogen bonded networks of increasing sizes and coordination. We modified the CHELPG charge scheme to constrain the total charge of the central, embedded water monomer to be zero and fit to the total *ab initio* dipole moment of the cluster. The results of Fig. 3.10 trace the quantitative response of the charge on the hydrogen atom of the central water molecule to the O–H bond elongation in the various solvation shells. In contrast to the

homolytic dissociation of the isolated monomer, for which $q_H \rightarrow 0$, the stretched O–H bond in the larger two- and three-solvation shell networks eventually enters a regime in which q_H increases with R_{OH}, indicating the onset of a heterolytic, proton-transfer-like dissociation channel in which the departing hydrogen atom acquires increasingly protonic character ($q_H \rightarrow +1$).

This behavior is fully consistent with the structural and vibrational trends discussed earlier. The clusters that exhibit substantial elongation of the equilibrium O–H bond and strong red-shifts of the OH stretches are also the ones for which the charge on the hydrogen on the stretched bond clearly departs from the gas phase monomer limit. In this sense, the charge analysis provides an *ab initio* interpretation of the same solvation effects already reflected in the equilibrium geometries and vibrational frequencies. The onset and magnitude of the upward charge evolution depend strongly on the local solvation environment. A quantitative analysis of the approximately linear increase in q_H at large R_{OH} is given in Fig. 3.11. The fitted slopes span a broad range, with the fully tetrahedrally coordinated clusters clearly separated from the hybrid clusters, especially the three *tet*-17 motifs exhibit different charge behavior despite having the same cluster size.

The CHELPG charges of the solvated monomer in the various networks (Fig. 3.10) make it possible to compare the fitted embedded-monomer dipole moment, μ (Debye, D), directly with earlier experimental and theoretical estimates of the dipole moment of a single water molecule in condensed phases (note that a quantum mechanical operator that yields the dipole moment of a single water molecule in an ensemble via the system's wavefunction does not exist). The gas phase dipole moment of water has been determined experimentally and theoretically to be $\mu = 1.85$ D,[156, 157, 158, 159] whereas experimental effective dipole moment of liquid water suggests is 2.9 ± 0.6 D, together with an effective charge transfer of about $0.5e$ along each O–H bond.[160] More recent theoretical work has emphasized that the monomer's dipole moment in liquid water is not uniquely defined.[90] In particular, Liu, Wang, and Bowman reported a broad liquid water distribution with an average near 2.94 D, while for ice Ih they found still larger values, with average dipoles of 3.54 D for core

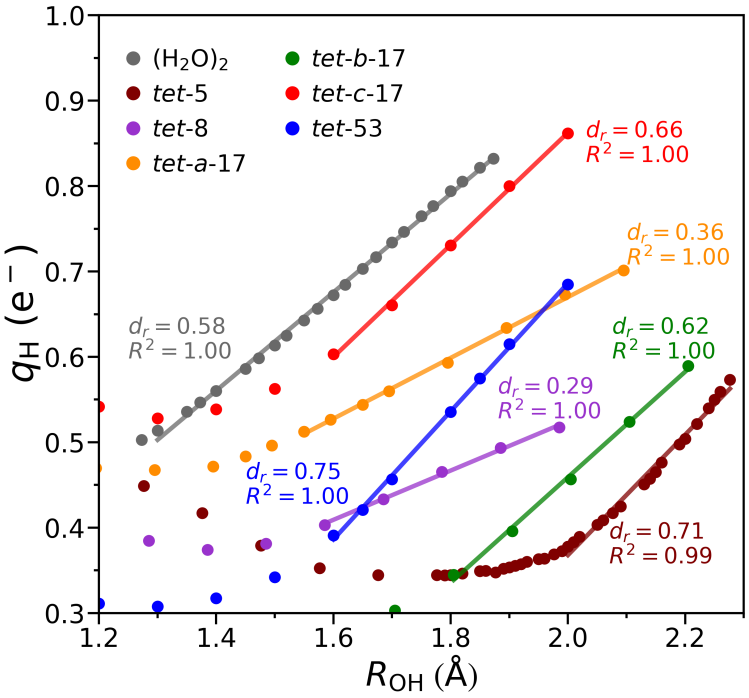

Figure 3.11: Fitted slope and R^2 values on the linearly increasing range of the modified CHELPG charge on hydrogen, q_H, as a function of the O–H bond length, R_{OH}, for $(H_2O)_n$, n = 1, 2, 5, 8, 17a, 17b, 17c, 53. The second R_{OH} and the angle H–O–H are fixed at their equilibrium values.

molecules and 3.02 D for surface molecules.[90] Similar values have also been reported in earlier theoretical studies of ice and extended hydrogen-bonded water networks.[161, 162, 163] Across the present set of clusters, the solvated monomer dipole moments span a broad range as shown in Fig. 3.12. Near the equilibrium region, *tet*-*b*-17 and *tet*-53 remain comparatively weakly polarized and close to the gas phase value, whereas the dimer, *tet*-8, and *tet*-*a*-17 fall in or near the liquid-water window, while *tet*-5 lies around 3.2 D. This pattern suggests that relaxation of the outer solvation shell plays a major role in determining the polarization of the embedded monomer.

The question naturally arises why classical potentials such as TTM2.1-F, MB-pol and

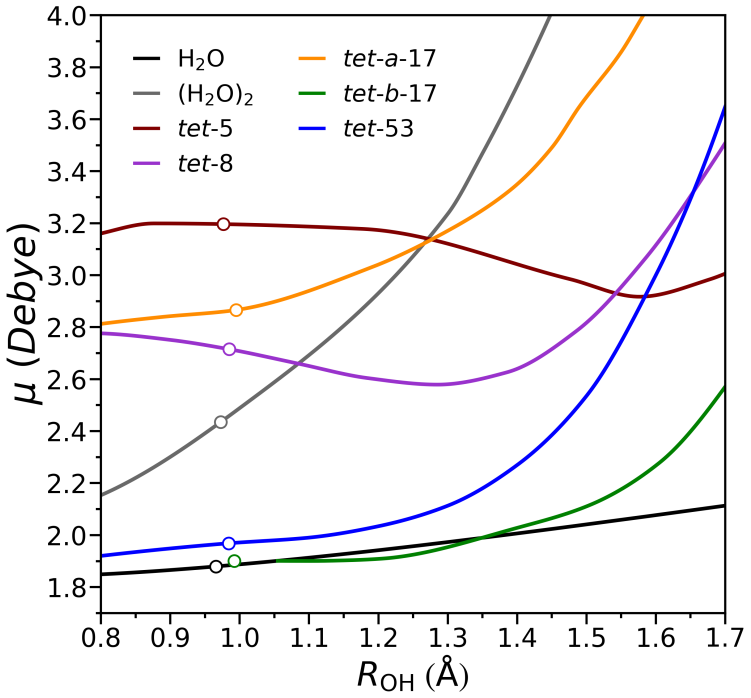

Figure 3.12: Dipole moment μ (Debye) of the solvated water monomer from the modified CHELPG charge, as a function of the O–H bond length, R_{OH}, for $(\mathrm{H_2O})_n$, $n = 1$, 2, 5, 8, 17a, 17b, 53. Open circles mark the location of the minimum for each case. The second R_{OH} and the angle H–O–H are fixed at their equilibrium values.

q-AQUA, which are based on the physically inaccurate (as regards the long range part) PS 1-body DMS associated with the homolytic cleavage of the O–H bond, produce reasonable vibrational spectra for liquid water? The shaded region in Fig. 3.10, which marks the range of O–H distances sampled during a molecular dynamics simulation of liquid water (see for example Fig. S1 in the SI of Ref. 69), is particularly revealing. Over this range of approximately 0.8–1.2 Å, the *tet*-53 and *tet-b*-17 curves remain very close to and nearly parallel to the PS gas phase monomer curve. Within the bond length range typically sampled during molecular dynamics simulations of liquid water, even strongly solvated environments exhibit a charge response to O–H elongation that closely resembles that of an isolated

monomer. Significant deviations from this monomer-like behavior only occur at larger O–H distortions, which are not encountered during standard MD simulations of liquid water with non-dissociative classical interaction potentials. This finding is important in the context of widely used many-body water models such as MB-pol, and q-AQUA, all of which employ the PS 1-body term.[20, 164, 165] The near-parallel behavior of the *tet*-53 and *tet-b*-17 curves to the monomer in the liquid-like window explains why the PS-based 1-body terms can still perform well for many equilibrium and near-equilibrium condensed-phase properties, even though they are not designed to reproduce the large R_{OH} heterolytic regime. However, this will not be the case for low-coordinated networks over a range of different O–O nearest neighbor separations.

At the same time, the absolute values of the fitted charges should be interpreted with caution. CHELPG and related electrostatic potential-based charge models represent the molecular electrostatics in terms of atom centered point charges fitted to reproduce the quantum mechanical electrostatic potential.[155, 166] Such charges are not unique observables, can vary with geometry and fitting protocol, and are known to have limitations when used as a compact description of anisotropic or penetrating charge distributions, especially in condensed or embedded environments.[166, 88] These caveats are particularly relevant when discussing embedded charges derived at the MP2 electrostatic potential level: the fitted charges are best regarded as useful descriptors of the evolving electrostatics, rather than as unique physical observables.

Overall, these results show that the solvating water molecules do more than simply perturb the gas phase monomer charge distribution. Rather, they stabilize the developing charge separation along the stretched O–H bond and progressively shift the dissociation pattern away from the neutral, gas phase homolytic limit toward a solvated proton transfer regime. The present calculations provide direct microscopic support for the physical picture that motivated the TTM3-F modification of the monomer dipole surface.[154] However, the variation of q_{H} with R_{OH} follows a parabola rather than a straight line used in TTM3-F. At the same time, the broad spread in fitted d_r values (see Fig. S2 in the SI) indicates that a single

universal correction cannot fully represent the diversity of condensed phase environments having different coordination and local order. While the TTM3-F form captures the correct qualitative long range trend, the *effective* 1-body response of a stretched O–H bond is clearly environment dependent.

3.4 Conclusions and outlook

The current study provides a systematic and quantitative characterization of the evolution of the intramolecular fundamental vibrational frequencies of a water molecule as its hydrogen bond environment grows from the isolated monomer to bulk-like networks. Using high level electronic structure calculations and large basis sets to produce CCSD(T)/CBS limit estimates for the vibrational frequency motifs of a water molecule embedded in networks resembling bulk-like environments, we identify distinct convergence patterns for its bending and stretching fundamental frequencies.

The H–O–H bending mode undergoes a rapid blue shift from its monomer value upon solvation and quickly approaches the band observed in liquid water already by the first solvation shell. Once the first solvation shell forms, the bending frequency falls within the experimental liquid full-width-at-half-maximum (FWHM) range and shows only minor variation with additional solvation shells. This suggests that the bending vibration is chiefly determined by the local hydrogen bond environment surrounding the embedded monomer in the first solvation shell.

In contrast, the OH stretching fundamentals show much larger red shifts and a markedly non-monotonic convergence pattern with the size and geometry of the surrounding network. Intermediate network sizes, particularly the tetrahedrally coordinated second solvation shell $(H_2O)_{17}$ motifs, are associated with extremely strong red shifts arising from cooperative hydrogen bonding within a highly ordered local network reminiscent of ice-like structures. For the larger $(H_2O)_{53}$ third solvation shell network the stretch frequencies shift back toward the experimental liquid water envelope and near the band center, reflecting the emergence of structural heterogeneity and a broader distribution of hydrogen bond strengths that

more closely resemble the liquid environment. Nevertheless, the current study provides a quantitative account of the vibrational shifts in low coordination environments, such as the air-liquid interface.

These results highlight the different physical sensitivities of water's bending and stretching vibrations to the underlying hydrogen bond structure and demonstrate how bulk-like vibrational signatures emerge from finite networks with the appropriate coordination. Future studies will extend this work to build an *effective* solvated monomer 1-body potential that incorporates the appropriate physics and is transferable across different coordination and local order environments.

Chapter 4

A CAUTIONARY TALE: FAILURE OF THE VALENCE CASSCF TO DESCRIBE THE HALLMARK OF HYDROGEN BONDING

Reproduced with permission from [Long H. Nguyen, Thom H. Dunning, Jr., Sotiris S. Xantheas. A Cautionary Tale: Failure of the Valence CASSCF to Describe the Hallmark of Hydrogen Bonding. Letter to the Editor, *J. Chem. Theory Comput.* in press (2026).[49]

Valence CASSCF (vCAS) calculations for the hydrogen-bonded $(H_2O)_2$, $(HF)_2$, and HF–H_2O dimers fail to predict the lengthening of the hydrogen donor bond and the corresponding red shift in the donor-bond vibrational stretching frequency. Analysis of the active orbitals in the vCAS calculation reveals that one of the weakly occupied orbitals in each dimer is used to correlate the electrons in the acceptor lone pairs rather than those in the donor bonding orbitals. This has significant implications for subsequent treatments of the dynamical correlation that are based on the vCAS wavefunction.

4.1 Introduction

Hydrogen bonding is a foundational intermolecular interaction, and small hydrogen-bonded dimers have long served as benchmarks for understanding bonding, charge redistribution, and electron-correlation effects in molecular clusters.[167, 113, 168, 29, 169, 170, 171, 172, 173] Upon hydrogen-bond formation, the donor X–H covalent bond typically elongates and its stretching frequency shifts to lower energy relative to the isolated monomer.[151, 174, 6, 175, 176] This coupled structural–spectroscopic response is one of the defining signatures of hydrogen bonding. It was recognized in the original work of Badger and Bauer,[151] extended through empirical correlations such as those of Iogansen,[177, 178] and more recently placed on a quantitative footing by Boyer *et al.*, who showed that O–H bond elongation and stretching-

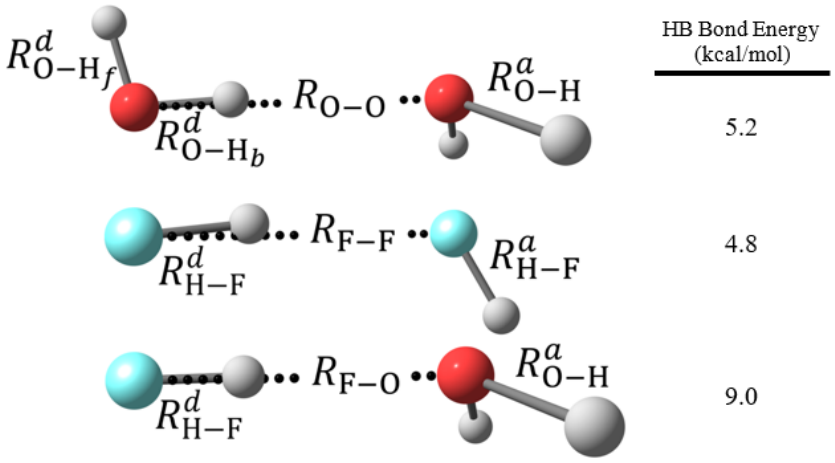

Figure 4.1: Optimum structures of the $(H_2O)_2$, $(HF)_2$, and HF–H_2O dimers together with the definition of selected geometrical parameters and the HB energies from CCSD(T)/aVTZ calculations.

frequency red shifts maintain an approximately linear relationship across aqueous clusters and liquid water.[6] Santis and Xantheas further developed analytic relationships connecting hydrogen-bond strength, covalent-bond elongation, and donor-stretch red shifts.[175, 176, 179]

Multiconfigurational self-consistent field (MCSCF) methods are often employed for systems where static (nondynamical) correlation is important, for instance when bonds are stretched or electronic configurations are nearly degenerate.[45, 180, 46, 181, 182, 183] A common choice is the valence CASSCF (vCAS) method, introduced by Roos and coworkers,[46] in which the active space includes all valence orbitals in the Hartree–Fock description of the constituent atoms. The vCAS approach has been applied successfully to countless molecular systems and is generally expected to yield reasonable potential-energy surfaces and equilibrium geometries, at least qualitatively, even for challenging cases.

In this chapter, we examine three dimers spanning a range of typical hydrogen-bond strengths, from moderate, $(H_2O)_2$ and $(HF)_2$, to relatively strong, HF–H_2O (Figure 4.1), making them ideal prototypes for testing the ability of theoretical methods to capture the

subtle electron-correlation and charge-transfer effects in hydrogen bonding. As shown below, the vCAS treatment breaks down even for these simple hydrogen-bonded dimers. Specifically, vCAS produces qualitatively incorrect geometries for the donor X–H bonds: the predicted donor-bond lengths are shorter than the donor free-bond lengths in the dimers and, in some cases, shorter than even the corresponding monomer bonds. This is accompanied by anomalous blue shifts of the donor X–H stretching frequencies, contrary to the observed red shifts. We analyze the vCAS wavefunction and identify the problem as a misallocation of the weakly occupied active orbitals. One of these orbitals in each dimer is used to correlate the electrons in the lone-pair orbitals of the acceptor, because doing so leads to a lower energy, thereby leaving the donor bond essentially uncorrelated and described as a doubly occupied bond orbital much like in the Hartree–Fock wavefunction.

4.2 Computational Details

For the vCAS calculations, Dunning's aug-cc-pVTZ (aVTZ) basis set was used throughout.[80, 82] The active spaces were specified by the valence orbitals of the atoms in each molecule, that is, the $2s$ and three $2p$ orbitals of O and F and the $1s$ orbitals of H. In all calculations, the $1s$ electrons of the heavy atoms were treated as inactive. Given the previously known geometries of the three dimers, C_s symmetry constraints were applied during the geometry optimizations, and harmonic frequencies were computed at the optimized geometries. The vCAS calculations were carried out with GAMESS,[184] and comparison CCSD(T) results were obtained with MOLPRO.[185, 186]

4.3 Results and Discussion

The equilibrium structures of the three dimers are shown in Figure 4.1, and selected geometrical parameters are collected in Table 4.1. For the isolated monomers, vCAS incorporates static correlation, especially left–right correlation of the electrons in the X–H bonds, and produces geometries close to the CCSD(T) values. For example, at the aVTZ level vCAS lengthens the H–F bond of HF from 0.8991 Å at the RHF level to 0.9176 Å, only 0.0034 Å shorter than

Table 4.1: Equilibrium parameters (bond lengths in Å, angles in degrees, total energies E_e in Hartree, and interaction energies ΔE_e in kcal mol^{-1}) for HF and H_2O and their hydrogen-bonded dimers at the RHF, vCAS, and CCSD(T) levels with the aVTZ basis set. Δ denotes the difference of the dimer intramolecular coordinates from the corresponding monomer values. Superscripts d and a denote hydrogen-bond donor and acceptor fragments, respectively; subscripts b and f denote hydrogen-bonded and free X–H bonds in the donor fragment.

System	Parameter	RHF	vCAS (Δ)	CCSD(T) (Δ)
HF	$R_{\mathrm{H-F}}$	0.8991	0.9176	0.9210
	E_e	−100.061464	−100.085328	−100.349577
H_2O	$R_{\mathrm{O-H}}$	0.9410	0.9641	0.9616
	θ_{HOH}	106.31	102.94	104.18
	E_e	−76.061203	−76.114025	−76.342326
$(HF)_2$	$R_{\mathrm{F-F}}$	2.8298	2.8526	2.7420
	$R^d_{\mathrm{H-F}}$	0.9030 (+0.0039)	0.9030 (−0.0146)	0.9266 (+0.0056)
	$R^a_{\mathrm{H-F}}$	0.9014 (+0.0023)	0.9217 (+0.0041)	0.9239 (+0.0029)
	θ_{HFF}	6.77	7.10	6.66
	E_e	−200.128876	−200.183205	−200.706849
	ΔE_e	−3.73	−7.87	−4.83
$(H_2O)_2$	$R_{\mathrm{O-O}}$	3.0392	3.0848	2.9133
	$R^d_{\mathrm{O-H_b}}$	0.9450 (+0.0040)	0.9448 (−0.0193)	0.9677 (+0.0061)
	$R^d_{\mathrm{O-H_f}}$	0.9403 (−0.0007)	0.9403 (−0.0237)	0.9608 (−0.0008)
	θ^d_{HOH}	106.33 (+0.0002)	106.31 (+3.38)	104.53 (+0.35)
	$R^a_{\mathrm{O-H}}$	0.9416 (+0.0006)	0.9630 (−0.0011)	0.9624 (+0.0008)
	θ^a_{HOH}	106.58 (+0.27)	104.91 (+1.97)	104.58 (+0.40)
	θ_{HOO}	3.11	3.12	5.66
	E_e	−152.128370	−152.258584	−152.692966
	ΔE_e	−3.74	−19.16	−5.22
HF–H_2O	$R_{\mathrm{F-O}}$	2.7145	2.7599	2.6477
	$R^d_{\mathrm{H-F}}$	0.9103 (+0.0112)	0.9092 (−0.0084)	0.9372 (+0.0162)
	$R^a_{\mathrm{O-H}}$	0.9416 (+0.0006)	0.9659 (+0.0018)	0.9625 (+0.0009)
	θ^a_{HOH}	107.20 (+0.89)	103.75 (+0.82)	105.14 (+0.96)
	θ_{HFO}	1.15	1.15	1.65
	E_e	−176.134231	−176.226537	−176.706196
	ΔE_e	−7.26	−17.06	−8.97

the CCSD(T) result. Similar agreement is obtained for H_2O.

4.3.1 Water dimer $(H_2O)_2$

Considered as the archetype of aqueous hydrogen bonding, the water dimer has been studied extensively by both experiment and high-level theory for decades.[167, 168, 123, 187, 188] Its vibrationally averaged O–O distance is 2.976 Å,[123] and its binding energy is about 5.0 ± 0.1 kcal mol^{-1}.[174, 189] Forming the hydrogen bond elongates the donor O–H bond by about 0.01 Å and correspondingly red shifts its stretching frequency by roughly 100 cm^{-1}.[167, 170, 174]

The RHF calculation yields a donor O–H bond lengthening from 0.9410 Å in the monomer to 0.9450 Å in the dimer, but, because it lacks electron correlation, RHF underestimates the interaction energy (-3.74 kcal mol^{-1}) and overestimates the O–O distance (3.0392 Å). In contrast, vCAS gives a qualitatively incorrect geometry: the optimized donor O–H bond in the dimer, 0.9448 Å, is shorter than the donor O–H bond in the vCAS monomer by 0.0193 Å. The RHF and vCAS donor-bond lengths in the dimer are nearly identical, strongly suggesting that the donor O–H bond in the vCAS calculation is effectively described at the RHF level. Moreover, vCAS predicts an O–O separation of 3.0848 Å, significantly longer than the CCSD(T) result of 2.9133 Å, indicating an underestimation of the attractive interaction between the two water molecules. Paradoxically, despite the overly long O–O distance, the vCAS interaction energy of -19.16 kcal mol^{-1} overestimates the CCSD(T) value of -5.22 kcal mol^{-1} by nearly a factor of four.

The harmonic frequencies reinforce the same picture. As shown in Table 4.2, CCSD(T) predicts the donor hydrogen-bonded O–H stretch at 3756 cm^{-1}, a red shift of -109 cm^{-1} relative to the average of the symmetric and antisymmetric O–H stretches of the monomer (3866 cm^{-1}). In contrast, vCAS places the donor hydrogen-bonded O–H stretch at 4073 cm^{-1}, corresponding to a blue shift of $+266$ cm^{-1} relative to the average monomer O–H stretches (3808 cm^{-1}). It further predicts a large blue shift ($+394$ cm^{-1}) for the donor free O–H stretch at 4201 cm^{-1}. Notably, the shorter donor O–H bond in the dimer compared with

the monomer at the vCAS/aVDZ level had already been reported by two of the authors in earlier work,[113] although the unphysical nature of that result was not discussed there.

4.3.2 HF dimer $(HF)_2$

In the hydrogen fluoride dimer, whose dissociation energy is approximately 4.58 kcal mol^{-1},[193] the donor H–F bond lengthens by roughly 0.006 Å and the donor stretching frequency is red shifted by about 100 cm^{-1} relative to the monomer.[194, 195, 196] This dimer shows the same qualitative vCAS failure as the water dimer. At the RHF level, the donor H–F bond is slightly longer than both the acceptor bond and the RHF monomer bond. At the vCAS level, however, the donor H–F bond is shorter than both the free H–F bond in the dimer and the H–F bond in the vCAS monomer. Again, the vCAS donor bond is essentially identical to the RHF donor bond, indicating that the donor H–F bond is described as if it were still at the Hartree–Fock level. The F–F distance is too long and the interaction energy is again overestimated, -7.87 kcal mol^{-1} versus -4.83 kcal mol^{-1} at CCSD(T).

The same inversion appears in the vibrational frequencies. CCSD(T) gives the donor H–F stretch at 4011 cm^{-1}, red shifted by -114 cm^{-1} relative to the monomer value of 4125 cm^{-1}. In contrast, vCAS predicts a donor H–F stretch at 4376 cm^{-1}, corresponding to a blue shift of $+260$ cm^{-1} relative to the vCAS monomer frequency of 4116 cm^{-1}.

4.3.3 HF–H_2O heterodimer

The heterodimer HF–H_2O forms a substantially stronger hydrogen bond, with $D_0 \approx -8.2$ kcal mol^{-1} and $D_e \approx -10.3$ kcal mol^{-1},[197] resulting in a significant lengthening of the donor H–F bond and a correspondingly larger red shift of its stretching frequency.[195, 198, 199] RHF captures the qualitative geometry of the complex, but it underestimates both the interaction energy and the donor-bond elongation because the heterodimer involves a large degree of polarization and charge transfer. The vCAS(16,11) result is again extreme: instead of elongating the donor H–F bond, it predicts a bond length shorter than that in the vCAS

Table 4.2: Harmonic vibrational frequencies (cm^{-1}) for the HF and H_2O monomers and intramolecular frequencies of their dimers computed at the vCAS and CCSD(T) levels of theory with the aVTZ basis set. The shifts are relative to the corresponding monomer frequencies. Monomer reference frequencies are from Refs. 190, 191; dimer frequencies for water clusters are from Refs. 174, 192.

System	Mode	vCAS	$\Delta_{\mathrm{dimer-mon.}}$	CCSD(T)	$\Delta_{\mathrm{dimer-mon.}}$
HF	H–F stretching	4116	—	4125	—
H_2O	H–O–H bend	1689	—	1646	—
	O–H symmetric stretch	3750	—	3811	—
	O–H antisymmetric stretch	3865	—	3920	—
$(HF)_2$	H–F donor stretching	4376	+260	4011	−114
	H–F acceptor stretching	4106	−10	4040	−85
$(H_2O)_2$	H–O–H acceptor bend	1676	−13	1643	−3
	H–O–H donor bend	1763	+74	1668	+22
	donor O–H (H-bonded) stretch	4073	+266	3756	−110
	acceptor O–H symmetric stretch	3814	+7	3830	−36
	donor O–H (free) stretch	4201	+394	3911	+46
	acceptor O–H antisymmetric stretch	3927	+120	3930	+65
HF–H_2O	H–O–H acceptor bend	1691	+2	1648	+2
	H–F donor stretching	4216	+100	3762	−363
	O–H acceptor symmetric stretch	3766	−42	3809	−57
	O–H acceptor antisymmetric stretch	3875	+68	3910	+45

monomer. At the same time, it overestimates the interaction energy by roughly a factor of two and predicts an F–O distance substantially longer than the CCSD(T) value.

The spectroscopic discrepancy is correspondingly the largest for this dimer. CCSD(T) predicts the donor H–F stretch at 3762 cm^{-1}, an expected red shift of -363 cm^{-1} from the monomer value of 4125 cm^{-1}. By contrast, vCAS yields 4216 cm^{-1} for the donor H–F stretch, corresponding to a blue shift of $+100$ cm^{-1} relative to the vCAS monomer value of 4116 cm^{-1}.

Across all three dimers, the same qualitative pattern emerges. vCAS, designed to account for static correlation, predicts a contraction rather than an elongation of the donor X–H bond and a blue rather than a red shift in the associated vibrational frequency with respect to the monomer values. This failure is accompanied by excessive interaction energies, indicating that vCAS adds electron correlation in an unbalanced way that overstabilizes the dimers. By contrast, CCSD(T) restores the expected structural and spectroscopic trends and yields values much closer to experiment. The stark differences between vCAS and CCSD(T) demonstrate that the *distribution* of electron correlation can be as important as the mere presence of correlation itself.

4.3.4 Origin of the vCAS failure: analysis of the natural orbitals

To determine the underlying cause of the inverted structural and spectroscopic trends, we examined the active-space natural orbitals (NOs) and their occupation numbers for each dimer. In a vCAS calculation, the active space is defined as the union of the valence atomic orbitals of the constituent atoms. The implicit expectation is that the weakly occupied active orbitals will include the antibonding $\sigma^*(\mathrm{X}-\mathrm{H})$ orbitals needed to provide left–right correlation in the donor bond and thereby stabilize the elongated donor X–H bond participating in hydrogen bonding. However, the variational optimization of the orbitals can instead divert one or more of these weakly occupied orbitals to correlate lone-pair orbitals if that assignment yields a lower total energy. This misallocation of the active orbitals reflects the size-inconsistency present in the vCAS description. In a consistent treatment, the active space of the dimer

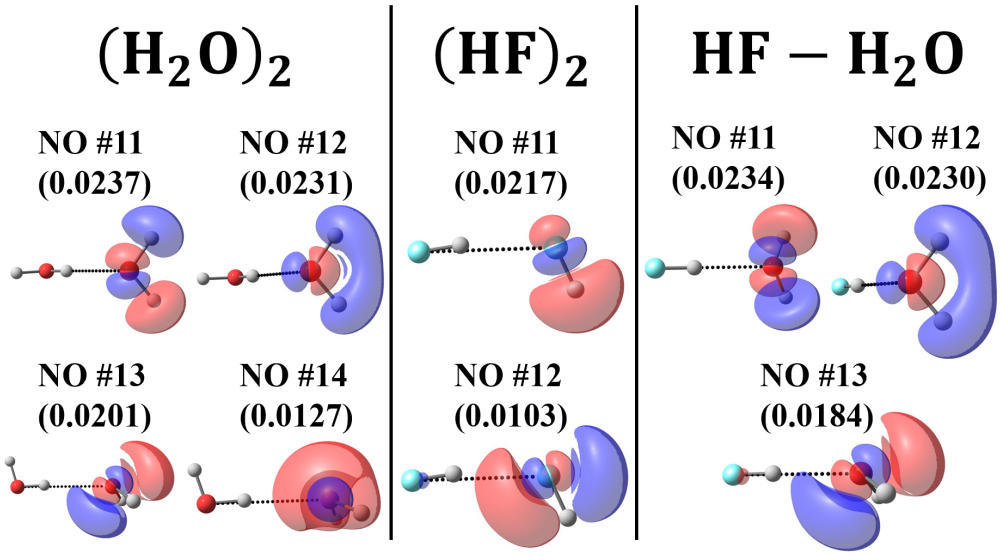

Figure 4.2: The 4/2/3 weakly occupied orbitals in the $(H_2O)_2$, $(HF)_2$, and HF–H_2O dimers from vCAS(16,12), vCAS(16,10), and vCAS(16,11)/aVTZ calculations, respectively. Occupation numbers are given in parentheses.

should reduce to the corresponding active spaces of the separated monomers. Here, however, the orbitals involved in the vCAS correlation change character upon dimer formation from bonding and antibonding orbitals in the monomers to a set of orbitals that are involved in correlating some of the lone pairs in the dimer. Consequently, the monomer and dimer vCAS wave functions are not described on the same footing, leading to a size-inconsistent reference. This imbalance can affect not only the optimized structures and vibrational features of the dimers, but also the resulting interaction energies. Similar size-consistency issues in supermolecular CASSCF/multiconfigurational interaction energies have been discussed previously.[200, 201, 202]

The weakly occupied vCAS NOs for the three dimers, namely 4, 2, and 3 NOs for the vCAS wave functions of $(H_2O)_2$, $(HF)_2$, and HF–H_2O with the aVTZ basis set, are shown in Figure 4.2. The vCAS calculation clearly illustrates that the weakly occupied orbitals are not dominated by the donor-bond antibonding orbital, $\sigma^*(O - H_d)$, which describes left–right

correlation in the hydrogen-bonded $O-H_d$ bond. Instead, they are primarily assigned to describe other valence correlation effects. For example, for $(H_2O)_2$, NO #11 is predominantly a $\sigma^*(O-H_a)$ orbital, the antisymmetric O–H antibonding orbital of the two O–H bonds in the acceptor water, and NO #12 is the symmetric $\sigma^*(O-H_a)$ orbital of the two O–H bonds in the acceptor. This explains why the bond lengths of the two acceptor O–H bonds, 0.9630 , in vCAS(16,12) are slightly longer than the isolated-water vCAS(8,6) value of 0.9641 and are largely elongated compared to the RHF result for the dimer, 0.9416 . Notably, the vCAS result for the acceptor O–H bond length is in good agreement with the CCSD(T) result, 0.9624 . Furthermore, NO #13 clearly shows that one of the weakly occupied orbitals is being used to correlate one of the oxygen lone pairs on the acceptor water, corresponding to in-and-out correlation, and NO #14 corresponds to a polarized O(3s) orbital. Similar arguments apply to $(HF)_2$ and HF–H_2O.

The results reported here correspond to the lowest-energy vCAS solution, which is often associated with a non-desirable choice of active space. It is sometimes possible to bias the active space toward a higher-energy but chemically more appropriate solution through the use of valence virtual orbitals (VVOs), freezing of core orbitals during the vCAS optimization, employing vCAS(SCGVB) as a core-safe reference, or following multi-step constrained workflows.[203, 204, 205] In the present systems, however, such approaches did not yield robust solutions. A more systematic investigation of these alternatives will be the subject of future work.

4.4 Conclusions

Hydrogen-bonded dimers provide a stringent test for multiconfigurational methods. We find that a valence CASSCF treatment fails to even qualitatively describe the $(H_2O)_2$, $(HF)_2$, and HF–H_2O dimers, predicting donor X–H bond contractions instead of elongations and vibrational blue shifts instead of red shifts relative to the isolated monomers. The origin of these anomalies is a misallocation of the active orbitals: one of the weakly occupied orbitals is used to correlate acceptor lone pairs rather than the donor bond pair, leaving the donor

bond effectively uncorrelated. These results underscore that vCAS is not a black-box tool for modeling intermolecular interactions, especially hydrogen bonding, and they cast doubt on the use of the resulting vCAS wavefunction as the reference for subsequent dynamical-correlation treatments such as CASPT2. Future work will examine this issue in a larger set of dimers and investigate alternative strategies for constructing more robust active spaces.

Chapter 5

CONCLUSIONS AND FUTURE DIRECTIONS

This thesis examined hydrogen bonding in water and related hydrogen bonded clusters by linking microscopic structure, vibrational spectroscopy, and electronic structure theory. The vibrational part of the work traced how the fundamentals of a water molecule evolve as it is embedded in progressively extended hydrogen bonded environments mimicking the networks in liquid water. The electronic structure part tested whether different theoretical methods recover the hallmark signatures of hydrogen bonding, including donor bond elongation and vibrational red shifts.

The first part investigated how the vibrational fundamentals of a water molecule evolve as it is embedded in increasingly extended hydrogen bonded environments. Using a hierarchy of water clusters, $(H_2O)_n$ with n = 1, 2, 5, 8, 17, and 53, the calculations traced the transition from the isolated monomer to bulk-like environments. The results showed that the bending and stretching vibrations converge in fundamentally different ways. The H–O–H bending frequency responds primarily to the immediate local hydrogen bond geometry and reaches values characteristic of liquid water already within the first solvation shell.In contrast, the O–H stretching fundamentals remain strongly sensitive to the broader hydrogen bond environment. Tetrahedrally constrained second shell structures produce very large red shifts associated with cooperative hydrogen bonding and ice-like local order, whereas larger and more relaxed networks shift the stretches back into the liquid water absorption envelope. These findings provide a microscopic explanation for the contrasting widths and environmental sensitivities of the bending and stretching bands in the infrared spectrum of water.

A further contribution of the vibrational analysis is methodological. By combining harmonic calculations, VPT2, and reduced dimensional DVR treatments, this work established

a practical route for obtaining best estimates of vibrational fundamentals while explicitly tracking uncertainties associated with electron correlation, basis-set size, and DVR grid convergence. The comparison between MP2 and CCSD(T), together with the transfer of monomer-based and cluster size corrections to larger embedded systems, made it possible to assess how robust the predicted trends are even when full high level anharmonic treatments are computationally prohibitive. This framework is important not only for interpreting the present cluster hierarchy, but also for future extensions to larger water networks and other hydrogen bonded systems.

These results have an important implication for the construction and interpretation of water potentials. An accurate gas phase monomer potential alone is not sufficient to guarantee an accurate description of monomers embedded in homogeneous or heterogeneous hydrogen bonded environments. The present results indicate that the effective intramolecular response of a water molecule, especially along the O–H stretching coordinates, depends strongly on its surroundings. In this sense, the one-body term relevant for spectroscopy in condensed-phase-like environments cannot be viewed as a fixed gas phase potential. Rather, the results support the need for an *effective* environment-dependent 1-body term capable of adapting to differences in coordination, local order, and hydrogen bond strength.

The second part demonstrated that valence complete active space self-consistent field (vCAS) wavefunctions can fail qualitatively even for simple hydrogen bonded dimers. For $(H_2O)_2$, $(HF)_2$, and HF–H_2O, the vCAS treatment predicts donor X–H bond contractions and corresponding vibrational blue shifts relative to the isolated monomers, in direct contradiction to the established hallmark of hydrogen bonding obtained from both experiments and highly correlated electronic structure calculations. The analysis of the active orbitals showed that this failure originates from an unbalanced allocation of the weakly occupied orbitals: instead of correlating the donor bond, one of these orbitals is preferentially used to correlate the acceptor lone pairs, leaving the donor bond effectively uncorrelated. These results show that vCAS should not be used as a black box tool for weak intermolecular interactions and that chemically meaningful observables, such as donor bond elongation and associated vibrational

red shifts, provide an essential test of whether the active space is physically appropriate. In contrast, HF and the single reference based mathods MP2 and CCSD(T) correctly recover the donor bond elongation and associated red shifts and therefore provide a more reliable framework for describing the properties of the hydrogen bond in these dimers.

This thesis provides a consistent microscopic picture of hydrogen bonding in which structure, spectroscopy, and electronic structure are tightly connected. Theoretical methods must also reproduce the correct structural response of donor bonds, the associated spectroscopic signatures, and the changing balance of many-body interactions across environments. For the systems studied here, MP2 and especially CCSD(T) provide reliable single-reference benchmarks, whereas vCAS can fail when the active space does not represent the physically relevant donor bond correlation. The vibrational analyses demonstrate that the bending mode is a relatively local probe of hydrogen bonding, while the stretching modes encode the broader and more heterogeneous organization of the surrounding network.

Several directions follow naturally from this work. One is the construction of *effective* solvated monomer one-body potentials and dipole moment surfaces for water,that incorporate the physics identified here and remain transferable across environments with different coordination motifs and local order. Another is the extension of the present vibrational analysis to larger clusters, additional isotopic substitutions, and more complete treatments of intermolecular mode coupling and the calculation of accurate intensities. A third is the continued examination of multi-configurational references for hydrogen bonded systems, including more robust active space choices and higher level post-CASSCF treatments such as Multi-Reference CI (MRCI), second order perturbation theory treatment on top of a CASSCF reference (CASPT2) etc. More broadly, the combination of wavefunction analysis and vibrational spectroscopy in this thesis provides a framework for studying how local electronic structure evolves into collective condensed phase behavior. In that sense, the present work contributes to a more unified microscopic picture of hydrogen bonding in water: one in which structure and spectroscopy are inseparable aspects of the same underlying many-body phenomena.

BIBLIOGRAPHY

[1] Elangannan Arunan, Gautam R. Desiraju, Roger A. Klein, Joanna Sadlej, Steve Scheiner, Ibon Alkorta, David C. Clary, Robert H. Crabtree, Joseph J. Dannenberg, Pavel Hobza, Henrik G. Kjaergaard, Anthony C. Legon, Benedetta Mennucci, and David J. Nesbitt. Definition of the hydrogen bond (IUPAC recommendations 2011). *Pure and Applied Chemistry*, 83(8):1637–1641, 2011.

[2] Bradley M. Auer and James L. Skinner. IR and Raman spectra of liquid water: Theory and interpretation. *The Journal of Chemical Physics*, 128:224511, 2008.

[3] NIST Chemistry WebBook. Water (H_2O)—vibrational and/or electronic energy levels and related spectroscopic data. National Institute of Standards and Technology. Accessed 2026-04-09.

[4] Alexei A. Kananenka and James L. Skinner. Fermi resonance in OH-stretch vibrational spectroscopy of liquid water and the water hexamer. *The Journal of Chemical Physics*, 148:244107, 2018.

[5] Richard M. Badger. A relation between internuclear distances and bond force constants. *The Journal of Chemical Physics*, 2(3):128–131, 1934.

[6] M. A. Boyer, O. Marsalek, J. P. Heindel, T. E. Markland, A. B. McCoy, and S. S. Xantheas. Beyond badger's rule: The origins and generality of the structure-spectra relationship of aqueous hydrogen bonds. *The Journal of Physical Chemistry Letters*, 10 (5):918–924, 2019.

[7] John E. Bertie and Zhida Lan. Infrared intensities of liquids XX: The intensity of the OH stretching band of liquid water revisited, and the best current values of the optical

constants of H_2O(l) at 25°c between 15,000 and 1 cm^{-1}. *Applied Spectroscopy*, 50(8): 1047–1057, 1996.

[8] John E. Bertie and Edward Whalley. Infrared spectra of ices Ih and Ic in the range 4000 to 350 cm^{-1}. *The Journal of Chemical Physics*, 40(6):1637–1645, 1964.

[9] Edward Whalley. A detailed assignment of the O–H stretching bands of ice I. *Canadian Journal of Chemistry*, 55(19):3429–3441, 1977.

[10] S. K. Reddy, S. C. Straight, P. Bajaj, C. H. Pham, M. Riera, D. R. Moberg, M. A. Morales, C. Knight, A. W. Götz, and Francesco Paesani. On the accuracy of the MB-pol many-body potential for water: Interaction energies, vibrational frequencies, and classical thermodynamic and dynamical properties from clusters to liquid water and ice. *The Journal of Chemical Physics*, 145:194504, 2016.

[11] J. E. Lennard-Jones. On the forces between atoms and ions. *Proceedings of the Royal Society of London. Series A*, 109(752):584–597, 1925.

[12] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, and J. Hermans. Interaction models for water in relation to protein hydration. In B. Pullman, editor, *Intermolecular Forces*, pages 331–342. Reidel, Dordrecht, 1981.

[13] H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma. The missing term in effective pair potentials. *The Journal of Physical Chemistry*, 91(24):6269–6271, 1987.

[14] William L. Jorgensen, Jayaraman Chandrasekhar, Jeffry D. Madura, Roger W. Impey, and Michael L. Klein. Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics*, 79(2):926–935, 1983.

[15] J. L. F. Abascal and C. Vega. A general purpose model for the condensed phases of water: TIP4P/2005. *The Journal of Chemical Physics*, 123(23):234505, 12 2005.

[16] J. L. F. Abascal, E. Sanz, R. García Fernández, and C. Vega. A potential model for the study of ices and amorphous water: Tip4p/ice. *J. Chem. Phys.*, 122:234511, 2005.

[17] Michael W. Mahoney and William L. Jorgensen. A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions. *J. Chem. Phys.*, 112(20):8910–8922, 2000.

[18] Saeed Izadi, Ramu Anandakrishnan, and Alexey V. Onufriev. Building water models: A different approach. *J. Phys. Chem. Lett.*, 5(21):3863–3871, 2014.

[19] Kristina M. Herman and Sotiris S. Xantheas. An extensive assessment of the performance of pairwise and many-body interaction potentials in reproducing ab initio benchmark binding energies for water clusters n = 225. *Phys. Chem. Chem. Phys.*, 25:7120–7143, 2023.

[20] Harry Partridge and David W. Schwenke. The determination of an accurate isotope dependent potential energy surface for water from extensive *ab initio* calculations and experimental data. *The Journal of Chemical Physics*, 106(11):4618–4639, 03 1997.

[21] D. W. Schwenke and H. Partridge. Concerning the ab initio determination of the dipole moment function of water. *J. Chem. Phys.*, 113:6592–6597, 2000.

[22] C. J. Burnham and S. S. Xantheas. Development of transferable interaction models for water. IV. a flexible, all-atom polarizable potential (TTM2-F) based on geometry dependent charges derived from an ab initio monomer dipole moment surface. *The Journal of Chemical Physics*, 116(12):5115–5124, 03 2002.

[23] George S. Fanourgakis and Sotiris S. Xantheas. The flexible, polarizable, thole-type interaction potential for water (TTM2-F) revisited. *The Journal of Physical Chemistry A*, 110(11):4100–4106, 2006.

[24] George S. Fanourgakis and Sotiris S. Xantheas. Development of transferable interaction potentials for water. V. extension of the flexible, polarizable, thole-type model potential

(TTM3-F) to describe the vibrational spectra of water clusters and liquid water. *The Journal of Chemical Physics*, 128:074506, 2008.

[25] Eleftherios Lambros and Francesco Paesani. How good are polarizable and flexible models for water: Insights from a many-body perspective. *The Journal of Chemical Physics*, 153(6):060901, 08 2020.

[26] Christian J. Burnham, Jichen Li, Sotiris S. Xantheas, and Maurice Leslie. The parametrization of a thole-type all-atom polarizable water model from first principles and its application to the study of water clusters (n=2–21) and the phonon spectrum of ice Ih. *The Journal of Chemical Physics*, 110(9):4566–4581, 03 1999.

[27] Christian J. Burnham and Sotiris S. Xantheas. Development of transferable interaction models for water. III. reparametrization of an all-atom polarizable rigid model (TTM2R) from first principles. *The Journal of Chemical Physics*, 116(4):1500–1510, 01 2002.

[28] Christian J. Burnham, David J. Anick, Paramjit K. Mankoo, and George F. Reiter. The vibrational proton potential in bulk liquid water and ice. *The Journal of Chemical Physics*, 128:154519, 2008.

[29] Christian J. Burnham and Sotiris S. Xantheas. Development of transferable interaction models for water. I. prominent features of the water dimer potential energy surface. *The Journal of Chemical Physics*, 116(4):1479–1492, 01 2002.

[30] G. S. Fanourgakis, G. K. Schenter, and S. S. Xantheas. A quantitative account of quantum effects in liquid water. *The Journal of Chemical Physics*, 125(14):141102, 10 2006. ISSN 0021-9606. doi: 10.1063/1.2358137.

[31] B. T. Thole. Molecular polarizabilities calculated with a modified dipole interaction. *Chemical Physics*, 59(3):341–350, 1981.

[32] Chao Qu and Joel M. Bowman. Permutationally invariant potential energy surfaces. *Annual Review of Physical Chemistry*, 69:151–175, 2018.

[33] Yimin Wang, Xinchuan Huang, Benjamin C. Shepler, Bastiaan J. Braams, and Joel M. Bowman. Flexible, *Ab Initio* potential, and dipole moment surfaces for water. I. tests and applications for clusters up to the 22-mer. *The Journal of Chemical Physics*, 134: 094509, 2011.

[34] R. Bukowski, K. Szalewicz, G. C. Groenenboom, and A. van der Avoird. Polarizable interaction potential for water from coupled cluster calculations. II. applications to dimer spectra and virial coefficients. *J. Chem. Phys.*, 128:094314, 2008.

[35] V. Babin, G. R. Medders, and F. Paesani. Development of a "first-principles" water potential with flexible monomers. I. methodology and bulk properties. *J. Chem. Theory Comput.*, 10:1599–1607, 2014.

[36] Qi Yu, Chen Qu, Paul L. Houston, Riccardo Conte, and Joel M. Bowman. q-aqua: A many-body CCSD(T) water potential, including four-body interactions, demonstrates the quantum nature of water from clusters to the liquid phase. *The Journal of Physical Chemistry Letters*, 13(22):5068–5074, 2022.

[37] Max Born and J. Robert Oppenheimer. Zur quantentheorie der molekeln. *Annalen der Physik*, 389(20):457–484, 1927.

[38] John C. Slater. The theory of complex spectra. *Physical Review*, 34(10):1293–1322, 1929.

[39] P. A. M. Dirac. Quantum mechanics of many-electron systems. *Proceedings of the Royal Society of London. Series A*, 123(792):714–733, 1929.

[40] D. R. Hartree. The wave mechanics of an atom with a non-coulomb central field. part ii. some results and discussion. *Mathematical Proceedings of the Cambridge Philosophical Society*, 24(1):111–132, 1928.

[41] V. Fock. Näherungsmethode zur lösung des quantenmechanischen mehrkörperproblems. *Zeitschrift für Physik*, 61(1–2):126–148, 1930.

[42] C. C. J. Roothaan. New developments in molecular orbital theory. *Reviews of Modern Physics*, 23(2):69–89, 1951.

[43] George G. Hall. The molecular orbital theory of chemical valency. viii. a method of calculating ionization potentials. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 205(1083):541–552, 1951.

[44] C. David Sherrill and Henry F. Schaefer. The configuration interaction method: Advances in highly correlated approaches. *Advances in Quantum Chemistry*, 34: 143–269, 1999.

[45] L. M. Cheung, K. R. Sundberg, and K. Ruedenberg. Dimerization of carbene to ethylene. *Journal of the American Chemical Society*, 100(25):8024–8025, 1978.

[46] Björn O. Roos, Peter R. Taylor, and Per E. M. Siegbahn. A complete active space SCF method (CASSCF) using a density matrix formulated super-CI approach. *Chemical Physics*, 48(2):157–173, 1980.

[47] K. Ruedenberg, M. W. Schmidt, M. M. Gilbert, and S. T. Elbert. Are atoms intrinsic to molecular electronic wavefunctions? i. the fors model. *Chemical Physics*, 71(1): 41–49, 1982. doi: 10.1016/0301-0104(82)87004-3.

[48] Hans-Joachim Werner and Peter J. Knowles. A second order multiconfiguration SCF procedure with optimum convergence. *The Journal of Chemical Physics*, 82(11): 5053–5063, 1985.

[49] L. Nguyen, T. H. Dunning, Jr., and S. S. Xantheas. A cautionary tale: Failure of the valence CASSCF to describe the hallmark of hydrogen bonding. *Journal of Chemical Theory and Computation*, 2026. Letter to the Editor, in press.

[50] Isaiah Shavitt. The method of configuration interaction. In Henry F. Schaefer, editor, *Methods of Electronic Structure Theory*, pages 189–275. Plenum Press, New York, 1977.

[51] Hans-Joachim Werner and Peter J. Knowles. An efficient internally contracted multiconfiguration-reference configuration interaction method. *Journal of Chemical Physics*, 89(9):5803–5814, 1988. doi: 10.1063/1.455556.

[52] Kerstin Andersson, Per-Åke Malmqvist, Björn O. Roos, Andrzej J. Sadlej, and Krzysztof Woliński. Second-order perturbation theory with a CASSCF reference function. *The Journal of Physical Chemistry*, 94(14):5483–5488, 1990. doi: 10.1021/j100377a012.

[53] K. Andersson, P.-Å. Malmqvist, and B. O. Roos. Second-order perturbation theory with a complete active space self-consistent field reference function. *The Journal of Chemical Physics*, 96(2):1218–1226, 1992.

[54] Chr. Møller and M. S. Plesset. Note on an approximation treatment for many-electron systems. *Physical Review*, 46(7):618–622, 1934.

[55] Michael J. Frisch, Martin Head-Gordon, and John A. Pople. A direct MP2 gradient method. *Chemical Physics Letters*, 166(3):275–280, 1990.

[56] Martin Head-Gordon, John A. Pople, and Michael J. Frisch. MP2 energy evaluation by direct methods. *Chemical Physics Letters*, 153(6):503–506, 1988.

[57] Michael J. Frisch, Martin Head-Gordon, and John A. Pople. Semi-direct algorithms for the MP2 energy and gradient. *Chemical Physics Letters*, 166(3):281–289, 1990.

[58] Martin Head-Gordon and Teresa Head-Gordon. Analytic MP2 frequencies without fifth-order storage. theory and application to bifurcated hydrogen bonds in the water hexamer. *Chemical Physics Letters*, 220(1–2):122–128, 1994.

[59] Léon Brillouin. *Les champs "self-consistents" de Hartree et de Fock*. Number 71 in Actualités Scientifiques et Industrielles. Hermann, Paris, 1934.

[60] Martin Head-Gordon, John A. Pople, and Michael J. Frisch. Mp2 energy evaluation by direct methods. *Chemical Physics Letters*, 153(6):503–506, 1988.

[61] Michael J. Frisch, Martin Head-Gordon, and John A. Pople. A direct MP2 gradient method. *Chemical Physics Letters*, 166(3):275–280, 1990.

[62] Sotiris S. Xantheas and Edoardo Aprà. The binding energies of the d2d and s4 water octamer isomers: High-level electronic structure and empirical potential results. *The Journal of Chemical Physics*, 120(2):823–828, 01 2004.

[63] Soohaeng Yoo, Edoardo Aprà, Xiao Cheng Zeng, and Sotiris S. Xantheas. High-level ab initio electronic structure calculations of water clusters $(H_2O)_{16}$ and $(H_2O)_{17}$: A new global minimum for $(H_2O)_{16}$. *The Journal of Physical Chemistry Letters*, 1(20): 3122–3127, 2010.

[64] Evangelos Miliordos and Sotiris S. Xantheas. An accurate and efficient computational protocol for obtaining the complete basis set limits of the binding energies of water clusters at the MP2 and CCSD(T) levels of theory: Application to $(H_2O)_m$, $m = 2$–6, 8, 11, 16, and 17. *The Journal of Chemical Physics*, 142(23):234303, 2015.

[65] Sotiris S. Xantheas. Ab initio studies of cyclic water clusters $(H_2O)_n$, $n = 1$–6. II. analysis of many-body interactions. *The Journal of Chemical Physics*, 100(10): 7523–7534, 1994.

[66] Sotiris S. Xantheas. Cooperativity and hydrogen bonding network in water clusters. *Chemical Physics*, 258(2–3):225–231, 2000.

[67] Berhane Temelso, Carla R. Renner, and George C. Shields. Importance and reliability of small basis set CCSD(T) corrections to MP2 binding and relative energies of water clusters. *Journal of Chemical Theory and Computation*, 11(4):1439–1448, 2015.

[68] Joseph P. Heindel and Sotiris S. Xantheas. The many-body expansion for aqueous systems revisited: I. water–water interactions. *Journal of Chemical Theory and Computation*, 16(11):6843–6855, 2020.

[69] Kristina M. Herman, Anthony J. Stone, and Sotiris S. Xantheas. Accurate calculation of many-body energies in water clusters using a classical geometry-dependent induction model. *Journal of Chemical Theory and Computation*, 19(19):6805–6815, 2023.

[70] Jeppe Olsen, Ove Christiansen, Henrik Koch, and Poul Jørgensen. Surprising cases of divergent behavior in Møller–Plesset perturbation theory. *The Journal of Chemical Physics*, 105(12):5082–5090, 1996.

[71] Matthew L. Leininger, Wesley D. Allen, Henry F. Schaefer, and C. David Sherrill. Is Møller–Plesset perturbation theory a convergent ab initio method? *The Journal of Chemical Physics*, 112(21):9213–9222, 2000.

[72] J. Coleman Howard and Gregory S. Tschumper. Benchmark structures and harmonic vibrational frequencies near the CCSD(T) complete basis set limit for small water clusters: $(H_2O)_{n=2,3,4,5,6}$. *Journal of Chemical Theory and Computation*, 11(5):2126–2136, 2015.

[73] Jiří Čížek. On the correlation problem in atomic and molecular systems. calculation of wavefunction components in ursell-type expansion using quantum-field theoretical methods. *The Journal of Chemical Physics*, 45(11):4256–4266, 1966.

[74] Rodney J. Bartlett and Monika Musiał. Coupled-cluster theory in quantum chemistry. *Reviews of Modern Physics*, 79(1):291–352, 2007.

[75] George D. Purvis and Rodney J. Bartlett. A full coupled-cluster singles and doubles model: The inclusion of disconnected triples. *The Journal of Chemical Physics*, 76(4): 1910–1918, 1982.

[76] Krishnan Raghavachari, Gary W. Trucks, John A. Pople, and Martin Head-Gordon. A fifth-order perturbation comparison of electron correlation theories. *Chemical Physics Letters*, 157(6):479–483, 1989.

[77] Evangelos Miliordos, Edoardo Aprà, and Sotiris S. Xantheas. Optimal geometries and harmonic vibrational frequencies of the global minima of water clusters $(H_2O)_n$, n=2-6, and several hexamer local minima at the CCSD(T) level of theory. *The Journal of Chemical Physics*, 139(11):114302, 09 2013.

[78] S. F. Boys. Electronic wave functions. i. a general method of calculation for the stationary states of any molecular system. *Proceedings of the Royal Society of London. Series A*, 200(1063):542–554, 1950.

[79] Warren J. Hehre, Richard Ditchfield, and John A. Pople. Self-consistent molecular orbital methods. XII. further extensions of gaussian-type basis sets for use in molecular orbital studies of organic molecules. *The Journal of Chemical Physics*, 56(5):2257–2261, 1972.

[80] Thom H. Dunning. Gaussian basis sets for use in correlated molecular calculations. I. the atoms boron through neon and hydrogen. *The Journal of Chemical Physics*, 90(2): 1007–1023, 1989.

[81] David E. Woon and Thom H. Dunning. Gaussian basis sets for use in correlated molecular calculations. III. the atoms aluminum through argon. *The Journal of Chemical Physics*, 98(2):1358–1371, 1993.

[82] Rick A. Kendall, Jr. Dunning, Thom H., and Robert J. Harrison. Electron affinities of the firstrow atoms revisited. systematic basis sets and wave functions. *The Journal of Chemical Physics*, 96(9):6796–6806, 05 1992.

[83] Jr. Wilson, E. Bright, J. C. Decius, and Paul C. Cross. *Molecular Vibrations: The Theory of Infrared and Raman Vibrational Spectra*. McGraw-Hill, New York, 1955.

[84] Vincenzo Barone. Anharmonic vibrational properties by a fully automated second-order perturbative approach. *The Journal of Chemical Physics*, 122:014108, 2005.

[85] Gregory A. Cisneros, K. T. Wikfeldt, Lars Ojamäe, Jun Lu, Yao Xu, Hoda Torabifard, Albert P. Bartók, Gábor Csányi, Valeria Molinero, and Francesco Paesani. Modeling molecular interactions in water: From pairwise to many-body potential energy functions. *Chemical Reviews*, 116(13):7501–7528, 2016.

[86] János Pipek and Paul G. Mezey. A fast intrinsic localization procedure applicable for *Ab Initio* and semiempirical linear combination of atomic orbital wave functions. *The Journal of Chemical Physics*, 90(9):4916–4926, 1989.

[87] Curt M. Breneman and Kenneth B. Wiberg. Determining atom-centered monopoles from molecular electrostatic potentials. the need for high sampling density in formamide conformational analysis. *Journal of Computational Chemistry*, 11(3):361–373, 1990.

[88] Bo Wang and Donald G. Truhlar. Partial atomic charges and screened charge models of the electrostatic potential. *Journal of Chemical Theory and Computation*, 8(6): 1989–1998, 2012.

[89] Richard F. W. Bader. *Atoms in Molecules: A Quantum Theory*. Oxford University Press, Oxford, 1990. ISBN 978-0198558651.

[90] Hanchao Liu, Yimin Wang, and Joel M. Bowman. Transferable ab initio dipole moment for water: Three applications to bulk water. *The Journal of Physical Chemistry B*, 120 (8):1735–1742, 2016.

[91] Francesco Paesani, Sotiris S. Xantheas, and Gregory A. Voth. Infrared spectroscopy and hydrogen-bond dynamics of liquid water from centroid molecular dynamics with an ab initio-based force field. *The Journal of Physical Chemistry B*, 113(39):13118–13130, 2009.

[92] David T. Colbert and William H. Miller. A novel discrete variable representation for quantum mechanical reactive scattering via the S-matrix Kohn method. *The Journal of Chemical Physics*, 96(3):1982–1991, 1992.

[93] Julián Echave and David C. Clary. Potential optimized discrete variable representation. *Chemical Physics Letters*, 190(3):225–230, 1992.

[94] I. M. Mills. Vibrational perturbation theory. *Journal of Molecular Spectroscopy*, 5(1–6): 334–340, 1961.

[95] Samuel L. Bore, Christopher E. B. Daducci, Gregory R. Medders, and Francesco Paesani. Realistic phase diagram of water from "first principles" data-driven quantum simulations. *Nat. Commun.*, 14:3353, 2023.

[96] P. L. Silvestrelli and M. Parrinello. Water molecule dipole in the gas and in the liquid phase. *Phys. Rev. Lett.*, 82:3308–3311, 1999.

[97] G. A. Cisneros, K. T. Wikfeldt, L. Ojamäe, and et al. Modeling molecular interactions in water: From pairwise to many-body potential energy functions. *Chem. Rev.*, 116: 7501–7528, 2016.

[98] I.E. Gordon, L.S. Rothman, R.J. Hargreaves, F.M. Gomez, T. Bertin, C. Hill, R.V. Kochanov, Y. Tan, P. Wciso, V. Yu. Makhnev, P.F. Bernath, M. Birk, V. Boudon, A. Campargue, A. Coustenis, B.J. Drouin, R.R. Gamache, J.T. Hodges, D. Jacquemart, E.J. Mlawer, A.V. Nikitin, V.I. Perevalov, M. Rotger, S. Robert, J. Tennyson, G.C. Toon, H. Tran, V.G. Tyuterev, E.M. Adkins, A. Barbe, D.M. Bailey, K. Bielska, L. Bizzocchi, T.A. Blake, C.A. Bowesman, P. Cacciani, P. ermák, A.G. Császár, L. Denis, S.C. Egbert, O. Egorov, A. Yu. Ermilov, A.J. Fleisher, H. Fleurbaey, A. Foltynowicz, T. Furtenbacher, M. Germann, E.R. Guest, J.J. Harrison, J.-M. Hartmann, A. Hjältén, S.-M. Hu, X. Huang, T.J. Johnson, H. Jówiak, S. Kassi, M.V. Khan, F. Kwabia-Tchana, T.J. Lee, D. Lisak, A.-W. Liu, O.M. Lyulin, N.A. Malarich, L. Manceron, A.A. Marinina, S.T. Massie, J. Mascio, E.S. Medvedev, V.V. Meshkov, G. Ch. Mellau, M. Melosso, S.N. Mikhailenko, D. Mondelain, H.S.P. Müller, M. ODonnell, A. Owens, A. Perrin, O.L. Polyansky, P.L. Raston, Z.D. Reed, M. Rey, C. Richard, G.B. Rieker, C. Röske, S.W. Sharpe, E. Starikova, N. Stolarczyk, A.V. Stolyarov, K. Sung, F. Tamassia, J. Terragni,

V.G. Ushakov, S. Vasilchenko, B. Vispoel, K.L. Vodopyanov, G. Wagner, S. Wójtewicz, S.N. Yurchenko, and N.F. Zobov. The hitran2024 molecular spectroscopic database. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 353:109807, 2026.

[99] John E. Bertie and Zhida Lan. Infrared intensities of liquids xx: The intensity of the OH stretching band of liquid water revisited, and the best current values of the optical constants of H_2O(l) at 25^o C between 15,000 and 1 cm^{-1}. *Applied Spectroscopy*, 50(8): 1047–1057, 1996.

[100] Sergei Yu. Venyaminov and Franklyn G. Prendergast. Water (H_2O and D_2O) molar absorptivity in the 1000-4000 cm^{-1} range and quantitative infrared spectroscopy of aqueous solutions. *Analytical Biochemistry*, 248(2):234–245, 1997.

[101] Jian Liu, William H. Miller, George S. Fanourgakis, Sotiris S. Xantheas, Sho Imoto, and Shinji Saito. Insights in quantum dynamical effects in the infrared spectroscopy of liquid water from a semiclassical study with an ab initio-based flexible and polarizable force field. *The Journal of Chemical Physics*, 135(24):244503, 12 2011.

[102] B. M. Auer and J. L. Skinner. Hydrogen bonding and raman, IR, and 2D-IR spectroscopy of dilute HOD in liquid D_2O. *Proc. Natl. Acad. Sci. U.S.A.*, 104:14215–14220, 2007.

[103] Benjamin M. Auer and James L. Skinner. Ir and raman spectra of liquid water: Theory and interpretation. *The Journal of Chemical Physics*, 128(22):224511, 2008.

[104] Gregory R. Medders and Francesco Paesani. Infrared and raman spectroscopy of liquid water through "first-principles" many-body molecular dynamics. *Journal of Chemical Theory and Computation*, 11(3):1145–1154, 2015.

[105] Sho Imoto, Sotiris S. Xantheas, and Shinji Saito. Ultrafast dynamics of liquid water: Frequency fluctuations of the OH stretch and the HOH bend. *The Journal of Chemical Physics*, 139(4):044503, 07 2013.

[106] Sho Imoto, Sotiris S. Xantheas, and Shinji Saito. Ultrafast dynamics of liquid water: Energy relaxation and transfer processes of the OH stretch and the HOH bend. *The Journal of Physical Chemistry B*, 119(34):11068–11078, 2015.

[107] Curt M. Breneman and Kenneth B. Wiberg. Determining atom-centered monopoles from molecular electrostatic potentials. the need for high sampling density in formamide conformational analysis. *J. Comput. Chem.*, 11(3):361–373, 1990.

[108] G. D. Purvis III and R. J. Bartlett. A full coupled-cluster singles and doubles model: The inclusion of disconnected triples. *J. Chem. Phys.*, 76(4):1910–1918, 1982.

[109] K. Raghavachari, G. W. Trucks, J. A. Pople, and M. Head-Gordon. A fifth-order perturbation comparison of electron correlation theories. *Chem. Phys. Lett.*, 157(6): 479–483, 1989.

[110] Jr. Dunning, Thom H. Gaussian basis sets for use in correlated molecular calculations. I. the atoms boron through neon and hydrogen. *The Journal of Chemical Physics*, 90 (2):1007–1023, 01 1989.

[111] J. D. Bernal and R. H. Fowler. A theory of water and ionic solution, with particular reference to hydrogen and hydroxyl ions. *The Journal of Chemical Physics*, 1(8): 515–548, 08 1933.

[112] Linus Pauling. The structure and entropy of ice and of other crystals with some randomness of atomic arrangement. *Journal of the American Chemical Society*, 57(12): 2680–2684, 1935.

[113] Sotiris S. Xantheas and Jr. Dunning, Thom H. Ab initio studies of cyclic water clusters $(H_2O)_n$, n=16. I. optimal structures and vibrational spectra. *The Journal of Chemical Physics*, 99(11):8774–8792, 12 1993.

[114] E. Aprà, E. J. Bylaska, W. A. de Jong, N. Govind, K. Kowalski, T. P. Straatsma, M. Valiev, H. J. J. van Dam, Y. Alexeev, J. Anchell, V. Anisimov, F. W. Aquino, R. Atta-Fynn, J. Autschbach, N. P. Bauman, J. C. Becca, D. E. Bernholdt, K. Bhaskaran-Nair, S. Bogatko, P. Borowski, J. Boschen, J. Brabec, A. Bruner, E. Cauët, Y. Chen, G. N. Chuev, C. J. Cramer, J. Daily, M. J. O. Deegan, Jr. Dunning, T. H., M. Dupuis, K. G. Dyall, G. I. Fann, S. A. Fischer, A. Fonari, H. Früchtl, L. Gagliardi, J. Garza, N. Gawande, S. Ghosh, K. Glaesemann, A. W. Götz, J. Hammond, V. Helms, E. D. Hermes, K. Hirao, S. Hirata, M. Jacquelin, L. Jensen, B. G. Johnson, H. Jónsson, R. A. Kendall, M. Klemm, R. Kobayashi, V. Konkov, S. Krishnamoorthy, M. Krishnan, Z. Lin, R. D. Lins, R. J. Littlefield, A. J. Logsdail, K. Lopata, W. Ma, A. V. Marenich, J. Martin del Campo, D. Mejia-Rodriguez, J. E. Moore, J. M. Mullin, T. Nakajima, D. R. Nascimento, J. A. Nichols, P. J. Nichols, J. Nieplocha, A. Otero-de-la Roza, B. Palmer, A. Panyala, T. Pirojsirikul, B. Peng, R. Peverati, J. Pittner, L. Pollack, R. M. Richard, P. Sadayappan, G. C. Schatz, W. A. Shelton, D. W. Silverstein, D. M. A. Smith, T. A. Soares, D. Song, M. Swart, H. L. Taylor, G. S. Thomas, V. Tipparaju, D. G. Truhlar, K. Tsemekhman, T. Van Voorhis, Á. Vázquez-Mayagoitia, P. Verma, O. Villa, A. Vishnu, K. D. Vogiatzis, D. Wang, J. H. Weare, M. J. Williamson, T. L. Windus, K. Woliski, A. T. Wong, Q. Wu, C. Yang, Q. Yu, M. Zacharias, Z. Zhang, Y. Zhao, and R. J. Harrison. NWChem: Past, Present, and Future. *The Journal of Chemical Physics*, 152(18):184102, 05 2020.

[115] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai,

T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox. Gaussian 16 Revision C.01, 2016. Gaussian Inc. Wallingford CT.

[116] John H. Frederick and Clemens Woywod. General formulation of the vibrational kinetic energy operator in internal bond-angle coordinates. *The Journal of Chemical Physics*, 111(16):7255–7271, 10 1999.

[117] Daniel T. Colbert and William H. Miller. A novel discrete variable representation for quantum mechanical reactive scattering via the smatrix kohn method. *The Journal of Chemical Physics*, 96(3):1982–1991, 02 1992.

[118] George H. Gardenier, Mark. A. Johnson, and Anne B. McCoy. Spectroscopic study of the ionradical h–bond in $H_4O_2^+$. *The Journal of Physical Chemistry A*, 113(16): 4772–4779, 2009.

[119] Hua Wei and Jr. Carrington, Tucker. The discrete variable representation of a triatomic hamiltonian in bond lengthbond angle coordinates. *The Journal of Chemical Physics*, 97(5):3029–3037, 09 1992.

[120] Ruth L. Jacobsen, Russell D. III Johnson, Karl K. Irikura, and Raghu N. Kacker. Anharmonic vibrational frequency calculations are not worthwhile for small basis sets. *Journal of Chemical Theory and Computation*, 9(2):951–954, 2013.

[121] Qin Yang, Marco Mendolicchio, Vincenzo Barone, and Julien Bloino. Accuracy and Reliability in the Simulation of Vibrational Spectra: A Comprehensive Benchmark of Energies and Intensities issuing from Generalized Vibrational Perturbation Theory to Second Order (GVPT2). *Frontiers in Astronomy and Space Sciences*, 8:77, May 2021.

[122] A.R. Hoy and P.R. Bunker. A precise solution of the rotation bending schrödinger equation for a triatomic molecule with application to the water molecule. *Journal of Molecular Spectroscopy*, 74(1):1–8, 1979.

[123] Thomas R. Dyke, Kenneth M. Mack, and J. S. Muenter. The structure of water dimer from molecular beam electric resonance spectroscopy. *The Journal of Chemical Physics*, 66(2):498–510, 01 1977.

[124] Joseph R. Lane. Ccsdtq optimized geometry of water dimer. *Journal of Chemical Theory and Computation*, 9(1):316–323, 2013.

[125] Emil Vogt and Henrik G. Kjaergaard. Vibrational spectroscopy of the water dimer at jet-cooled and atmospheric temperatures. *Annual Review of Physical Chemistry*, 73 (Volume 73, 2022):209–231, 2022.

[126] Y. Bouteiller, B. Tremblay, and J.P. Perchard. The vibrational spectrum of the water dimer: Comparison between anharmonic ab initio calculations and neon matrix infrared data between 14,000 and 90 cm^{-1}. *Chemical Physics*, 386(1):29–40, 2011.

[127] K. Ichikawa, Y. Kameda, T. Yamaguchi, H. Wakita, and M. Misawa. Neutron-diffraction investigation of the intramolecular structure of a water molecule in the liquid phase at high temperatures. *Molecular Physics*, 73(1):79–86, 1991.

[128] J.G. Powles. The structure of the water molecule in liquid water. *Molecular Physics*, 42(4):757–765, 1981.

[129] Lawrie B. Skinner, Congcong Huang, Daniel Schlesinger, Lars G. M. Pettersson, Anders Nilsson, and Chris J. Benmore. Benchmark oxygen-oxygen pair-distribution function of ambient water from x-ray diffraction measurements with a wide q-range. *The Journal of Chemical Physics*, 138(7):074506, 02 2013.

[130] W. F. Kuhs and M. S. Lehmann. The structure of the ice ih by neutron diffraction. *The Journal of Physical Chemistry*, 87(21):4312–4313, 1983.

[131] V. F. Petrenko and R. W. Whitworth. *Physics of Ice*. Oxford University Press, Oxford, 1999.

[132] Xiao He, Olaseni Sode, Sotiris S. Xantheas, and So Hirata. Second-order many-body perturbation study of ice Ih. *The Journal of Chemical Physics*, 137(20):204505, 11 2012.

[133] W. F. Kuhs and M. S. Lehmann. Bond-lengths, bond angles and transition barrier in ice Ih by neutron scattering. *Nature*, 294(5840):432–434, 1981.

[134] Victor F. Petrenko and Robert W. Whitworth. *Physics of Ice*. Oxford University Press, 01 2002. ISBN 9780198518945.

[135] Sho Imoto, Sotiris S. Xantheas, and Shinji Saito. Molecular origin of the difference in the HOH bend of the IR spectra between liquid water and ice. *The Journal of Chemical Physics*, 138(5):054506, 02 2013.

[136] J. Paul Devlin, Joanna Sadlej, and Victoria Buch. Infrared spectra of large h_2o clusters: New understanding of the elusive bending mode of ice. *The Journal of Physical Chemistry A*, 105(6):974–983, 2001.

[137] J. E. Bertie and E. Whalley. Infrared spectra of ices Ih and Ic in the range 4000 to 350 cm^{-1}. *The Journal of Chemical Physics*, 40(6):1637–1645, 03 1964.

[138] Edward Whalley. A detailed assignment of the OH stretching bands of ice I. *Canadian Journal of Chemistry*, 55:3429–3441, 1977.

[139] M. J. Taylor and E. Whalley. Raman spectra of ices Ih, Ic, II, III, and V. *The Journal of Chemical Physics*, 40(6):1660–1664, 03 1964.

[140] Daniel R. Moberg, Shelby C. Straight, Christopher Knight, and Francesco Paesani. Molecular origin of the vibrational structure of ice Ih. *The Journal of Physical Chemistry Letters*, 8(12):2579–2583, 2017.

[141] Liang Shi, J. L. Skinner, and Thomas L. C. Jansen. Two-dimensional infrared spectroscopy of neat ice ih. *Phys. Chem. Chem. Phys.*, 18:3772–3779, 2016.

[142] Qiang Sun and Haifei Zheng. Raman OH stretching vibration of ice ih. *Progress in Natural Science*, 19(11):1651–1654, 2009.

[143] Alexander A. Kananenka and James L. Skinner. Fermi resonance in OH-stretch vibrational spectroscopy of liquid water and the water hexamer. *The Journal of Chemical Physics*, 148(24):244107, 2018.

[144] Luigi De Marco, Krupa Ramasesha, and Andrei Tokmakoff. Experimental evidence of Fermi resonances in isotopically dilute water from ultrafast broadband IR spectroscopy. *The Journal of Physical Chemistry B*, 117(49):15319–15327, 2013.

[145] Zhaohui Wang, Andrei Pakoulev, Yoonsoo Pang, and Dana D. Dlott. Vibrational substructure in the OH stretching transition of water and HOD. *The Journal of Physical Chemistry A*, 108(42):9054–90, 2004.

[146] Francesco Paesani. Getting the right answers for the right reasons: Toward predictive molecular simulations of water with many-body potential energy functions. *Accounts of Chemical Research*, 49(9):1844–1851, 2016.

[147] Yves Maréchal. The molecular structure of liquid water delivered by absorption spectroscopy in the whole IR region completed with thermodynamics data. *Journal of Molecular Structure*, 1004(1):146–155, 2011.

[148] Sietse T. van der Post, Cho-Shuen Hsieh, Masanari Okuno, Yuki Nagata, Huib J. Bakker, Mischa Bonn, and Johannes Hunger. Strong frequency dependence of vibrational relaxation in bulk and surface water reveals sub-picosecond structural heterogeneity. *Nature Communications*, 6:8384, 2015.

[149] Takakazu Seki, Kuo-Yang Chiang, Chun-Chieh Yu, Xiaoqing Yu, Masanari Okuno, Johannes Hunger, Yuki Nagata, and Mischa Bonn. The bending mode of water: A

powerful probe for hydrogen bond structure of aqueous systems. *The Journal of Physical Chemistry Letters*, 11(19):8459–8469, 2020.

[150] Gregory R. Medders and Francesco Paesani. Infrared and Raman spectroscopy of liquid water through "first-principles" many-body molecular dynamics. *Journal of Chemical Theory and Computation*, 11(3):1145–1154, 2015.

[151] Richard M. Badger and Simon H. Bauer. Spectroscopic studies of the hydrogen bond. ii. the shift of the oh vibrational frequency in the formation of the hydrogen bond. *The Journal of Chemical Physics*, 5(11):839–851, 11 1937.

[152] Mark A. Boyer, Ondrej Marsalek, Joseph P. Heindel, Thomas E. Markland, Anne B. McCoy, and Sotiris S. Xantheas. Beyond badgers rule: The origins and generality of the structurespectra relationship of aqueous hydrogen bonds. *The Journal of Physical Chemistry Letters*, 10(5):918–924, 2019.

[153] Jr. Wilson, E. Bright. Some mathematical methods for the study of molecular vibrations. *The Journal of Chemical Physics*, 9(1):76–84, 01 1941.

[154] George S. Fanourgakis and Sotiris S. Xantheas. Development of transferable interaction potentials for water. V. extension of the flexible, polarizable, thole-type model potential (TTM3-F, v. 3.0) to describe the vibrational spectra of water clusters and liquid water. *The Journal of Chemical Physics*, 128(7):074506, 02 2008.

[155] Curt M. Breneman and Kenneth B. Wiberg. Determining atom-centered monopoles from molecular electrostatic potentials. the need for high sampling density in formamide conformational analysis. *Journal of Computational Chemistry*, 11(3):361–373, 1990.

[156] Shelley L. Shostak, William L. Ebenstein, and J. S. Muenter. The dipole moment of water. I. dipole moments and hyperfine properties of H_2O and HDO in the ground and excited vibrational states. *The Journal of Chemical Physics*, 94(9):5875–5882, 05 1991.

[157] Shelley L. Shostak and J. S. Muenter. The dipole moment of water. II. analysis of the vibrational dependence of the dipole moment in terms of a dipole moment function. *The Journal of Chemical Physics*, 94(9):5883–5890, 05 1991.

[158] Shepard A. Clough, Yardley Beers, Gerald P. Klein, and Laurence S. Rothman. Dipole moment of water from stark measurements of H_2O, HDO, and D_2O. *The Journal of Chemical Physics*, 59(5):2254–2259, 09 1973.

[159] Thomas R. Dyke and J. S. Muenter. Electric dipole moments of low J states of H_2O and D_2O. *The Journal of Chemical Physics*, 59(6):3125–3127, 09 1973.

[160] Y. S. Badyal, M.-L. Saboungi, D. L. Price, S. D. Shastri, D. R. Haeffner, and A. K. Soper. Electron distribution in water. *The Journal of Chemical Physics*, 112(21): 9206–9208, 06 2000.

[161] Enrique R. Batista, Sotiris S. Xantheas, and Hannes Jósson. Molecular multipole moments of water molecules in ice Ih. *The Journal of Chemical Physics*, 109(11): 4546–4551, 09 1998.

[162] Enrique R. Batista, Sotiris S. Xantheas, and Hannes Jósson. Multipole moments of water molecules in clusters and ice Ih from first principles calculations. *The Journal of Chemical Physics*, 111(13):6011–6015, 10 1999.

[163] Enrique R. Batista, Sotiris S. Xantheas, and Hannes Jósson. Electric fields in ice and near water clusters. *The Journal of Chemical Physics*, 112(7):3285–3292, 02 2000.

[164] Volodymyr Babin, Claude Leforestier, and Francesco Paesani. Development of a "first-principles" water potential with flexible monomers. I. dimer potential energy surface, vrt spectrum, and second virial coefficient. *Journal of Chemical Theory and Computation*, 9(12), 2013.

[165] Qi Yu, Chen Qu, Paul L. Houston, Riccardo Conte, Apurba Nandi, and Joel M. Bowman. q-aqua: A many-body CCSD(T) water potential, including four-body interactions,

demonstrates the quantum nature of water from clusters to the liquid phase. *The Journal of Physical Chemistry Letters*, 13(22):5068–5074, 2022.

[166] Emma Sigfridsson and Ulf Ryde. Comparison of methods for deriving atomic charges from the electrostatic potential and moments. *Journal of Computational Chemistry*, 19 (4):377–395, March 1998.

[167] Z. S. Huang and R. E. Miller. High-resolution near-infrared spectroscopy of water dimer. *The Journal of Chemical Physics*, 91(11):6613–6631, 1989.

[168] R. S. Fellers, L. B. Braly, R. J. Saykally, and C. Leforestier. Fully coupled six-dimensional calculations of the water dimer vibration-rotation-tunneling states with split wigner pseudospectral approach. ii. improvements and tests of additional potentials. *The Journal of Chemical Physics*, 110(13):6306–6318, 1999.

[169] A. Mukhopadhyay, W. T. S. Cole, and R. J. Saykally. The water dimer i: Experimental characterization. *Chemical Physics Letters*, 633:13–26, 2015.

[170] G. E. Douberly, R. E. Miller, and S. S. Xantheas. Formation of exotic networks of water clusters in helium droplets facilitated by the presence of neon atoms. *Journal of the American Chemical Society*, 139(11):4152–4156, 2017.

[171] A. Mukhopadhyay, S. S. Xantheas, and R. J. Saykally. The water dimer ii: Theoretical investigations. *Chemical Physics Letters*, 700:163–175, 2018.

[172] M. A. Suhm, Jr. Farrell, J. T., A. McIlroy, and D. J. Nesbitt. High resolution 1.3 μm overtone spectroscopy of hf dimer in a slit jet: Ka=0$\leftarrow$0 and ka=1$\leftarrow$0 subbands of $u_{acc} = 2 \leftarrow 0$. *The Journal of Chemical Physics*, 97(8):5341–5354, 1992.

[173] K. A. Peterson and Jr. Dunning, T. H. Benchmark calculations with correlated molecular wave functions. vii. binding energy and structure of the hf dimer. *The Journal of Chemical Physics*, 102(5):2032–2041, 1995.

[174] E. Miliordos, E. Apra, and S. S. Xantheas. Optimal geometries and harmonic vibrational frequencies of the global minima of water clusters $(h_2o)_n$, n=2–6, and several hexamer local minima at the ccsd(t) level of theory. *The Journal of Chemical Physics*, 139(11), 2013.

[175] G. D. Santis and S. S. Xantheas. Extending badger's rule. i. the relationship between energy and structure in hydrogen bonds. *The Journal of Chemical Physics*, 162(4), 2025.

[176] G. D. Santis and S. S. Xantheas. Extending badger's rule. ii. the relationship between energy and vibrational spectra in hydrogen bonds. *The Journal of Chemical Physics*, 162(24), 2025.

[177] A. V. Iogansen and B. V. Rassadin. Dependence of intensity and shift in the infrared bands of the hydroxyl group on hydrogen bonding energy. *Journal of Applied Spectroscopy*, 11(5):1318–1325, 1969.

[178] A. V. Iogansen. Direct proportionality of the hydrogen bonding energy and the intensification of the stretching u(xh) vibration in infrared spectra. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 55(7):1585–1612, 1999.

[179] G. D. Santis and S. S. Xantheas. On the strength of ouroboros hydrogen bonds. *The Journal of Physical Chemistry A*, 129(43):9930–9941, 2025.

[180] K. Ruedenberg, M. W. Schmidt, M. M. Gilbert, and S. T. Elbert. Are atoms intrinsic to molecular electronic wavefunctions? i. the fors model. *Chemical Physics*, 71(1): 41–49, 1982.

[181] J. Olsen, B. O. Roos, P. Jørgensen, and H. J. A. Jensen. Determinant based configuration interaction algorithms for complete and restricted configuration interaction spaces. *The Journal of Chemical Physics*, 89(4):2185–2192, 1988.

[182] I. Shavitt. The unitary group and the electron correlation problem. In P.-O. L"owdin and B. Pullman, editors, *New Horizons of Quantum Chemistry*, pages 279–293. Springer Netherlands, Dordrecht, 1983.

[183] C. D. Sherrill, A. I. Krylov, E. F. C. Byrd, and M. Head-Gordon. Energies and analytic gradients for a coupled-cluster doubles model using variational brueckner orbitals: Application to symmetry breaking in o_4^+. *The Journal of Chemical Physics*, 109(11): 4171–4181, 1998.

[184] G. M. J. Barca, C. Bertoni, L. Carrington, D. Datta, N. De Silva, J. E. Deustua, D. G. Fedorov, J. R. Gour, A. O. Gunina, E. Guidez, et al. Recent developments in the general atomic and molecular electronic structure system. *The Journal of Chemical Physics*, 152(15), 2020.

[185] H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, and M. Sch"utz. Molpro: a general-purpose quantum chemistry program package. *WIREs Computational Molecular Science*, 2(2):242–253, 2012.

[186] H.-J. Werner, P. J. Knowles, F. R. Manby, J. A. Black, K. Doll, A. Heßelmann, D. Kats, A. K"ohn, T. Korona, D. A. Kreplin, et al. The molpro quantum chemistry package. *The Journal of Chemical Physics*, 152(14), 2020.

[187] B. E. Rocher-Casterline, L. C. Ch'ng, A. K. Mollner, and H. Reisler. Communication: Determination of the bond dissociation energy (d_0) of the water dimer, $(h_2o)_2$, by velocity map imaging. *The Journal of Chemical Physics*, 134(21), 2011.

[188] S. Scheiner. Ab initio studies of hydrogen bonds: The water dimer paradigm. *Annual Review of Physical Chemistry*, 45:23–56, 1994.

[189] A. Halkier, H. Koch, P. Jørgensen, O. Christiansen, I. M. B. Nielsen, and T. Helgaker. A systematic ab initio study of the water dimer in hierarchies of basis sets and correlation models. *Theoretical Chemistry Accounts*, 97(1):150–157, 1997.

[190] T. A. Ruden, T. Helgaker, P. Jørgensen, and J. Olsen. Coupled-cluster connected quadruples and quintuples corrections to the harmonic vibrational frequencies and equilibrium bond distances of hf, n_2, f_2, and co. *The Journal of Chemical Physics*, 121 (12):5874–5884, 2004.

[191] D. Feller and K. A. Peterson. High level coupled cluster determination of the structure, frequencies, and heat of formation of water. *The Journal of Chemical Physics*, 131(15): 154306, 2009.

[192] R. Kalescky, W. Zou, E. Kraka, and D. Cremer. Local vibrational modes of the water dimer – comparison of theory and experiment. *Chemical Physics Letters*, 554:243–247, 2012.

[193] R. I. Ovsyannikov, V. Y. Makhnev, N. F. Zobov, J. Koput, J. Tennyson, and O. L. Polyansky. Highly accurate hf dimer ab initio potential energy surface. *The Journal of Chemical Physics*, 156(16), 2022.

[194] J. C. Howard, J. L. Gray, A. J. Hardwick, L. T. Nguyen, and G. S. Tschumper. Getting down to the fundamentals of hydrogen bonding: Anharmonic vibrational frequencies of $(hf)_2$ and $(h_2o)_2$ from ab initio electronic structure computations. *Journal of Chemical Theory and Computation*, 10(12):5426–5435, 2014.

[195] A. S. Pine, W. J. Lafferty, and B. J. Howard. Vibrational predissociation, tunneling, and rotational saturation in the hf and df dimers. *The Journal of Chemical Physics*, 81 (7):2939–2950, 1984.

[196] T. M. Sexton, J. C. Howard, and G. S. Tschumper. Dissociation energy of the $h_2o\cdots hf$ dimer. *The Journal of Physical Chemistry A*, 122(21):4902–4908, 2018.

[197] A. C. Legon, D. J. Millen, and H. M. North. Experimental determination of the dissociation energies d_0 and d_e of $h_2o\cdots hf$. *Chemical Physics Letters*, 135(3):303–306, 1987.

[198] R. K. Thomas. Hydrogen bonding in the vapour phase between water and hydrogen fluoride: the infrared spectrum of the 1:1 complex. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 344(1639):579–592, 1975.

[199] V. P. Bulychev, I. M. Grigoriev, E. I. Gromova, and K. G. Tokhadze. Study of the ν_1 band shape of the $h_2o\cdots hf$, $h_2o\cdots df$, and $h_2o\cdots hcl$ complexes in the gas phase. *Physical Chemistry Chemical Physics*, 7(11):2266–2278, 2005.

[200] Mohammad Reza Jangrouei, Agnieszka Krzemi'nska, Michał Hapka, Ewa Pastorczak, and Katarzyna Pernal. Dispersion interactions in exciton-localized states. theory and applications to π–π and n–π excited states. *Journal of Chemical Theory and Computation*, 18(6):3497–3511, 2022. doi: 10.1021/acs.jctc.2c00221.

[201] Michał Hapka, Agnieszka Krzemi'nska, and Katarzyna Pernal. How much dispersion energy is included in the multiconfigurational interaction energy? *Journal of Chemical Theory and Computation*, 16(10):6280–6293, 2020. doi: 10.1021/acs.jctc.0c00681.

[202] Lu T. Xu and Jr. Dunning, Thom H. A cautionary tale: Problems in the valence-casscf description of the ground state ($X^1\Sigma^+$) of bf. *The Journal of Chemical Physics*, 153 (11):114113, 2020. doi: 10.1063/5.0024134.

[203] W. C. Lu, C. Z. Wang, M. W. Schmidt, L. Bytautas, K. M. Ho, and K. Ruedenberg. Molecule intrinsic minimal basis sets. i. exact resolution of *ab initio* optimized molecular orbitals in terms of deformed atomic minimal-basis orbitals. *The Journal of Chemical Physics*, 120(6):2629–2637, 2004.

[204] W. C. Lu, C. Z. Wang, M. W. Schmidt, L. Bytautas, K. M. Ho, and K. Ruedenberg. Molecule intrinsic minimal basis sets. ii. bonding analyses for si_4h_6 and si_2 to si_{10}. *The Journal of Chemical Physics*, 120(6):2638–2651, 2004.

[205] M. W. Schmidt, E. A. Hull, and T. L. Windus. Valence virtual orbitals: An unambiguous

ab initio quantification of the lumo concept. *The Journal of Physical Chemistry A*, 119 (41):10408–10427, 2015.