EMBEDDED-FRAGMENT STUDIES OF ICE, WATER, AND MOLECULAR CLUSTER CATIONS

BY

MICHAEL SALIM

DISSERTATION

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Chemistry with a concentration in Computational Science and Engineering in the Graduate College of the University of Illinois at Urbana-Champaign, 2017

Urbana, Illinois

Doctoral Committee:

Professor So Hirata, Chair
Professor Nancy Makri
Professor Catherine Murphy
Associate Professor Dallas Trinkle
The embedded-fragment scheme known as the binary interaction method (BIM) was extended to enable *ab initio* second-order many-body perturbation (MP2) theory simulations of molecular liquids and solids with nonorthogonal lattice vectors. Three related anomalies of water ice Ih were investigated with BIM-MP2. The negative thermal expansion and pressure-induced amorphization of ice Ih were computationally observed and linked to hydrogen bond-bending acoustic phonons that soften with increasing pressure. The anomalous volume isotope effect, where D$_2$O possesses a larger molar volume than H$_2$O ice, was determined to be a more complex phenomenon relying on cancellations between volume-shrinking and expanding contributions from several anharmonic mode groups.

In application of BIM-MP2 to the Gibbs free energy of other ice phases, it was found that basis sets larger than aug-cc-pVDZ are essential to predict reasonable lattice energies and bulk moduli of the ices. While the smaller aug-cc-pVDZ basis set predicts a nearly backwards lattice energy ordering among the ices Ih, II, III/IX, V, and VI, the order is largely corrected by moving towards the complete basis set limit (CBS). The bulk modulus is also systematically reduced to accurate values, although density continues to be overestimated by BIM-MP2 with the atomic embedding field. The phase diagram of solid water proves too difficult even for a qualitatively correct prediction by BIM-MP2/CBS, although the small free energy differences between phases may be improved by a simple increase of BIM lattice truncation radii or disorder-averaging lattice energy over proton configurations of the H-disordered ices.

BIM-MP2 with spin component scaling was used to generate on-the-fly forces for *ab initio* molecular dynamics of liquid water at constant volume and temperature. The simulated structure of the first and second coordination shells, as judged by the oxygen-oxygen radial distribution, is in excellent agreement with experiment. This requires a slightly lower than ambient temperature of 250 K and fixed density, however, resulting in an average pressure of -0.6 GPa. The significant negative pressure exerted to maintain the experimental density of water is related to the overestimated dispersion interactions of MP2. The simulated diffusivity, infrared, and Raman spectra (especially intensity and line width) were in reasonable agreement with the observed. The difference between isotropic and anisotropic Raman components was traced to the diverse hydrogen-bonding environments experienced by the waters, which is corroborated by the broad distribution of monomer geometries and induced molecular dipole moments. The strong dispersion interaction of BIM-MP2 stabilizes the occupation of a fifth non-hydrogen-bonded location relative to each tetra-coordinated water molecule. Hence,
the water structure is significantly less ice-like and the probability of tetra- and penta-
coordination were found to be nearly equal. The transient attachment, detachment, and
exchange of this weakly bound fifth molecule seems to play a large role in fluctuations
of the hydrogen bonding network.

Because the viability of BIM and related fragment methods depends on weakly in-
teracting monomers with integer electron counts, the methodology has been restricted
to molecular crystals, leaving systems with significant electron delocalization out of
reach. As a first step towards handling significant charge transfer in a BIM-like frame-
work, the valence bond charge transfer (VBCT) method was proposed for molecular
cluster cations. VBCT approaches construction of an effective Hamiltonian in a basis of
charge-localized configurations, which are intuitively amenable to fragmentation. The
Hamiltonian matrix elements are directly defined in terms of embedded-fragment cal-
culations of monomers and dimers with the goal of approximating the cation ground
state energy. VBCT was applied to the UHF and UMP2 energies of several helium,
argon, ethylene, water, and water-sodium clusters to evaluate its performance. The
preliminary results are generally encouraging, although erratic behaviors are observed
and related to the UHF symmetry breaking in significantly multireference systems. In
regions of the potential energy surface where charge is delocalized or localized without
the existence of near-degenerate UHF solutions, VBCT can reproduce UHF and UMP2
ground state energies quite reliably.
To Baba, Papa, and Dr. Havey.
Acknowledgements

First, I must acknowledge my advisor, So Hirata, for introducing me to frontiers in computational chemistry and sharing such a contagious enthusiasm for the subject. I could not imagine learning so much in graduate school or completing this thesis without his guidance. I am eternally grateful for all his patience and encouragement along the way.

I have been fortunate to share the lab with such good friends and inspiring colleagues. I owe a lot to former group members Ryan and Matt, who taught me so much about chemistry, useful algorithms, and scientific computing in general. It was exciting to collaborate with Soohaeng on the molecular dynamics project and see a long-standing goal of the lab come to fruition. His expertise in the subject truly broadened my horizons. My current labmates Cole, Alex, Jake, and Evan have shared a truly infectious passion for all things computing. They’ve inspired me to learn about new technologies and strive to become a more well-rounded programmer, not to mention the number of times they’ve helped me around a puzzling issue in my own research. The new lab postdocs Alex and Jun strike me as truly exceptional scientists, and I am excited to read about their work in the coming years.

I must also acknowledge all of the great teachers and mentors I’ve had here and at my undergrad institution, James Madison University. My first advisor, Kevin Caran, was the reason I tried research in the first place. To this day, I owe my excitement about chemistry to the completely new perspectives that he shared with me. I had more than a few wonderful teachers in physical chemistry and the Mathematics department who shaped my decision to try computational chemistry. I am indebted to my first quantum chemistry teacher, Daniel Havey, who believed so strongly in me and supported my graduate school applications.

Most importantly, I must thank my family and friends who mean everything to me. My mom and dad have been endlessly loving and supportive. I have sincerely missed my best friends and sister from home and worked in part to make them proud. My friends and roommates in Champaign have been amazing and truly made this town feel like a second home. I am lucky to have met them and excited to see what their futures have in store.
# Table of Contents

List of Abbreviations ........................................ viii

1 Introduction ................................................. 1
  1.1 Background ........................................... 1
  1.2 Molecular electronic structure .......................... 3
    1.2.1 The molecular Hamiltonian ......................... 3
    1.2.2 The Born–Oppenheimer approximation ............... 4
  1.3 The many-body expansion ................................ 5
    1.3.1 Total electronic energy ............................ 5
    1.3.2 Electrostatic embedding ............................ 6
  1.4 The binary interaction method ......................... 7
    1.4.1 Energy ........................................... 7
    1.4.2 Energy gradient and Hessian ....................... 9

2 Ice Ih anomalies ........................................... 11
  2.1 Introduction .......................................... 11
  2.2 Computational methods .................................. 13
    2.2.1 Electronic structures .............................. 13
    2.2.2 Quasiharmonic approximation ...................... 14
    2.2.3 Born–Oppenheimer molecular dynamics ............ 16
  2.3 Results and discussion ................................ 16
    2.3.1 Structure and lattice energy ...................... 16
    2.3.2 Thermal contraction and expansion ............... 16
    2.3.3 Bulk modulus .................................... 22
    2.3.4 Anomalous volume isotope effect ................. 22
    2.3.5 Pressure-induced amorphization ................. 25
  2.4 Embedding-field and basis-set dependence .......... 30
  2.5 Conclusions ........................................... 33

3 Water ices beyond ambient pressure ..................... 35
  3.1 Introduction .......................................... 35
  3.2 Computational methods ................................ 36
    3.2.1 Embedded-fragment crystal optimization and CBS energy 36
    3.2.2 Phonon dispersion ................................ 37
    3.2.3 Quasiharmonic approximation .................... 37
  3.3 Results ................................................ 38
    3.3.1 Structural optimization and lattice energy .......... 38
    3.3.2 Inelastic neutron scattering intensity ............ 39
    3.3.3 Pseudo-translational mode analysis .............. 41
    3.3.4 Thermal expansion ................................ 44
    3.3.5 Equation of state ................................ 45
    3.3.6 Phase diagram .................................. 46
  3.4 Conclusions ........................................... 52
4 Molecular dynamics of liquid water .................................. 53
4.1 Introduction .......................................................... 53
4.2 Methods .............................................................. 55
  4.2.1 Born-Oppenheimer molecular dynamics ....................... 55
  4.2.2 Self-diffusion coefficient .................................... 55
  4.2.3 IR and Raman spectra ........................................... 55
  4.2.4 Energy and gradient evaluation ............................... 56
4.3 Results .............................................................. 57
  4.3.1 Potential energy surface ....................................... 57
  4.3.2 Radial distribution function ................................. 59
  4.3.3 Self-diffusion coefficient .................................... 60
  4.3.4 Water molecular geometry in the liquid phase ................. 60
  4.3.5 Coordination numbers and intershell exchange ............... 61
  4.3.6 Dipole moments ............................................ 65
  4.3.7 IR and Raman spectra ........................................ 65
4.4 Conclusions ........................................................ 68

5 Embedded-fragment valence bond charge transfer method ........ 70
5.1 Introduction ........................................................ 70
5.2 Computational methods .......................................... 72
  5.2.1 Diagonal elements ........................................... 72
  5.2.2 Off-diagonal elements ....................................... 74
  5.2.3 Total energy and charge distribution ......................... 75
  5.2.4 Rationale .................................................... 76
5.3 Results .............................................................. 76
  5.3.1 He$_2^+$ .................................................... 76
  5.3.2 He$_3^+$ .................................................... 78
  5.3.3 He$_4^+$ .................................................... 79
  5.3.4 He$_5^+$ .................................................... 82
  5.3.5 Ar$_4^+$ .................................................... 82
  5.3.6 [Na(H$_2$O)$_3$]$^+$ ........................................ 85
  5.3.7 (C$_2$H$_4$)$_3^+$ ........................................... 86
  5.3.8 (C$_2$H$_4$)$_4^+$ ........................................... 89
  5.3.9 (H$_2$O)$_4^+$ ............................................... 92
5.4 Conclusions ........................................................ 94

6 Conclusions .......................................................... 96

References ............................................................... 99
# List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIM</td>
<td>Binary-interaction method.</td>
</tr>
<tr>
<td>BOMD</td>
<td>Born–Oppenheimer molecular dynamics.</td>
</tr>
<tr>
<td>CBS</td>
<td>Complete basis set.</td>
</tr>
<tr>
<td>CC</td>
<td>Coupled cluster.</td>
</tr>
<tr>
<td>DFT</td>
<td>Density-functional theory.</td>
</tr>
<tr>
<td>HF</td>
<td>Hartree–Fock.</td>
</tr>
<tr>
<td>MBPT</td>
<td>Many-body perturbation theory.</td>
</tr>
<tr>
<td>MBE</td>
<td>Many-body expansion.</td>
</tr>
<tr>
<td>MO</td>
<td>Molecular orbital.</td>
</tr>
<tr>
<td>MP</td>
<td>Møller-Plesset.</td>
</tr>
<tr>
<td>MPI</td>
<td>Message-Passing Interface.</td>
</tr>
<tr>
<td>PES</td>
<td>Potential energy surface.</td>
</tr>
<tr>
<td>QHA</td>
<td>Quasiharmonic approximation.</td>
</tr>
<tr>
<td>UHF</td>
<td>Unrestricted Hartree–Fock.</td>
</tr>
<tr>
<td>VBCT</td>
<td>Valence bond charge transfer.</td>
</tr>
<tr>
<td>ZPE</td>
<td>Zero-point energy.</td>
</tr>
</tbody>
</table>
1 Introduction

1.1 Background

On the atomic scale, molecular clusters, liquids, and solids can be partitioned into discrete molecular units, where each tightly binds a whole number of electrons and weakly interacts with other molecules in its vicinity. The weak interactions are often loosely categorized as hydrogen-bond, electrostatic, or dispersion forces, and they dictate the physical properties of ubiquitous substances like water. At sufficiently high pressure, water and other inorganic molecules, like carbon dioxide or dihydrogen, form molecular crystals (ices) that play an important role in the geochemistry of planets.\textsuperscript{1, 2} Liquid water, of course, is central to biology and yet many of its fundamental properties remain poorly understood.\textsuperscript{3, 4}

From the interplay of many weak interactions comes an extraordinary complexity to the structure and stability of molecular crystals. When different structures occupy neighboring regions of stability in the phase diagram, the solid state exhibits polymorphism, where various solid phases can exist stably or coexist metastably. In the pharmaceutical industry, undesired crystalline polymorphs of a drug with low bioavailability are a major problem.\textsuperscript{5} This polymorphism arises in solid water, too, which has no less than fifteen crystalline forms formed experimentally at high pressure or low temperature.\textsuperscript{6}

Characterizing the structure and intermolecular forces in molecular crystals is experimentally challenging and perhaps even more difficult in liquids. It requires building a coherent model from diffraction experiments, vibrational spectroscopy, and equation of state measurements.\textsuperscript{1, 2} In molecular crystals, polymorphism and metastability complicates the isolation of single crystals and delineation of phase boundaries. Probing phases under extreme conditions is also inherently difficult; for instance, studies of water under hydrostatic compression require the use of pressure-transmitting fluids that can either freeze or form unwanted inclusions.\textsuperscript{7} The vibrational density of states probed by inelastic neutron scattering is often difficult to interpret without computational assistance.\textsuperscript{8–10}

For these reasons, atomistic computer simulations that go beyond parameterized force fields\textsuperscript{11, 12} can serve as an important tool for guiding and interpreting experimental measurements or even making independent predictions.\textsuperscript{13} Analytic force fields have been enormously successful, and indeed they sometimes continue to outperform first-principles quantum mechanical approaches in areas like the phase diagram of water.\textsuperscript{14–16} However, these methods are intrinsically limited to the set of empirical or \textit{ab}
**initio** reference data to which they are trained, and systematic improvement is difficult in systems for which little reference data exists. Moreover, approaches that neglect the electronic degrees of freedom cannot faithfully simulate spectroscopic intensities unless a dipole moment surface is parameterized along with the potential energy surface.\(^{12}\)

Density functional theory (DFT) is a first-principles quantum mechanical approach that has been tremendously successful in all types of condensed-phase simulations.\(^{17}\) Although it has none of the limitations of parameterized force fields, DFT exhibits serious deficiencies when dispersion (van der Waals) interactions contribute significantly to the electronic energy. Semilocal exchange correlation functionals using the generalized gradient approximation with or without hybrid exchange do not describe the attractive dispersion that decays as \(R^{-6}\) in the intermolecular separation \(R\). Without this essential cohesive interaction, simulations of water tend to predict a liquid that is not dense enough and overstructured with low diffusivity.\(^{18}\) For solid water, the same methods grossly overestimate energy differences between the low and high pressure phases of ice, because it is the missing dispersion interaction that stabilizes water in interstitial spaces that become filled at high density.\(^{17}\) Alternative dispersion-inclusive DFT methods have therefore been developed which improve substantially on the conventional functionals for describing water. However, the errors incurred by various dispersion-corrected approaches are still not systematically controllable, and according to a recent perspective “erratic differences between the various representations of dispersion already manifest themselves in the energetics of the (water) dimer.”\(^{17}\)

In this study, therefore, we resort to electronic structure methods based on *ab initio* molecular orbital (MO) theory. These methods include various hierarchies of approximations to electron correlation (such as many-body perturbation (MP) theory and coupled cluster (CC) theory) that can be systematically converged to the exact solution of the electronic Schrödinger equation. The simplest of these approaches, second-order MP (MP2) theory, is fortunately already quite accurate for a variety molecular systems including water,\(^{2,19}\) describing all covalent, electrostatic, and dispersion interactions on equal footing. The main drawback of MO-based methods is the steep scaling of computational cost with system size, increasing at least as the fifth power of the number of molecules. This is usually so prohibitive that systems of a mere 50-100 atoms are too large for any meaningful investigation of the potential energy surface. Consequently, there has been considerable effort in devising efficient and reduced-scaling implementations to enable application of these MO theories to larger systems.

The approach followed in this work is an example of the so-called *fragmentation* methods, where a large molecular system is broken into smaller fragment calculations from which full-system properties are predicted by reconstruction in one sense or another.\(^{2,5,19–21}\) In the remaining sections of this Introduction, the fragmentation scheme known as the binary interaction method\(^{2,22–24}\) (BIM) is introduced. It has proven versatility for a variety of molecular condensed phase calculations, ranging from the thermodynamic properties of carbon dioxide ices\(^2\) to the Born–Oppenheimer molecular dynamics simulation of Raman spectra in liquid water.\(^{25}\) The BIM is systematic in the
sense that errors related to fragmentation and electron correlation theory can be controlled, at least in principle. In practice, it is accurate and efficient, having enabled MP2 calculations near the complete basis set limit for condensed water with an asymptotically linear-scaling algorithm. With distributed-memory parallelism, the method scales well from small clusters to large supercomputers with good parallel efficiency.

In Chapter 2, BIM-MP2 is used to study several thermodynamic anomalies of water in the predominant solid phase on earth: ice Ih. Calculations of the phonon density of states and the quasiharmonic approximation are employed to analyze the origins of negative thermal expansion, the anomalous effect of deuterium substitution on volume, and the pressure-induced amorphization of ice Ih. In Chapter 3, BIM-MP2 is then applied to the higher-pressure crystalline forms of water: ices II, III, IX, V, and VI. The performance of the method in predicting relative phase stabilities via the quasiharmonic Gibbs free energy is evaluated. In Chapter 4, we illustrate the application of BIM-MP2 as a force field for ab initio molecular dynamics of liquid water. Predictions of the liquid structure, diffusivity, and vibrational spectra are compared with experimental data.

In Chapter 5, we begin to address the serious restriction of BIM and related fragment methods to weakly interacting molecular crystals. As explained in the following sections, fragmentation relies on rapid convergence of the many-body expansion of the total electronic energy. This is generally possible only when the system consists of weakly interacting fragments with whole number electron counts. Thus ionic, covalent, and especially metallic solid electronic structures are presently outside the scope of BIM. Inspired by valence bond-type methods and fragment approaches to cluster ionization potentials, we propose a new scheme in the spirit of BIM for the ground state of singly-ionized molecular cluster cations with significant charge transfer. Finally, Chapter 6 closes with a summary of the present work and outlook for future developments and applications of fragment methods.

1.2 Molecular electronic structure

1.2.1 The molecular Hamiltonian

The starting point of quantum chemistry is the nonrelativistic time-independent Schrödinger equation, which describes the quantized motion of electrons and point nuclei interacting through the electric potential:

$$\hat{H}\Psi(r, R) = E\Psi(r, R), \quad (1.1)$$

where $\hat{H}$ is the molecular Hamiltonian containing kinetic and potential energy operators for the electrons and nuclei. In atomic units, the Hamiltonian for a molecule of $N$ electrons and $M$ nuclei reads
\[
\hat{H} = -\sum_{i=1}^{N} \frac{1}{2} \nabla_i^2 - \sum_{A=1}^{M} \frac{1}{2M_A} \nabla_A^2 - \sum_{i=1}^{N} \sum_{A=1}^{M} \frac{Z_A}{r_{iA}} \\
+ \sum_{i=1}^{N} \sum_{j>i}^{N} \frac{1}{r_{ij}} + \sum_{A=1}^{M} \sum_{B=A}^{M} \frac{Z_A Z_B}{R_{AB}}
\]

(1.2)

In this equation, \( i \) and \( j \) index electrons while \( A \) and \( B \) index nuclei. The first two terms containing Laplacians are the kinetic energy operators for the electrons and nuclei, where \( M_A \) is the ratio of the \( A \)th nucleus mass to the electron mass. The last three terms contain all Coulomb interactions among the electrons and nuclei, where \( Z_A \) is the nuclear charge of the \( A \)th nucleus, and \( r_{iA}, r_{ij}, \) and \( r_{AB} \) are electron-nucleus, electron-electron, and nucleus-nucleus distances, respectively.

### 1.2.2 The Born–Oppenheimer approximation

We seek approximate solutions to equation 1.1 for the molecular energy \( E \) and wave function \( \Psi(r, R) \) expressed in electronic (\( r \)) and nuclear (\( R \)) position coordinates. To make progress in systems containing multiple electrons and nuclei, it is essential to invoke the Born-Oppenheimer approximation, which entails a factorization of \( \Psi(r, R) \) into an electronic wave function \( \Psi_{\text{elec}}(r; R) \) that only depends parametrically on nuclear coordinates and a nuclear wave function \( \Psi_{\text{nuc}}(R) \). This approximate separation of variables is justified because even the lightest hydrogen nucleus is almost 2000 times heavier than the electron, and \( M_A \) may be considered so large that the electrons move in a field of fixed nuclear charges. In this case, a separate electronic Hamiltonian is obtained by neglecting the second term of equation 1.2 and taking the last term as a constant nuclear repulsion, \( V_{\text{nn}}(R) \):

\[
\hat{H}_{\text{elec}} = -\sum_{i=1}^{N} \frac{1}{2} \nabla_i^2 - \sum_{i=1}^{N} \sum_{A=1}^{M} \frac{Z_A}{r_{iA}} + \sum_{i=1}^{N} \sum_{j>i}^{N} \frac{1}{r_{ij}} + V_{\text{nn}}(R)
\]

(1.3)

This electronic Hamiltonian depends explicitly on electronic coordinates \( r \) and parametrically on nuclear coordinates \( R \). Solutions of the electronic Schrödinger equation,

\[
\hat{H}_{\text{elec}} \Psi_{\text{elec}}(r; R) = V(R) \Psi_{\text{elec}}(r; R)
\]

(1.4)

therefore yield an electronic wave function which also depends parametrically on the nuclear coordinates. The energy eigenvalue \( V(R) \) contains the total electronic energy of a molecule with nuclei fixed at the coordinate \( R \). Electronic structure methods are concerned with finding approximate solutions to this equation. In this work, we employ Hartree–Fock theory to obtain approximate reference wave functions and energies, which are then systematically improved by applying \textit{ab initio} electron correlation methods like many-body perturbation (MP) or coupled-cluster (CC) theory.29
Once \( V(\mathbf{R}) \) is known, nuclear motion in the Born-Oppenheimer approximation can finally be addressed by revisiting equation 1.1 with the factorized wave function and integrating out electronic degrees of freedom. This yields a Schrödinger equation for nuclear motion with the Hamiltonian

\[
\hat{H}_{\text{nuc}} = -\sum_{A=1}^{M} \frac{1}{2M_A} \nabla_A^2 + V(\mathbf{R}) \quad (1.5)
\]

In this context, \( V(\mathbf{R}) \) is often referred to as the potential energy surface (PES), because it contains the entire potential for nuclear motion in terms of the nuclear coordinates alone. Taylor expansion of \( V(\mathbf{R}) \) about a local minimum enables an exact second-order solution of the nuclear eigenvalue problem known as the harmonic approximation. Alternatively, Newtonian simulations that neglect quantization of nuclear motion can be formulated, where \(-\nabla_\mathbf{R} V\) contains the forces experienced by each atom. Both approaches are taken in the following chapters.

1.3 The many-body expansion

1.3.1 Total electronic energy

For a molecular system in which the electrons and nuclei can be partitioned into \( M \) distinct monomers, the total electronic energy \( E_{\text{total}} \) can be written in a many-body expansion (MBE):

\[
E^{(1)} = \sum_{i} E_i \\
E^{(2)} = \sum_{i<j} E_{ij} - E_i - E_j \\
E^{(3)} = \sum_{i<j<k} E_{ijk} - E_{ij} - E_{ik} - E_{jk} + E_i + E_j + E_k \\
\vdots \\
E_{\text{total}} = \sum_{n} E^{(n)} \quad (1.6)
\]

where \( E_i \) denotes the electronic energy \( V(\mathbf{R}_i) \) of monomer \( i \) with nuclear coordinates \( \mathbf{R}_i \), obtained by solving equation 1.4 with the electronic Hamiltonian of monomer \( i \) in free space. \( E_{ij} \) is the electronic energy of the dimer containing both fragments \( i \) and \( j \), and so on. In solids or liquids described with periodic boundary conditions, the total energy per unit cell, \( E_{\text{cell}} \), can be expressed similarly by restricting all \( i \) summations to one unit cell, introducing distance-based truncations for \( E^{(n>1)} \), and including appropriate factors to avoid overcounting.

In the last line of the MBE, the summation of all \( n \)-body terms in the expansion up to the \( M \)-body term yields nothing but an identity for \( E_{\text{total}} \). The MBE is only
useful for treating large systems when the series quickly converges to $E^{\text{total}}$ at small $n$. This is often the case for insulating molecular solids and liquids, where the electron density can be partitioned into discrete, weakly interacting regions associated with each monomer. In water, for instance, it has been convincingly shown\textsuperscript{31} that the series converges rapidly by $E^{(4)}$, with five- and higher-body interactions having a negligible contribution to $E^{\text{total}}$.

The electronic Hamiltonian contains at most two-body operators, so why can three- and four-body terms be so substantial in the MBE? To a large extent, they contain cooperative effects like the self-consistent polarization of monomer electron densities with their environment. This is drastic in water,\textsuperscript{32} where the dipole moment increases 40% by moving from the vapor to the solid phase, for example.\textsuperscript{19} In effect, each hydrogen bond between two molecules is strongly influenced by the presence of third and fourth neighbors that polarize the interacting pair.

### 1.3.2 Electrostatic embedding

The insight that electrostatic induction accounts for the majority of $E^{(3)}$ and higher-order terms led to numerous methods for accelerating the MBE convergence. These methods approximately fold higher-order effects into the low-order terms using a so-called electrostatic embedding field\textsuperscript{5,19,30,33}

With an embedding field, the form of equation 1.6 is unaltered, but each fragment energy is evaluated with additional one-electron operators in the molecular Hamiltonian for interaction with an external field. This field can be represented in a number of ways, including molecular multipole moments\textsuperscript{34} or atomic charges fit to the monomer electrostatic potentials.\textsuperscript{35} Regardless of this choice, the field must be determined self-consistently to capture the induction effect. It has been shown diagrammatically that the double-counted electrostatic interactions in $E^{(1)}$ are cancelled at every higher order of the embedded MBE.\textsuperscript{20}

The electrostatically-embedded MBE truncated at second order forms the basis for many of the so-called embedded–fragment methods described in the literature.\textsuperscript{20} Conventional MP and CC electron correlation theories have computational costs scaling at least as the fifth power of system size, and so their application to large molecular clusters, solids, and liquids is usually intractable. The embedded-fragment methods, on the other hand, scale no worse than $M^2$ (closer to $M^1$ in large systems where widely separated dimers are omitted from $E^{(2)}$), are almost embarrassingly parallel, and allow accurate electron-correlation theories to be applied to these systems. The binary interaction method described in the following section is one example which has been specifically developed with simulation of molecular solids and liquids in mind.
1.4 The binary interaction method

1.4.1 Energy

In the binary interaction method (BIM), the electronic energy per unit cell is

\[ E_{\text{cell}} = \sum_i E_i \text{Q}_i + \sum_{i \leq j} (E_{ij; \text{Q}_i \cup \text{Q}_j} - E_i \text{Q}_i \cup \text{Q}_j - E_j \text{Q}_i \cup \text{Q}_j) \]

\[ + E_C + E_{LR} \] (1.7)

The summations run over all monomers \( i \) in the primary unit cell and monomers \( j \) in all unit cells. The primed summation indicates that terms where \( j = i \) should only be included if the unit cell indices \((a, b, c)\) of \( j \) are greater than the primary cell index \((0, 0, 0)\) by lexicographic comparison. This restricted summation leverages translational symmetry to avoid repeating costly equivalent dimer calculations. \( E_i \text{Q}_i \) is the electronic energy of the \( i \)-th monomer embedded in the field of point charges \( Q_i \). The embedding field \( Q_i \) is constructed by placing point charges on all atoms of the molecules whose centers of mass fall within a distance \( R_{BQ} \) of monomer \( i \)'s center of mass. 

\( E_i \text{Q}_i \) is obtained by solving equation 1.4 with the electronic Hamiltonian \( \hat{H}_{\text{elec},i} \) from 1.3, plus two terms for the electronic and nuclear interactions of monomer \( i \) with field \( Q_i \):

\[ \hat{H}_{i; \text{Q}_i} = \hat{H}_{\text{elec},i} - \sum_n \sum_{A \in \text{Q}_i} \frac{q_A}{R_{nA}} + \sum_{A_i} \sum_{B \in \text{Q}_i} \frac{Z_A q_B}{R_{AB}} \] (1.8)

where \( n \) indexes the electrons of monomer \( i \), while \( A \) and \( B \) index atomic centers. \( Z_A \) is the charge of nucleus \( A \) and \( q_B \) is the partial charge on embedding atom \( B \). \( E_{ij; \text{Q}_i \cup \text{Q}_j} \) is the total energy of the dimer formed by joining monomers \( i \) and \( j \) in the union of their two embedding fields. \( E_i \text{Q}_i \cup \text{Q}_j \cup \text{Q}_j \) is the total energy of the \( i \)-th monomer in the union of the embedding fields of the \( i \)-th and \( j \)-th monomers, plus point charges on the atoms of the \( j \)-th monomer. Dimers \( ij \) are only computed if the centers of mass of \( i \) and \( j \) lie less than a distance of \( R_{QM} \) apart. See Figure 1.1 for a cartoon representation of the three types of fragment calculations in this formalism.

The first term in the right-hand side of equation 1.7 double-counts the Coulomb interactions between monomers \( i \) and \( j \) via the embedding field when \( R_{ij} \leq R_{BQ} \). This spurious interaction is cancelled in the second term only if \( R_{ij} \) is also less than \( R_{QM} \). Therefore, dimers \( ij \) with center of mass separation residing in the shell \( R_{QM} < R_{ij} \leq R_{BQ} \) incur a double-counting error which is approximately removed by the third term \( E_C \):

\[ E_C = -\frac{1}{2} \sum_{i,j} \sum_{A \in i} \sum_{B \in j} \frac{q_A q_B}{R_{AB}}, \text{ if } R_{QM} < R_{ij} \leq R_{BQ} \] (1.9)
where $A$ and $B$ run over atoms in the $i$th and $j$th monomers, respectively. $E_C$ can be made to vanish simply by setting $R_{QM} = R_{BQ}$. Finally, $E_{LR}$ contains the long-range Coulomb interactions truncated at distance $R_{LR}$:

$$E_{LR} = \frac{1}{2} \sum_{i,j} \sum_{A \in i} \sum_{B \in j} \frac{q_A q_B}{R_{AB}}, \text{ if } R_{BQ} < R_{ij} \leq R_{LR}. \quad (1.10)$$

The BIM described here improves on older methods that require explicit construction of a supercell for the embedding field, which introduces surface effects somewhat arbitrarily and breaks the translational symmetry. This scheme is completely specified by the three truncation radii $R_{QM}$, $R_{BQ}$, and $R_{LR}$ and restores translational symmetry to the fragment energies, in that dimer $i(0)j(n)$ is equivalent to $j(0)i(-n)$, where $n$ denotes the unit cell index relative to primary cell $0$. This enables the cost-saving restricted summation in the second term of equation 1.7. Computationally, the embedding fields $Q_i$ and long range interaction $E_{LR}$ are easily generated, even for nonorthogonal unit cells, using a breadth first search-type algorithm. In short, the procedure accumulates pairs between the primary unit cell and all others, moving outward from cell $(0,0,0)$ and terminating at cells that do not contain any monomers $j$ satisfying $R_{ij} \leq R_{LR}$. The fragment calculations are performed in a load-balanced master-worker scheme implemented with the message passing interface (MPI). This form of distributed-memory parallelism allows fragment calculations entering equation 1.7 to be evaluated concurrently and with minimal communication overhead.

The embedding charges constituting $Q_i$ are determined in a self-consistent cycle of
monomer calculations at the Hartree-Fock level. In the first iteration, each monomer \(i\) wave function is computed in a vacuum. The atomic charges best representing the electrostatic potential due to \(i\) are generated by least-squares fitting the atomic charge potential on a grid according to the CHELPG procedure. In the next iteration, the monomer \(i\) calculation is repeated in the field \(Q_i\) determined from the previous iteration. This procedure converges rapidly and only requires Hartree-Fock calculations of the monomers. All subsequent fragment calculations needed to determine \(E_{\text{cell}}\) can be performed in parallel.

### 1.4.2 Energy gradient and Hessian

The gradient of the BIM energy with respect to atomic coordinates is easily derived from the energy summation formula. Specifically, the derivative of the unit cell energy with respect to the in-phase \(x\)-coordinate of the \(k\)th atom, \(x_k\), is determined by

\[
\frac{\partial E_{\text{cell}}}{\partial x_k} = \sum_i \frac{\partial E_{i;Q_i}}{\partial x_k} + \sum_{i<j} \left( \frac{\partial E_{i;Q_i\cup Q_j}}{\partial x_k} - \frac{\partial E_{i;Q_i\cup Q_j;\cup j}}{\partial x_k} - \frac{\partial E_{j;Q_i\cup Q_j;\cup i}}{\partial x_k} \right) + \frac{\partial E_C}{\partial x_k} + \frac{\partial E_{\text{LR}}}{\partial x_k},
\]

where \(\partial E_{i;Q_i}/\partial x_k, \partial E_{i;Q_i\cup Q_j}/\partial x_k,\) etc. are derivatives of the monomer and dimer energies, taking into account variations of both fragment nuclear coordinates and embedding field atomic coordinates. Thus the derivative \(\partial E_{i;Q_i}/\partial x_k\) is nonzero even when \(k \notin i\), because \(k\) may reside in the embedding field \(Q_i\). The gradients with respect to nuclear coordinates are readily obtained from the analytical gradient capabilities of most molecular quantum chemistry software. The gradients with respect to embedding field coordinates are approximated by evaluating the fragment \(i\) electric field at the atomic site positions of \(Q_i\), which is also possible in most molecular software.

The virial stress tensor component \(P_{xy}\) is computed from the gradients as

\[
P_{xy} = -\frac{1}{V} \sum_i \sum_{k \in i} \frac{\partial E_{i;Q_i}}{\partial y_k}
\]

where \(V\) is the unit cell volume and \(r_{i;x}^{\text{com}}\) is the \(x\)-component of the center of mass of fragment \(i\). For molecular dynamics, a kinetic energy term is added to this quantity. It is symmetrized and the pressure is computed as one third of the trace of \(P\).
Finally, the Hessian matrix elements are available as

\[
\frac{\partial^2 E_{\text{cell}}}{\partial x \partial y} = \sum_i \frac{\partial^2 E_{\text{C}i}}{\partial x \partial y} + \sum_{i \leq j} \left( \frac{\partial^2 E_{ij;Q_i \cup Q_j}}{\partial x \partial y} - \frac{\partial^2 E_{i;Q_i} \cup Q_j}{\partial x \partial y} - \frac{\partial^2 E_{j;Q_j} \cup Q_i}{\partial x \partial y} \right)
\]

+ \frac{\partial^2 E_{\text{C}}}{\partial x \partial y} + \frac{\partial^2 E_{\text{LR}}}{\partial x \partial y},
\]

(1.13)

where \( x \) and \( y \) need not be in-phase collective coordinates; they can be individual coordinates of atoms belonging to different unit cells \( 0 \) and \( n \). In this case, the force constants computed for dimer \( i(0) j(n) \) are also accumulated into the \( j(0)i(−n) \) block of the interaction force constants matrix. Once again, the fragment Hessians are evaluated with molecular software, most often using numerical differentiation of the analytic gradients.
2 Ice Ih anomalies

2.1 Introduction

Hexagonal, proton-disordered ice Ih,\textsuperscript{1,36} the predominant solid phase of water on Earth, displays several anomalous thermodynamic properties. It thermally contracts at low temperatures (< 70 K), unlike most solids, which expand upon heating. Substitution of hydrogen by deuterium in ice Ih results in a volume expansion by 0.09 % at 10 K,\textsuperscript{37,38} which is also the opposite behavior from most other solids. Furthermore, this anomalous volume isotope effect (VIE) grows with temperature and persists even in the liquid phase, despite the fact that isotope-dependence of volume is a quantum effect and is thus expected to vanish in the high-temperature (classical) limit.

Both temperature- and isotope-dependence of volume are quantum as well as anharmonic effects. Anharmonicity usually causes the frequencies of lattice vibrations to decrease upon expansion, since longer atom-atom distances tend to experience a softer part of the anharmonic potential. It, in turn, makes more vibrational states thermally accessible to the solid for an entropy gain, energetically favoring expansion. The fact that the opposite is the case with ice Ih means that many of its thermally accessible lattice vibrations have reverse anharmonicity, where frequency increases upon expansion\textsuperscript{39,40} (i.e., their mode Grüneisen parameters are negative). Specifically, the modes displaying such characteristics are the hydrogen-bond bending modes (transverse acoustic phonons)\textsuperscript{13,41} that squeeze hollow hydrogen-bond cages of ice Ih into denser structures.

With pressure loading, one expects the frequencies of these modes to decrease further and ultimately reach zero, causing a structural instability in compressed ice Ih. This is believed to be the process underlying the “mechanical melting”\textsuperscript{42} (to be distinguished from thermodynamic melting) of ice Ih into a noncrystalline, glass-like, denser phase known as high-density amorphous (HDA) ice\textsuperscript{39,42–44} at 1.0 GPa and 77 K. Both the negative thermal expansion and this pressure-induced amorphization therefore share a common lattice-dynamical origin: the reverse anharmonicity of quantized hydrogen-bond bending modes, which strongly modulate the volume of hollow cages in ice Ih. Corroborating this is the observation that other tetrahedrally coordinated hollow crystals such as silica\textsuperscript{45} also exhibit thermal contraction. On the other hand, which modes are responsible for the anomalous VIE is a far more complex issue.

These anomalous behaviors of ice Ih have striking parallels with those of liquid water, underscoring the broader significance of the issue beyond just ice. Liquid water also thermally contracts at lower temperatures (< 4°C) before it expands. Deuterated
liquid water has a greater molar volume than normal liquid water. Compression of ice Ih between 250 and 273 K causes melting to the denser liquid phase. Indeed, the Ih-HDA transition pressure and temperature can be roughly extrapolated from the melting curve of ice Ih, which has a negative slope reflecting the volume contraction of water upon melting. These comparisons suggest that the anomalous behaviors of liquid water, including the negative slope of the melting curve, may have a similar (if not the same) origin as the ice Ih anomalies, the understanding of which sheds light on the former. On this note, it is worth mentioning that the existence of two distinct phases of liquid water, the low- and high-density liquids, has been proposed as a continuation of the low- and high-density amorphous (LDA and HDA) phases, respectively, though this remains to be a highly controversial issue.

The ice anomalies are macroscopic manifestations of the quantum effects of nuclei moving in anharmonic potentials. These potentials are, in turn, created by intramolecular or intermolecular covalent, hydrogen-bond, and dispersion interactions in the proton-disordered or amorphous solid environment. Therefore, quantitative or even just the correct descriptions of these effects are expected to require high-accuracy quantum many-body methods for both electrons and nuclei. Nevertheless, Ramírez et al. were able to reproduce the temperature-dependence of the volume of ice Ih accurately for a wide range of temperatures using the empirical force field, q-TIP4P/F. The nuclear quantum effects were taken into account with either path-integral molecular dynamics (PIMD) or the quasiharmonic approximation (QHA), which were shown to differ from each other to an insignificant extent. Strässle et al. experimentally demonstrated the pressure-induced softening of transverse acoustic phonons in ice Ih and, combining the experimental data with a simple lattice-dynamical model, showed that it was responsible for the negative thermal expansion and pressure-induced amor- phosphization at 2.5 GPa.

Pamuk et al. used density-functional theory (DFT) with and without dispersion correction in the QHA and determined the primary source of the anomalous VIE as reverse anharmonicity of the intramolecular O–H/O–D stretching modes, though they did not offer an explanation of its growth with temperature. Remarkably, on the basis of the calculations, Pamuk et al. predicted the normalcy of the VIE of ice Ih upon 16O/18O substitution, which was borne out by experiment. On the other hand, when Murray and Galli used another dispersion-corrected DFT, the anomalous VIE of H/D substitution was not reproduced.

These prior studies suggest that the negative thermal expansion and pressure-induced amorphization of ice Ih are relatively well understood and robustly reproducible computationally. However, the anomalous VIE is a much subtler effect and its origin is still not fully understood. Its increase with temperature has not been explained, either.

In this chapter, therefore, we present embedded-fragment ab initio second-order many-body perturbation (MP2) calculations of the structures of ice Ih as a function of temperature and pressure, including the quantum effects of anharmonic lattice vibrations in the QHA. We show that they too reliably reproduce the thermal contraction of ice Ih at low temperatures and thermal expansion at higher temperatures, though the
latter is underestimated. Our MP2 calculations confirm that the contraction is caused by reverse anharmonicity (negative mode Grüneisen parameters) of the hydrogen-bond bending modes, whereas the subsequent expansion is due to the hydrogen-bond stretching modes.

Our MP2 calculations show that the net VIE at 0 K is the result of an extremely close competition between volume-expanding and contracting effects involving all fundamental lattice vibrations (not just the O–H/O–D stretching modes) that exist in the zero-point state. Consequently, even MP2 gives contradicting predictions for the VIE, depending on minor calculation details (specifically, the choice of the embedding field). However, the thermal growth of the anomalous VIE is reliably reproduced and primarily ascribed to the isotope-dependent volume-expanding effect of the librational modes.

The calculations also detect near-abrupt softening of the hydrogen-bond bending (acoustic) modes at 2.35–3 GPa, accompanied by difficulty in converging the geometry optimization. This signals a mechanical instability in ice Ih under pressure, leading to the pressure-induced amorphization, which has been observed experimentally. This is further verified by Born–Oppenheimer molecular dynamics (BOMD) using on-the-fly MP2 forces; at 2.5 GPa and 75 K, the structure of ice degrades from crystalline Ih to amorphous with an initial 20% reduction in volume in a short time.

### 2.2 Computational methods

#### 2.2.1 Electronic structures

The internal energy, equilibrium structure, and phonon dispersion and density of states (DOS) of an infinite crystal of ice Ih were determined with the binary-interaction method (BIM) detailed in Chapter 1. It was combined with MP2 with the aug-cc-pVDZ or aug-cc-pVTZ basis set. The energy was calculated also at the coupled-cluster singles and doubles (CCSD) level with the aug-cc-pVDZ or aug-cc-pVTZ basis set. Here, the energy per unit cell of a crystal was obtained as a sum of MP2/CCSD energies of water dimers and monomers embedded in the self-consistently determined electrostatic field of the crystal.

Two types of the embedding field were considered: atomic partial charges obtained by the CHelpG procedure and molecular dipole moments, both at the Hartree–Fock (HF) level with the corresponding basis set. These embedding fields are hereafter referred to as “atomic” and “dipolar.” First and second derivatives of the energy with respect to atomic coordinates and lattice constants were also obtained efficiently as sums of the corresponding derivatives of the fragment energies, which were then utilized to determine the equilibrium crystal structures and phonon dispersion curves and DOS. The electronic-structure backend for the BIM calculations was a version of NWChem modified for calculating embedding-field derivative contributions to analytical gradients.

Orthorhombic unit cells of ice Ih containing 32 or 64 water molecules were generated by a Monte Carlo method. They had quasi-random proton configurations obey-
ing the ice rules with net zero unit-cell dipole moment. Various lattice-sum truncation radii of the BIM calculations in the notation of our previous work\(^\text{25}\) were as follows: \(R_{\text{QM}} = 8\ \text{Å}, R_{\text{BQ}} = 10\ \text{Å},\) and \(R_{\text{LR}} = 200\ \text{Å}\). At a typical density, each monomer interacted quantum-mechanically with its 66 nearest-neighbor monomers, going beyond the third coordination shell, at which the MP2 correlation interaction decayed to a negligible level.\(^\text{21}\) Each of these monomers was embedded in the field of about 128 nearest neighbors. Each dimer was embedded in the union of the fields of the two monomers. Classical electrostatic interactions beyond the embedding field with an additional 1,140,000 monomer-monomer pairs were also included in the internal energy.

The crystal structure was optimized at 0 GPa and 0 K with tolerances for the atomic and lattice gradients being \(6 \times 10^{-4}\) and \(3 \times 10^{-4}\) a.u., respectively. The unconstrained optimization introduces imperceptibly small positional disorder of the oxygen atoms, which is a physical outcome of the proton disorder neglected in most structural refinements.\(^\text{6}\) Furthermore, as discussed below, a more visible loss of symmetry and difficulty converging the geometry optimization occur at higher pressures, owing to the pressure-induced amorphization.

The interaction force-constants were computed in the \(3 \times 3 \times 3\) unit cells. The normal modes and harmonic frequencies of both \(\text{H}_2\text{O}\) and \(\text{D}_2\text{O}\) ices were calculated at \(9 \times 9 \times 9\) evenly-spaced \(k\) points in the reciprocal unit cell. A regularization of acoustic phonon branches at the \(\Gamma\) point was invoked.\(^\text{23}\)

### 2.2.2 Quasiharmonic approximation

The thermal expansion and VIE of \(\text{H}_2\text{O}\) and \(\text{D}_2\text{O}\) ice Ih were determined from their vibrationally-corrected volumes \((V_0)\) as a function of pressure \((P)\) and temperature \((T)\). The vibrationally-corrected volume \(V_0\) is the one that minimizes the Gibbs energy evaluated in the QHA.\(^\text{13, 41, 60–62}\) In atomic units, the Gibbs energy per unit cell is written as

\[
G(P, T) = F(V_0, T) + PV_0, \tag{2.1}
\]

with the Helmholtz energy being

\[
F(V, T) = E(V) + \frac{1}{2K} \sum_n \sum_k \omega_{nk}(V) + \frac{1}{\beta K} \sum_n \sum_k \ln \left( 1 - e^{-\beta \omega_{nk}(V)} \right), \tag{2.2}
\]

where \(E(V)\) is the internal (i.e., electronic) energy per unit cell as a function of volume \((V)\), \(K\) is the number of \(k\) points in the reciprocal unit cell, \(\omega_{nk}(V)\) is the harmonic frequency of the \(n\)th (fundamental) phonon branch with wave vector \(k\), and \(\beta = (k_B T)^{-1}\). The residual entropy due to proton disorder is negligible for our problems and was therefore neglected.
The Helmholtz energy $F(V, T)$ was evaluated at 0 to 250 K with a 1-K interval at 29 evenly-spaced volumes about the equilibrium structure at 0 GPa and 0 K. In the MP2/aug-cc-pVDZ calculations with the atomic-embedding field, these volumes ranged from 27.3 to 32.0 Å$^3$ per molecule, corresponding to electronic (virial) pressures from 1.25 to $-1.4$ GPa, respectively. The force constants were evaluated and the harmonic frequencies were calculated at each volume, wherein no approximate linearization of frequencies$^{13}$ was invoked. The volumes were obtained by first isotropically scaling the lattice constants and atomic coordinates and then relaxing them under a fixed pressure for each volume. This relaxation step caused slight changes in the unit-cell shape, except when an abrupt and dramatic change occurred near the limit of mechanical stability leading to amorphization (see below). The structures near the stability limit were not included in the QHA calculations. The vibrationally-corrected volume ($V_0$) at finite pressure (including both virial and phonon pressures) and temperature was then obtained by solving

$$
\left( \frac{\partial}{\partial V} \right)_{V=V_0} \{F(V, T) + PV\} = 0, \quad (2.3)
$$

using a third-order polynomial fit of $F(V, T)$, $V_0(P, T)$ thus obtained directly provides both thermal expansion and VIE.

The QHA relies on the analytical expression of the partition function of a harmonic insulating crystal, and the effect of anharmonicity, responsible for thermal expansion and VIE, is implicit in the parametric volume-dependence of the harmonic frequencies, $\omega_{nk}(V)$. It is qualitatively superior to Grüneisen’s theory,$^{63}$ which expresses the isochoric thermal expansion coefficient ($\alpha_V$) in terms of the total Grüneisen parameter ($\gamma$), isothermal bulk modulus ($B_0$), and isochoric heat capacity ($C_V$) as

$$
\alpha_V(T) = \frac{C_V(T)\gamma(T)}{V_0B_0}. \quad (2.4)
$$

Here, $\gamma$ is related to the mode Grüneisen parameters by Eq. (2) of Li et al.$^{63}$ The latter are defined by

$$
\gamma_{nk} = -\frac{\partial \ln \omega_{nk}(V)}{\partial \ln V}, \quad (2.5)
$$

which is positive if the frequency increases upon compression and negative if the frequency decreases upon compression, corresponding to the reverse anharmonicity.

Grüneisen’s theory neglects phonon pressure and also assumes constancy of $\gamma_{nk}$ with volume. It therefore cannot account for VIE and may give poorer results for thermal expansion.$^{62,63}$ Nonetheless, the mode Grüneisen parameters defined above remain useful for understanding the origin of these effects semi-quantitatively. Owing to the congestion of phonon dispersion curves for a large unit cell, a tiny fraction of $\gamma_{nk}$ are not calculated correctly by a finite-difference method because of reordering of phonon dispersion branches between two volumes. Conclusions drawn from the analyses using them are unaffected. The QHA calculations are not affected at all by
reordering of phonon dispersion branches.

2.2.3 Born–Oppenheimer molecular dynamics

At a pressure beyond the limit of mechanical stability, geometry optimization of ice Ih was found to become increasingly difficult, signaling the pressure-induced amorphization. To confirm this, a short (320 fs), constant-pressure (2.5 GPa), constant-temperature (75 K) BOMD simulation was carried out starting from the 64-molecule-unit-cell crystalline structure optimized at 2.35 GPa, using on-the-fly MP2/aug-cc-pVDZ forces with the Nosé–Hoover thermostat \(^{64,65}\) and Berendsen pressure coupling algorithm.\(^{66}\) A time step of 1 fs was used. Additional algorithmic details of MP2 BOMD can be found elsewhere.\(^{25}\)

2.3 Results and discussion

2.3.1 Structure and lattice energy

First, we assess the performance of MP2 for the structural prediction of ice Ih.\(^{5,19,21,51,52}\) Table 2.1 compares the structural parameters obtained with MP2 using different basis sets and embedding fields against the experimental data.\(^{1,38}\) Differences between 32- and 64-molecule unit cells are minor, indicating that the 32-molecule unit cell is large enough to represent an average proton-disordered structure. From here on, all the aug-cc-pVDZ calculations are based on the 64-molecule unit cell, while the aug-cc-pVTZ results are obtained with the 32-molecule unit cell.

In all cases, MP2 underestimates the average oxygen-oxygen distance \(\langle R_{OO} \rangle\), lattice constant \(a\), and the volume \(V_e\), underscoring the well-known tendency of the theory to overbind molecular crystals.\(^{23,24,52,53,63,67}\) The dipolar embedding suffers less from the overbinding, displaying a systematic improvement of the calculated quantities with a basis-set extension.

Table 2.2 compares the calculated and observed\(^{68}\) binding energies of ice Ih. Consistent with the structural data, MP2 overbinds, but not significantly more than various approximations of DFT.\(^{13,55}\) Both the equilibrium and vibrationally-corrected binding energies of CCSD/aug-cc-pVTZ with the dipolar-embedding field are in accurate agreement with the observed (within 4%), while the same method using the atomic-embedding field overshoots the latter significantly. We thus conclude that the atomic-embedding field tends to overestimate the water-water interaction strength in ice.

2.3.2 Thermal contraction and expansion

The vibrationally-corrected volumes of H\(_2\)O and D\(_2\)O ice Ih are plotted as a function of temperature in Fig. 2.1. The equilibrium and vibrationally-corrected volumes, VIE, and bulk moduli are compiled in Table 2.3. The figure shows that the MP2/aug-cc-pVDZ calculation with atomic-embedding field correctly reproduces the thermal contraction at low temperatures and thermal expansion at higher temperatures occurring in both
Table 2.1: Structural parameters of ice Ih.

<table>
<thead>
<tr>
<th>Method</th>
<th>(N)^b</th>
<th>(\langle R_{OO}\rangle)^c / Å</th>
<th>(a) / Å</th>
<th>(c/a)</th>
<th>(V_e) / Å³</th>
</tr>
</thead>
<tbody>
<tr>
<td>MP2/aug-cc-pVDZ (atomic)</td>
<td>32</td>
<td>2.65</td>
<td>8.70</td>
<td>1.64</td>
<td>28.80</td>
</tr>
<tr>
<td>MP2/aug-cc-pVDZ (atomic)</td>
<td>64</td>
<td>2.66</td>
<td>8.71</td>
<td>1.63</td>
<td>28.95</td>
</tr>
<tr>
<td>MP2/aug-cc-pVTZ (atomic)</td>
<td>32</td>
<td>2.64</td>
<td>8.65</td>
<td>1.64</td>
<td>28.47</td>
</tr>
<tr>
<td>MP2/aug-cc-pVDZ (dipolar)</td>
<td>32</td>
<td>2.69</td>
<td>8.84</td>
<td>1.60</td>
<td>29.69</td>
</tr>
<tr>
<td>MP2/aug-cc-pVDZ (dipolar)</td>
<td>64</td>
<td>2.69</td>
<td>8.82</td>
<td>1.62</td>
<td>29.70</td>
</tr>
<tr>
<td>MP2/aug-cc-pVTZ (dipolar)</td>
<td>32</td>
<td>2.70</td>
<td>8.87</td>
<td>1.61</td>
<td>30.15</td>
</tr>
<tr>
<td>DFT (PBE)(^{13})</td>
<td>96</td>
<td>...</td>
<td>8.78</td>
<td>1.63</td>
<td>29.92</td>
</tr>
<tr>
<td>DFT (vdW-DF(^{PBE}))(^{13})</td>
<td>96</td>
<td>...</td>
<td>8.88</td>
<td>1.63</td>
<td>30.90</td>
</tr>
<tr>
<td>Experiment(^{1,38})</td>
<td></td>
<td>2.75</td>
<td>8.99</td>
<td>1.63</td>
<td>32.05</td>
</tr>
</tbody>
</table>

\(^a\)The term in parentheses indicates the type of embedding field.
\(^b\)The number of water molecules in a unit cell.
\(^c\)The average nearest-neighbor oxygen-oxygen distance.
\(^d\)The equilibrium (non-vibrationally-corrected) volume per molecule.

Table 2.2: Binding energies of ice Ih.

<table>
<thead>
<tr>
<th>Method</th>
<th>(\Delta E)^b / meV</th>
<th>(\Delta E_{\text{H}_2\text{O}})^c / meV</th>
</tr>
</thead>
<tbody>
<tr>
<td>MP2/aug-cc-pVDZ (atomic)</td>
<td>773</td>
<td>626</td>
</tr>
<tr>
<td>MP2/aug-cc-pVDZ (dipolar)</td>
<td>705</td>
<td>558</td>
</tr>
<tr>
<td>MP2/aug-cc-pVTZ (dipolar)</td>
<td>673</td>
<td>529</td>
</tr>
<tr>
<td>CCSD/aug-cc-pVDZ (atomic)</td>
<td>715</td>
<td>568</td>
</tr>
<tr>
<td>CCSD/aug-cc-pVTZ (dipolar)</td>
<td>653</td>
<td>506</td>
</tr>
<tr>
<td>CCSD/aug-cc-pVTZ (dipolar)</td>
<td>618</td>
<td>474</td>
</tr>
<tr>
<td>DFT (PBE)(^{55})</td>
<td>663</td>
<td>535</td>
</tr>
<tr>
<td>DFT (vdW-DF(^{PBE}))(^{13})</td>
<td>770</td>
<td>...</td>
</tr>
<tr>
<td>DFT (vdW-DF(^2))(^{55})</td>
<td>648</td>
<td>518</td>
</tr>
<tr>
<td>Experiment(^{58})</td>
<td>610</td>
<td>491</td>
</tr>
</tbody>
</table>

\(^a\)See the corresponding footnote of Table 2.1. The CCSD energy calculations were performed at the MP2-optimized geometry with the corresponding basis set and embedding field. The phonon DOS for \(\Delta E_{\text{H}_2\text{O}}\) was also supplied by MP2.
\(^b\)The equilibrium (non-vibrationally-corrected) binding energy per molecule.
\(^c\)The vibrationally-corrected binding energy per molecule of \(\text{H}_2\text{O}\) ice Ih at 0 K.
Figure 2.1: Change in the vibrationally-corrected volumes of H\textsubscript{2}O and D\textsubscript{2}O ice Ih as a function of temperature relative to the vibrationally-corrected volume of H\textsubscript{2}O ice Ih at 0 K. The experimental data are taken from Röttger et al.\textsuperscript{37, 38}

H\textsubscript{2}O and D\textsubscript{2}O ices, though the expansion at higher temperatures is underestimated considerably.

To analyze their origin, we first categorize in Fig. 2.2 the phonons into five groups with different volume-dependence: the hydrogen-bond bending (HB), hydrogen-bond stretching (HS), librational (L), intramolecular HOH/DOD bending (B), and intramolecular O–H/O–D stretching (S) modes. The hydrogen-bond stretching modes increase their frequencies upon compression and thus have positive mode Gruneisen parameters. This is because the water molecules have greater difficulty pushing one another away in a more crowded environment. In contrast, the hydrogen-bond bending modes (transverse acoustic modes) act to alleviate the congestion in a hollow hydrogen-bonded cage of ice Ih, decreasing the frequencies upon compression and thus having negative mode Gruneisen parameters. Since the strengths of an O–H/O–D bond and of its hydrogen bond have inverse correlation, the O–H/O–D stretching frequencies decrease with pressure and have negative mode Gruneisen parameters. The librational modes increase their frequencies upon compression, while the bending-mode frequencies seem nearly stationary with pressure.

The volume-contracting effect of the hydrogen-bond bending modes has been implicated\textsuperscript{39, 41, 71} in the negative thermal expansion. Our analysis confirms this. Figure 2.3 plots the phonon-group contributions to $C_V(T)\gamma(T)$, a factor in the thermal expansion coefficient according to Grüneisen’s theory [Eq. (2.4)]. Although our calculations are based on the more accurate QHA than Grüneisen’s theory, the total $C_V(T)\gamma(T)$ function (black curve) semi-quantitatively reproduces the temperature-dependence of volume in Fig. 2.1. We find that the thermal contraction is entirely due to the hydrogen-bond bending modes, whereas the subsequent thermal expansion is mostly caused by the hydrogen-bond stretching modes and to a much lesser extent by the librational
<table>
<thead>
<tr>
<th>Method</th>
<th>( V_e ) / Å(^3)</th>
<th>( V_{\text{H}_2\text{O}} ) / Å(^3)</th>
<th>( V_{\text{D}_2\text{O}} ) / Å(^3)</th>
<th>VIE(^e)</th>
<th>( B_e ) / GPa</th>
<th>( B_{\text{H}_2\text{O}} ) / GPa</th>
</tr>
</thead>
<tbody>
<tr>
<td>MP2/aug-cc-pVDZ (0 K)</td>
<td>28.95</td>
<td>29.16</td>
<td>29.21</td>
<td>-0.14%</td>
<td>19.4</td>
<td>15.9</td>
</tr>
<tr>
<td>MP2/aug-cc-pVDZ (100 K)</td>
<td>28.95</td>
<td>29.14</td>
<td>29.18</td>
<td>-0.16%</td>
<td>19.4</td>
<td>15.4</td>
</tr>
<tr>
<td>MP2/aug-cc-pVDZ (220 K)</td>
<td>28.95</td>
<td>29.24</td>
<td>29.31</td>
<td>-0.25%</td>
<td>19.4</td>
<td>13.7</td>
</tr>
<tr>
<td>DFT (PBE) (0 K)(^i)</td>
<td>29.91</td>
<td>29.93</td>
<td>30.04</td>
<td>-0.35%</td>
<td>14.7(^h)</td>
<td>14.3(^h)</td>
</tr>
<tr>
<td>DFT (PBE) (200 K)(^i)</td>
<td>30.60</td>
<td>30.00</td>
<td>30.16</td>
<td>-0.53%</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>DFT (vdW-DF(^{PBE})) (0 K)(^i)</td>
<td>30.90</td>
<td>31.16</td>
<td>31.23</td>
<td>-0.23%</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>DFT (vdW-DF2) (0 K)(^i)</td>
<td>33.29</td>
<td>33.94</td>
<td>33.92</td>
<td>+0.06%</td>
<td>13.0</td>
<td>11.6</td>
</tr>
<tr>
<td>Experiment (10 K)(^37,38)</td>
<td>...</td>
<td>32.05</td>
<td>32.08</td>
<td>-0.09%</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>Experiment (100 K)(^37,38)</td>
<td>...</td>
<td>32.05</td>
<td>32.07</td>
<td>-0.08%</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>Experiment (100 K)(^i)</td>
<td>...</td>
<td>32.08</td>
<td>32.10</td>
<td>-0.08%</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>Experiment (220 K)(^37,38)</td>
<td>...</td>
<td>32.37</td>
<td>32.44</td>
<td>-0.21%</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>Experiment (220 K)(^i)</td>
<td>...</td>
<td>32.37</td>
<td>32.43</td>
<td>-0.19%</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>Experiment (257 K)(^69)</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>9.0</td>
</tr>
<tr>
<td>Experiment (145 K)(^70)</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>9.8</td>
</tr>
<tr>
<td>Experiment (0 K)(^69)</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>12.1(^k)</td>
</tr>
</tbody>
</table>

\(^a\) The MP2 calculations use the atomic-embedding field.
\(^b\) The equilibrium (non-vibrationally-corrected) volume per molecule of ice Ih, which is independent of isotope or temperature.
\(^c\) The vibrationally-corrected volume per molecule of \( \text{H}_2\text{O} \) ice Ih at the temperature given in the first column.
\(^d\) The vibrationally-corrected volume per molecule of \( \text{D}_2\text{O} \) ice Ih at the temperature given in the first column.
\(^e\) The volume isotope effect, as defined by \( (V_{\text{H}_2\text{O}}/V_{\text{D}_2\text{O}} - 1) \). A negative value indicates the anomalous expansion of the heavier isotopomer relative to the lighter one.
\(^f\) The equilibrium (non-vibrationally-corrected) bulk modulus of ice Ih, which is independent of isotope or temperature.
\(^g\) The vibrationally-corrected bulk modulus of \( \text{H}_2\text{O} \) ice Ih at the temperature given in the first column.
\(^h\) Taken from Murray and Galli\(^55\).
\(^i\) Ice XI result of Pamuk\ et al.\(^13\).
\(^j\) Ice Ic result of Murray and Galli\(^55\).
\(^k\) Extrapolated value at 0 K.
Figure 2.2: The phonon density of states of H$_2$O ice Ih calculated at virial pressures of −1.40 and 1.25 GPa using MP2/aug-cc-pVDZ with the atomic-embedding field. The phonon groups with positive and negative mode Grüneisen parameters are highlighted by right- and left-going arrows, respectively. ‘HB’ stands for the hydrogen-bond bending modes (0–125 cm$^{-1}$), ‘HS’ for the hydrogen-bond stretching modes (125–500 cm$^{-1}$ in H$_2$O and 125–425 cm$^{-1}$ in D$_2$O), ‘L’ for the librational modes (500–1500 cm$^{-1}$ in H$_2$O and 425–900 cm$^{-1}$ in D$_2$O), ‘B’ for the HOH/DOD bending modes (1500–2000 cm$^{-1}$ in H$_2$O and 900–1300 cm$^{-1}$ in D$_2$O), and ‘S’ for the O–H/O–D stretching modes (> 2000 cm$^{-1}$ in H$_2$O and > 1300 cm$^{-1}$ in D$_2$O).

The hydrogen-bond bending and the hydrogen-bond stretching modes have the opposite temperature-dependence of volume, while their DOS adjoin each other. These two groups are, however, clearly distinguishable by the mode-stretching parameter defined as

$$s(\omega, \omega + \Delta\omega) = \left\{ \frac{\sum_{i,j} [(\mathbf{u}_i^k - \mathbf{u}_j^k) \cdot \mathbf{n}_{i,j}]^2}{\sum_{i,j} |\mathbf{u}_i^k - \mathbf{u}_j^k|^2} \right\}^{1/2}_{\omega, \omega + \Delta\omega},$$

where the average is taken over all modes labeled by $k$ with frequencies in the window of $\omega$ to $\omega + \Delta\omega$. $\mathbf{u}_i^k$ is the displacement of the oxygen atom in the $i$th molecule and the $k$th normal mode, and $\mathbf{n}_{i,j}$ is the unit vector from the $j$th to $i$th molecule. The $i$ and $j$ summation runs over all hydrogen-bonded pairs in the unit cell.

Figure 2.4 shows that the mode-stretching parameter sharply rises from near zero to about 0.5 at 125 cm$^{-1}$, which demarcates the lower-frequency hydrogen-bond bending modes and the higher-frequency hydrogen-bond stretching modes. This demarcation accurately coincides with the sign change in the average mode Grüneisen parameter. In the hydrogen-bond bending region, its values range from −0.5 to −3 with an average of −1.2, whereas in the hydrogen-bond stretching region, its values fall between +1 and +2.
Figure 2.3: The breakdown of $C_{V}(T)\gamma(T)$ (in units of gas constant $R$) into phonon-group contributions of H$_2$O ice Ih calculated by MP2/aug-cc-pVDZ with the atomic-embedding field. See the caption of Fig. 2.2 for the definitions of phonon groups: HB, HS, and L.

![Figure 2.3](image)

Figure 2.4: The mode stretching parameter [Eq. (2.6) with $\Delta \omega = 4\text{ cm}^{-1}$] and $\Gamma$-point mode Grüneisen parameters of H$_2$O ice Ih at 0 GPa calculated using MP2/aug-cc-pVDZ with the atomic-embedding field. The dashed line at 125 cm$^{-1}$ separates the hydrogen-bond bending modes (‘HB’) from the hydrogen-bond stretching modes (‘HS’).

![Figure 2.4](image)
2.3.3 Bulk modulus

The thermal expansion of ice Ih at higher temperatures is underestimated by a factor of 3 (see Fig. 2.1), causing also the temperature of maximum density to be shifted to a somewhat higher value than the observed (ca. 70 K). It is unlikely that the underestimation is caused by higher-order anharmonicity neglected in the QHA, since Ramírez et al.\textsuperscript{11} showed that the QHA results for the temperature-dependence of volume of ice Ih agree accurately with the results of PIMD, which is exact in the limit of an infinite number of beads. Furthermore, the QHA results seem to slightly overestimate (instead of underestimate) the thermal expansion as compared with the PIMD results. Interestingly, the dispersion-corrected DFT calculation in the QHA by Pamuk et al.\textsuperscript{13} also displayed a similar underestimation of thermal expansion.

A more probable source of this discrepancy is the overestimation of the bulk modulus ($B_0$), as evidenced in Table 2.3. MP2 systematically overbinds ice Ih (see Sec. 2.3.1) and molecular crystals in general, and thus underestimates their volumes.\textsuperscript{23,24,52,63,67} The error in the volumes is amplified in $B_0$ because the latter involves the derivative of pressure with respect to volume.\textsuperscript{63} Since this quantity enters Grüneisen’s expression [Eq. (2.4)] of the thermal expansion coefficient ($\alpha_V$) as a denominator (implicitly in the case of the QHA), its overestimation results in the underestimation of $\alpha_V$. In other words, MP2 tends to predict too stiff a molecular crystal, which therefore cannot expand thermally as much as it should. In our previous study\textsuperscript{63} of solid CO$_2$-I, the underestimation of $\alpha_V$ by a factor of nearly two by MP2 was ascribed to the overestimation of the bulk modulus by ca. 100%. This error was subsequently erased by Heit et al.,\textsuperscript{62} who used more accurate volumes calculated by CCSD with noniterative triples [CCSD(T)] in the complete-basis-set (CBS) limit.

Figure 2.5 plots the MP2 equation of state of H$_2$O ice Ih, which illustrates this point. The equilibrium bulk modulus ($B_e$) is proportional to the reciprocal of the slope of the red curve at 0 GPa, which plots the equilibrium (non-vibrationally-corrected) volume. The reciprocal slope is greater than the corresponding reciprocal slope of the experimental curve (black curve). Clearly, this is caused by the overbinding of ice Ih by MP2. The inclusion of anharmonic vibrational corrections at 150 K in the QHA (blue curve) softens the crystal and brings the value of $B_e$ from 19.4 GPa down to $B_{H_2O} = 14.8$ GPa, which is much closer to the experimental value of 9.8 GPa at 145 K,\textsuperscript{70} but is still too large. This is largely responsible for the underestimation of the thermal expansion around that temperature. In other words, this discrepancy is primarily traced to the smallness of the basis set.

2.3.4 Anomalous volume isotope effect

Next, we seek to explain the anomalous VIE and its temperature dependence. Figure 2.1 and Table 2.3 show that the MP2/aug-cc-pVDZ calculation with the atomic-embedding field reproduces the anomalous VIE at 0 K and also its subsequent increase with temperature, both nearly quantitatively. Experimentally,\textsuperscript{37,38} the VIE at 0 K is $-0.09\%$ (the negative value means anomalous behavior), which remains nearly the
same at −0.08% at 100 K, but increases in magnitude to −0.21% at 220 K. Computationally, we find the corresponding values at −0.14% (0 K), −0.16% (100 K), and −0.25% (220 K), which are in more accurate agreement than other previous calculations.

However, as will be discussed in the Appendix, the calculated results change qualitatively once such a minor detail as the choice of the embedding field is altered; the calculations with the dipolar-embedding field do not reproduce the anomalous (negative) VIE. This means that the agreement in the sign of the VIE (but not its temperature-dependence) from the atomic-embedding field is likely accidental. Computationally reproducing it for the right reason remains to be an exceedingly difficult task even for ab initio MP2. This difficulty can be understood by analyzing the origin of the VIE.

At 0 K, the Gibbs energy reduces to

\[
G(P, T = 0 \text{ K}) = E(V) + \frac{1}{2k} \sum_n \sum_k \omega_{nk}(V) + PV. \tag{2.7}
\]

At \( V = V_e \), the enthalpy is at a minimum, and so we can write

\[
E(V) + PV \approx c_1 + c_2(V - V_e)^2, \tag{2.8}
\]

where \( c_1 \) and \( c_2 > 0 \) are some constants. In this vicinity, \( \omega_{nk}(V) \) can be approximated\textsuperscript{13} as

\[
\omega_{nk}(V) \approx \omega_{nk}(V_e) \left(1 - \frac{V - V_e}{V_e} \gamma_{nk}\right), \tag{2.9}
\]
Table 2.4: Contribution from each phonon group to the anharmonic vibrational correction to the volume and VIE of ice Ih at 0 K, calculated as $\langle \omega \gamma \rangle$ by MP2/aug-cc-pVDZ with the atomic-embedding potential.

<table>
<thead>
<tr>
<th>Phonon group$^a$</th>
<th>H$_2$O ice</th>
<th>D$_2$O ice</th>
<th>VIE$^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>H-bond bending (HB)</td>
<td>$-16.8%$</td>
<td>$-15.8%$</td>
<td>$-1.0%$</td>
</tr>
<tr>
<td>H-bond stretch (HS)</td>
<td>$198.1%$</td>
<td>$185.7%$</td>
<td>$+12.4%$</td>
</tr>
<tr>
<td>Libration (L)</td>
<td>$277.2%$</td>
<td>$204.7%$</td>
<td>$+72.5%$</td>
</tr>
<tr>
<td>Bending (B)</td>
<td>$-1.7%$</td>
<td>$-4.3%$</td>
<td>$+2.6%$</td>
</tr>
<tr>
<td>Stretch (S)</td>
<td>$-356.8%$</td>
<td>$-256.8%$</td>
<td>$-100.0%$</td>
</tr>
<tr>
<td><strong>Total</strong>$^c$</td>
<td>$100.0%$</td>
<td>$113.5%$</td>
<td>$-13.5%$</td>
</tr>
</tbody>
</table>

$^a$ See the caption of Fig. 2.2 for the definitions of phonon groups: HB, HS, L, B, and S.

$^b$ Phonon-group contributions to the VIE as measured by the difference in $\langle \omega \gamma \rangle$ between H$_2$O and D$_2$O ices. The negative value means the anomalous behavior.

$^c$ The total correction for the volume of H$_2$O ice Ih is taken as 100%.

where $\gamma_{nk}$ is the mode Gr"uneisen parameter. The vibrationally-corrected volume $V_0$ that minimizes the Gibbs energy then satisfies

$$V_0 - V_e \propto \frac{1}{K} \sum_n \sum_k \omega_{nk}(V) \gamma_{nk} \equiv \langle \omega \gamma \rangle.$$  \hspace{1cm} (2.10)

Table 2.4 lists phonon-group contributions to the anharmonic vibrational correction to the volume of H$_2$O and D$_2$O ices as estimated by $\langle \omega \gamma \rangle$. Its total VIE does not agree with the value from the QHA, but it reproduces its sign and thus the anomalous VIE. This table paints an exceedingly complex picture of this anomaly. No single phonon group determines even the sign of the VIE. This is because this effect exists even at 0 K, at which it is caused by anharmonicity of zero-point vibrational levels involving all fundamental modes. This is in striking contrast to thermal expansion, in which only low-frequency thermally-accessible lattice vibrations play a dominant role and they tend to share the same volume-expanding or contracting effect. Consequently, none of the phonon groups is negligible in VIE. Even the bending modes, whose contribution to $\langle \omega \gamma \rangle$ is the smallest, has a VIE of the same order of magnitude as the total VIE. It is true that ice Ih exhibits this anomaly owing to the large volume-contracting effect of the O–H/O–D stretching modes, as pointed out by Pamuk et al.,$^{13}$ but whether the VIE is negative or merely small positive is determined by close competition of volume-expanding and contracting effects of all phonon groups. It is doubtful if any method short of a nearly exact one, such as CCSD(T) in the CBS limit, can reproduce it for the right reason.

Our MP2 calculation, however, reproduces the thermal increase of the anomalous VIE. This can be unequivocally attributed to the volume-expanding effect of the librational modes, which are the lowest-energy phonons whose frequencies are strongly isotope-dependent (see Fig. 2.3). Unlike the VIE, this property is not the result of cancellation of competing effects and is thus more reliably reproduced by theory.
Figure 2.6: Phase diagram of H$_2$O. Experimental data were taken from Mishima.$^{43}$ The dashed line is the crystallization line and the dotted-dashed curve is an extrapolation of the melting curve.

### 2.3.5 Pressure-induced amorphization

In 1984, Mishima $et$ $al.$$^{42}$ extrapolated the negative-slope melting curve of ice Ih to low temperatures, predicting and then observing the pressure-induced transition of ice Ih to HDA at 1.0 GPa and 77 K. Combining inelastic neutron scattering from ice Ih under pressure and a lattice-dynamical model calculation, Strässle $et$ $al.$$^{39,40}$ argued that the $\Gamma$–$M$ transverse acoustic phonon dispersion branch softens with pressure, ultimately becoming flat zero at ca. 2.5 GPa, at which point a mechanical instability exists, causing the pressure-induced amorphization. These phonons are believed to be the same as what we call the hydrogen-bond bending modes with the volume-contracting effect. Hence, both the negative thermal expansion and pressure-induced amorphization originate from this particular phonon group.

The phase diagram of H$_2$O is drawn in Fig. 2.6. Simulations with empirical potentials$^{72}$ and inelastic neutron scattering measurements$^{40}$ have established two amorphization regimes: mechanical melting at low temperatures and thermal melting at high temperatures. The thermal melting of ice Ih in the range of 160–250 K results in the metastable formation of supercooled water (with subsequent nucleation of other crystalline phases) at transition pressures that extrapolate smoothly from the melting curve of ice Ih.$^{43}$ Below 160 K, the slope of the transition pressure curve changes abruptly, signaling the entry into the regime of pressure-induced amorphization by mechanical melting. The latter (as opposed to thermal melting) occurs at low temperatures because thermal energy is insufficient for ice to reach structures with lower Gibbs energies. In other words, with increasing pressure, the Gibbs energy of HDA becomes lower than that of ice Ih, where the latter is thus metastable and is unable to cross an energy bar-
Figure 2.7: Equilibrium (non-vibrationally-corrected) volume and lattice constant ratio \((c/a)\) of H\(_2\)O ice Ih as a function of virial pressure calculated by MP2/aug-cc-pVDZ with the atomic-embedding field. The experimental value of \(c/a = 1.628\) is superimposed as the blue line.

rrier for thermal melting. With further pressure loading, flattening of transverse acoustic phonons occurs and Born’s stability condition is violated,\(^72\) and ice Ih thus undergoes mechanical melting to HDA.

Our MP2 calculations detect sudden reduction in volume, loss of symmetry\(^73\) and of long-range order in structure, and softening of the acoustic phonons, all in the pressure range of 2.35–3 GPa, strongly suggesting the occurrence of pressure-induced amorphization during the computer simulations.

Figure 2.7 plots the pressure-dependence of the structural parameters of H\(_2\)O ice Ih calculated by MP2/aug-cc-pVDZ with the atomic-embedding field. With increasing pressure, the geometry optimization starting from an initial crystalline structure became more difficult at \(> 2.35\) GPa. By 3 GPa, it did not converge within an allotted number of iterative cycles. The nonconverged geometry at 3 GPa (Fig. 2.8) shows a sudden drop in volume by 15% relative to the ambient-pressure value and also loss of symmetry, as evidenced by an elevated room-mean-square-deviation of 0.04 in the oxygen fractional coordinates from the experiment values (it is nearly constant at 0.01 below 2.5 GPa). The volume reduction of 15% is in good agreement with the corresponding value of 15% at 2.5 GPa predicted by the lattice-dynamical model based on the inelastic neutron scattering data.\(^40\) Our calculated mechanical transition pressure (2.35–3 GPa) is also roughly consistent with the observed value of 1.0 GPa by Mishima \textit{et al.},\(^42\) 1.4 GPa predicted by the \(q\)-TIP4P/F simulation,\(^11,74\) or 2.5 GPa quoted above.\(^40\)

Figure 2.8 draws the crystalline and disordered structures at 0 and 3 GPa, respectively. While the latter is not converged, it shows that molecules permeate through the otherwise hollow spaces of the hydrogen-bonded crystalline cage and the solid becomes denser. It is an initial structural change towards complete amorphization.
Figure 2.8: The structures of H$_2$O ice Ih at 0 and 3 GPa obtained by MP2/aug-cc-pVDZ with the atomic-embedding field. The 0-GPa structure is fully optimized and stable. The 3-GPa structure is a snapshot of a nonconverging geometry optimization starting from an initial crystalline structure.

Superimposed in Fig. 2.6 is the calculated thermodynamic phase boundary (blue line) between Ih and HDA. It was calculated as the temperature and pressure at which the Gibbs energies (minus the vibrational contributions) of the two phases, i.e., $E(V) + PV - TS_{\text{conf.}}$, agree with each other at the MP2/aug-cc-pVDZ level with the atomic-embedding field. Here, $S_{\text{conf.}} = 0$ was assumed for ice Ih and $S_{\text{conf.}} = 2$ J K$^{-1}$ mol$^{-1}$ for HDA, and $E(V)$ and $V$ for HDA corresponded to the snapshot structure in Fig. 2.8. The calculated thermodynamic transition pressure (blue curve) is about 1.6 GPa at 0 K with little temperature-dependence, above which ice Ih is metastable. This transition pressure is a crude estimate, as we neglected vibrational contributions and used a single nonconverged HDA structure whose energy and volume were not ensemble-averaged. Nevertheless, it is consistent with the extrapolated melting curve of ice Ih (dotted-dashed curve) at 0.6 GPa with little temperature-dependence. The calculated mechanical transition pressure is no higher than 2.35 GPa at 0 K, above which ice Ih is no longer metastable but becomes unstable and transforms into HDA. This is also consistent with the experimental transition pressures (dots) in the mechanical melting regime (< 160 K).

To further confirm that the failure of geometry optimization has a physical origin and is not a computational artifact, we performed a BOMD simulation using on-the-fly MP2/aug-cc-pVDZ forces. Figure 2.9 attests to a rapid reduction in volume by nearly 20% once the initial crystalline structure is allowed to relax according to classical equation of motion. The lattice constant ratio ($c/a$) also fluctuates, unequivocally demonstrating that ice Ih is unstable at 2.5 GPa. With the Berendsen pressure coupling algorithm employed here, the relaxation time of unit cell volume does not have much physical significance, but the dynamics and ensemble averages are meaningful.
Concomitant with the detection of the structural instability, our MP2 calculations also observe pressure-induced red-shifting of DOS due to acoustic phonon branches at \( \text{ca.} \ 2 \ \text{GPa} \), as shown in Figure 2.10. A clearer piece of evidence is given in Fig. 2.11, where rapid softening of the frequency of a representative acoustic mode near the edge of the Brillouin zone is recorded. The softening becomes complete at 2.35 GPa, coinciding with the onset of the geometry optimization difficulty. This pressure is an upper bound for the mechanical instability predicted by MP2/aug-cc-pVDZ. The lag in the transition pressure (2.35–3 GPa) suggested by the geometry optimization procedure is likely due to an insensitivity of the structure to a small number of zero or imaginary frequencies. It is, however, interesting to note that, experimentally also,\(^\text{32}\) it takes an additional 0.5 GPa of pressure loading for the amorphization to complete after it starts at 1.0 GPa.
Figure 2.10: Virial pressure dependence of the phonon DOS of H$_2$O ice Ih calculated by MP2/aug-cc-pVDZ with atomic-embedding field.

Figure 2.11: The frequency of a representative acoustic mode at $\mathbf{k} = (\pi/a, \pi/b, 3\pi/4c)$ as a function of virial pressure calculated by MP2/aug-cc-pVDZ with the atomic-embedding field.
2.4 Embedding-field and basis-set dependence

Here, we examine how robust the qualitative or quantitative predictions about the ice Ih anomalies are, upon changing embedding fields and basis sets.

Figures 2.12 and 2.13 show the temperature-dependence of the ice Ih volume calculated by MP2 with aug-cc-pVDZ and aug-cc-pVTZ, respectively, and the dipolar-embedding field. In both calculations as well as ones shown in Fig. 2.1, the thermal contraction at low temperatures followed by the thermal expansion at higher temperatures is reproduced (though the latter is underestimated), indicating that this qualitative behavior is neither subtle nor difficult to describe computationally. A similar analysis (not shown) based on mode Grüneisen parameters as above confirms that the thermal contraction is caused almost exclusively by the hydrogen-bond bending modes (transverse acoustic phonons), while the thermal expansion is caused by the hydrogen-bond stretching modes. The underestimation of the thermal expansion is to a great extent caused by the overestimation of bulk modulus. Here, we do not observe a systematic improvement upon a basis-set extension, suggesting that approximations in the electronic structure theory, fragmentation scheme, and/or embedding field are as significant as the basis-set incompleteness.

As already mentioned in Sec. 2.3.4, MP2 with the dipolar-embedding field does not reproduce the anomalous VIE; the predicted sign of the VIE is wrong. This is traced to the fact that the sign is determined by an extremely close competition between the volume-expanding and contracting effects of all fundamental modes even at 0 K (and of their overtones and combinations at higher temperatures). Even ab initio MP2 with the aug-cc-pVTZ basis set does not seem to have the sufficient fidelity to reproduce this effect qualitatively correctly, although the basis-set extension seems to bring the calculated VIE closer to the experiment. Note that dispersion-corrected DFT calculations by Pamuk et al.\textsuperscript{13} and by Murray and Galli\textsuperscript{55} gave varied results for this quantity as well. On the other hand, the greater rate of thermal expansion in D\textsubscript{2}O ice Ih, causing the temperature growth of the VIE (provided its anomalous value at 0 K is correctly reproduced), is consistently reproduced by all three calculations, as this is a relatively straightforward consequence of the isotope-dependence of the volume-expanding librational modes.

One may at this point wonder if the atomic-embedding field is simply more accurate than the dipolar-embedding field and, therefore, the former is predicting the anomalous VIE correctly for the right reason (not accidentally). Unfortunately, this does not seem to be the case. First, as we have seen in Sec. 2.3.1, MP2 with the atomic-embedding field overbinds ice Ih to a greater extent than MP2 with the dipolar-embedding field, the latter yielding more accurate structures and binding energies. Figure 2.14 testifies the tendency of the atomic-embedding field to overbind the water clusters to an increasing degree with size. This is, in turn, traced to the overestimation of the dipole moment in the atomic-embedding field, as illustrated in Fig. 2.15. The atomic partial charges that make up the atomic-embedding field are determined so as to reproduce a short-range electrostatic field accurately\textsuperscript{56} and, therefore, they do not necessarily have accurate
Figure 2.12: Same as Fig. 2.1 but using MP2/aug-cc-pVDZ with the dipolar-embedding field.

Figure 2.13: Same as Fig. 2.1 but using MP2/aug-cc-pVTZ with the dipolar-embedding field.
Figure 2.14: The error in the energy of (H₂O)ₙ per molecule incurred by the BIM at the HF/aug-cc-pVDZ or aug-cc-pVTZ level using the atomic- or dipolar-embedding field. The cluster geometries were taken from Maheshwary et al. The negative values indicate overbinding.

Figure 2.15: The distribution of the molecular dipole moments in H₂O ice Ih calculated by MP2/aug-cc-pVDZ with the atomic- or dipolar-embedding field. The unit cell contained 64 molecules and the histogram was convoluted with a Gaussian with the standard deviation of 0.01 Debye. The geometries used were optimized with their respective embedding fields.

dipole moments or the correct long-range electrostatic field.

Since MP2 with the dipolar-embedding field predicts weaker binding in ice Ih, its acoustic modes have lower, near-zero frequencies, which are, therefore, subjected to a greater degree of numerical errors such as those arising from the truncation of long-range interactions and the finite precision of optimized geometries. The volume-temperature curves in Figs. 2.12 and 2.13 should be understood to contain considerable numerical uncertainty and are used here only to demonstrate the sensitivity of the calculated VIE on the embedding fields and basis sets. The results of the atomic-embedding field discussed in the main text should be converged and reliable.
2.5 Conclusions

We have applied BIM-MP2 to anomalous thermodynamic properties of proton-disordered ice Ih. In combination with the QHA, we have computed the Gibbs energies of ice Ih across a range of pressures and temperatures. The calculations have taken into account the quantum effect of nuclei moving in the anharmonic potentials created by the covalent, hydrogen-bond, and dispersion interactions within and between the water molecules, both from first principles without any empiricism or nonsystematic approximations to the electronic or vibrational Hamiltonian. This has enabled a prediction of the vibrationally-corrected volume of ice Ih as a function of temperature and isotope, reproducing quantitatively the thermal contraction at low temperatures and qualitatively the thermal expansion at higher temperatures. An analysis using the mode Gr"uneisen parameters has confirmed that the negative thermal expansion is almost exclusively ascribed to the hydrogen-bond bending modes (transverse acoustic phonons), while the subsequent expansion is mostly due to the hydrogen-bond stretching modes and to a lesser extent to the librational modes. The underestimation of the rate of thermal expansion at higher temperatures is traced largely to a general overestimation of the bulk modulus by MP2, although the same difficulty has been encountered in dispersion-corrected DFT.\textsuperscript{13}

The anomalous VIE is exceedingly difficult to predict correctly, owing to its small magnitude determined by subtle cancellation among many opposing phonon-group contributions. The atomic-embedding field succeeds in reproducing its correct sign, but the dipolar-embedding field does not, highlighting the sensitivity of the result to minor calculation details. This effect was ascribed largely to the negative mode Gr"uneisen parameters of the O–H/O–D stretching modes, which indeed have the largest contribution at 0 K,\textsuperscript{13} but our detailed analysis has shown that there is another phonon group (libration) with a contribution of nearly equal magnitude, and even the smallest phonon-group (bending) contribution is on the same order of magnitude as the final sum. It is, therefore, rather doubtful if any method short of a nearly exact one such as CCSD(T) in the CBS limit has the ability to reproduce this effect for the right reason. The temperature-dependence of the VIE, on the other hand, has been reproduced by our MP2 calculations and is ascribed to the librational modes. This conclusion is more robust and largely independent of calculation details, as this property does not rely on any cancellations.

MP2 has experienced difficulty optimizing the geometry of ice Ih at high pressures (> 2.35 GPa) and a loss of symmetry and a volume collapse was observed at 3 GPa. These are highly suggestive of the mechanical instability of ice Ih, leading to the pressure-induced amorphization to HDA detected experimentally. The MP2 BOMD simulation at 2.5 GPa, starting with an initial crystalline geometry, has also shown a loss of symmetry and long-range order and a 20% volume reduction in a short time. This computational detection of amorphization has been further corroborated by the red-shifting of acoustic phonons accelerating around 2.0 GPa, consistent with the experimental observation of the same by Str"assle \textit{et al.}\textsuperscript{39} Our crude estimate of the
thermodynamic Ih-HDA phase boundary is 1.6 GPa at 0 K with small temperature-dependence, which is reasonably close to, if larger than, the experimentally inferred value of \( \approx 0.6 \) GPa. The onset of the mechanical instability is higher both in experiment (1.0–2.5 GPa) and our calculation (< 2.35 GPa).
3 Water ices beyond ambient pressure

3.1 Introduction

The binary interaction method (BIM) has been used to compute ab initio electron-correlated structures, vibrational spectra, and thermochemical properties for a variety of molecular crystals using second-order many-body perturbation (MP2) theory. The predictive capabilities of BIM-MP2 have been established for discerning the relative stabilities and spectroscopic signatures of solid carbon dioxide and formic acid polymorphs. For solid water, however, only the ambient pressure phase ice Ih and high-pressure phase VIII have been investigated with BIM. The studies focused either on vibrational (infrared, Raman, and inelastic neutron scattering) spectroscopy or thermodynamic attributes like thermal expansion, which were discussed in Chapter 2.

However, ice Ih is only one of at least fifteen known phases of solid water, which exhibit a diversity of crystal structures with densities spanning 0.9 to 1.4 g cm$^{-3}$ (excluding the non-molecular phases) and hydrogen configurations ranging from highly ordered ferroelectric systems to completely disordered. The solid water phase diagram has been completely mapped out and high-quality structural, spectroscopic, and thermodynamic data are published for many of these phases. In general, predicting the phase diagram of molecular solids without a priori information is one of the most exciting prospects of ab initio embedded-fragment methods. It is a major challenge that requires identifying global minima in the potential energy surface (PES) and accurately evaluating the free energy to discern relative stabilities as small as 1 kJ mol$^{-1}$.

Studying ice allows us to focus on the latter problem of discriminating free energies, since fifteen polymorphs are already well-characterized.

Ramírez applied the quasi-harmonic approximation (QHA) to evaluate the empirical q-TIP4P/F phase diagram of the ice phases. That model predicted too great a stability of ice IX relative to ices II and V, which were therefore metastable in the entire phase diagram. With density functional theory (DFT), exchange-correlation functionals lacking dispersion corrections overestimate transition pressures by at least an order of magnitude. Santra et al. found that several dispersion-corrected functionals, on the other hand, predict unphysically large stability of ice IX at the expense of thermodynamic stability of ice Ih and II. In general, dispersion interactions become more significant with increasing density of the ice phases, but the lattice constants and energies predicted by DFT are highly sensitive to the choice of dispersion-corrected functional. According to a 2016 perspective, although significant progress has been
made in DFT descriptions of water, “current dispersion-inclusive methods are still immature, and over-correction can be as much of a peril as under-correction.”\textsuperscript{17}

In this Chapter, therefore, our goal is to evaluate the performance of BIM-MP2 in the simultaneous prediction of thermodynamic properties for several ice phases. We employ the QHA, which is a particularly effective method to compute solid vibrational free energies as analytic functions of temperature and volume, using only volume-dependent normal mode frequencies. In this manner, we can compute the Gibbs free energy of the ices with a balanced, nonempirical treatment of the electron-correlated internal energy and quantized anharmonic lattice vibrations. We study the ice phases Ih, II, III, V, and VI, which occupy a contiguous region of the experimental phase diagram at ambient and medium (<3 GPa) pressures. Additionally, we simulate ice IX, which is the proton-ordered form of ice III. We first compute the ice structures and phonon dispersion using MP2 in the aug-cc-pVDZ basis set. The free energies are subsequently evaluated using electronic energies from both MP2/aug-cc-pVDZ and MP2 extrapolated to the complete basis set (CBS) limit. We make comparisons to experimental data on the inelastic neutron scattering (INS) spectra, $PV$ equations of state, thermal expansion, and finally the phase diagram.

### 3.2 Computational methods

#### 3.2.1 Embedded-fragment crystal optimization and CBS energy

The crystal energy per unit cell, $E_{\text{cell}}$ and energy gradient with respect to atomic coordinates were computed with the binary interaction method (BIM) using density-fitted second-order Möller-Plesset perturbation theory (DF-MP2) in the aug-cc-pVDZ basis set. The BIM lattice truncation radii were chosen to be $R_{QM} = 8 \, \text{Å}$, $R_{BQ} = 8 \, \text{Å}$, and $R_{LR} = 200 \, \text{Å}$, and the embedding field of self-consistent atomic charges fit to reproduce each monomer electrostatic potential was used. For more information, refer to the formalisms and algorithmic details of BIM in Chapter 1.\textsuperscript{25,26}

For each ice phase, an optimization of both atomic coordinates and lattice parameters was carried out at ambient pressure. The gradients were relaxed to $1 \times 10^{-3}$ a.u. for atomic positions and $3 \times 10^{-4}$ a.u. for lattice parameters. All cells had a net-zero dipole moment; for proton-disordered phases, a random hydrogen bond configuration obeying the ice rules with vanishing dipole was constructed. The optimal cell geometry was then scaled isotropically to generate a grid of 11 to 19 volumes. The atomic coordinates were relaxed at each volume while keeping the lattice parameters fixed. The MP2 energy in the complete basis set (MP2/CBS) was calculated at each volume from two calculations in the aug-cc-pVTZ and aug-cc-pVQZ basis sets. The CBS limit was computed from the HF aug-cc-pVQZ energy and the standard Helgaker two-point extrapolation of the correlation energy.\textsuperscript{78} The Psi4 program package\textsuperscript{79} was used as a backend for the fragment calculations.
3.2.2 Phonon dispersion

Hessians were computed by finite difference of BIM MP2/aug-cc-pVDZ gradients at the three lowest energy volumes for each phase. The interaction force constants were determined in 3x3x3 unit cells for ices I, III, V, and VI, while 5x5x5 cells were used for the proton-ordered ices II and IX. The phonon dispersion $\omega_{nk}$ was evaluated in the first Brillouin zone using as few as 125 $k$-points in the large ice VI cell and as many as 729 $k$-points for the small proton-ordered cells of II and IX. The inelastic neutron scattering intensity was computed as the hydrogen amplitude-weighted phonon density of states, histogrammed into 1 cm$^{-1}$ bins and convoluted with gaussians of 3 cm$^{-1}$ standard deviation. The low energy pseudotranslational normal modes were classified according to Tanaka’s mode stretching parameter:

$$ s(\omega, \omega + \Delta \omega) = \left( \frac{\sum_{i,j} |(u^i_k - u^j_k) \cdot n_{i,j}|^2}{\sum_{i,j} |u^i_k - u^j_k|^2} \right)^{1/2}_{\omega,\omega+\Delta \omega} $$

(3.1)

where the average is taken over all normal modes in the frequency range $\omega$ to $\omega + \Delta \omega$, $u^i_k$ is the displacement of oxygen atom $i$ in mode $k$, and $n_{i,j}$ is the unit vector pointing from oxygen $i$ to oxygen $j$. The summation is only taken over hydrogen-bonded pairs. A value of $s$ near 0 suggests that modes in a given frequency range consist primarily of hydrogen bond (O-O-O) bending-type motion.

3.2.3 Quasiharmonic approximation

The Grüneisen parameters corresponding to modes with frequency $\omega_{nk}$,

$$ \gamma_{nk} = -\frac{\partial \ln \omega_{nk}}{\partial \ln V} $$

(3.2)

were computed by finite difference of the frequencies computed at the three lowest energy volumes of each ice phase. Each frequency $\omega_k$ was subsequently ascribed a linear volume dependence from the first-order Taylor expansion about the volume $V_e$ that minimizes electronic energy:

$$ \omega_k(V) = \omega_k(V_e) \left(1 - \frac{V - V_e}{V_e} \gamma_k \right). $$

(3.3)

This form of the frequencies was then used directly to evaluate the Helmholtz free energy at temperature $T$:

$$ F(V, T) = E(V) + \frac{1}{2K} \sum_n \sum_k \omega_{nk}(V) 
+ \frac{1}{\beta K} \sum_n \sum_k \ln \left[1 - e^{-\beta \omega_{nk}(V)}\right] - TS_{\text{conf}} $$

(3.4)

where $E(V)$ is the electronic energy per unit cell as a function of volume ($V$), $K$ is the number of $k$-points in the the first Brillouin zone, $\omega_{nk}$ is the harmonic frequency of the
nth fundamental branch with wave vector $k$, and $\beta = (k_B T)^{-1}$. The final term contains the contribution of the proton configurational entropy $S_{\text{conf}}$. It vanishes for ordered phases II and IX, and accurate values from recent simulations were used for phases I, III, V, and VI.

The Helmholtz energy is evaluated from 0 to 250 K in 1 K intervals at each of the optimized volumes generated in section 3.2.1. A cubic interpolating polynomial in volume is used to find the volume $V_0$ that minimizes the Gibbs free energy at a given temperature and pressure ($P$):

$$G(T, P) = F(V_0, T) + PV_0$$
$$P = -\frac{\partial F(V, T)}{\partial V} \bigg|_{V=V_0}$$

$V_0(T, P)$ thus contains the equation of state and thermal expansion behavior of each phase. $G(T, P)$ is used to construct the phase diagram in a brute-force approach: at each point in the $P-T$ plane the phase with minimum $G$ is identified as the thermodynamically stable form of ice. The isothermal bulk modulus $B_0$ is then obtained analytically as

$$B_0 = V_0 \left( \frac{\partial^2 F}{\partial V^2} \right)_{V=V_0}.$$  

### 3.3 Results

#### 3.3.1 Structural optimization and lattice energy

Table 3.3.1 lists the structural parameters of ices Ih, II, III, IX, V, and VI after cell optimization by BIM with MP2/aug-cc-pVDZ. It is made clear in Figure 3.1 that MP2 tends to systematically overbind ice crystals and predict volumes that are about 10% smaller than the observed. The lattice energy is significantly improved with basis set extension from aug-cc-pVDZ to the CBS limit, moving 100 meV closer to the experimental value of 610 meV / H$_2$O for ice Ih. The curvature of the lattice energy also decreases with increasing basis set, which results in a lower bulk modulus as discussed.

<table>
<thead>
<tr>
<th>System</th>
<th>$a$/Å</th>
<th>$b$/Å</th>
<th>$c$/Å</th>
<th>$\alpha$</th>
<th>$\beta$</th>
<th>$\gamma$</th>
<th>$N^a$</th>
<th>H-disorder$^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ice Ih</td>
<td>8.63</td>
<td>7.49</td>
<td>14.20</td>
<td>90.0</td>
<td>90.0</td>
<td>90.0</td>
<td>32</td>
<td>Yes</td>
</tr>
<tr>
<td>Ice II</td>
<td>7.41</td>
<td>7.41</td>
<td>7.41</td>
<td>113.1</td>
<td>113.1</td>
<td>113.1</td>
<td>12</td>
<td>No</td>
</tr>
<tr>
<td>Ice III</td>
<td>6.41</td>
<td>6.41</td>
<td>6.53</td>
<td>90.0</td>
<td>90.0</td>
<td>90.0</td>
<td>12</td>
<td>Yes</td>
</tr>
<tr>
<td>Ice V</td>
<td>8.61</td>
<td>7.20</td>
<td>9.77</td>
<td>90.9</td>
<td>109.3</td>
<td>90.0</td>
<td>28</td>
<td>Yes</td>
</tr>
<tr>
<td>Ice VI</td>
<td>11.96</td>
<td>12.00</td>
<td>11.00</td>
<td>90.0</td>
<td>90.0</td>
<td>90.0</td>
<td>80</td>
<td>Yes</td>
</tr>
<tr>
<td>Ice IX</td>
<td>6.58</td>
<td>6.58</td>
<td>6.56</td>
<td>90.0</td>
<td>90.0</td>
<td>90.0</td>
<td>12</td>
<td>No</td>
</tr>
</tbody>
</table>

$^a$ Number of water molecules in the simulation unit cell.

$^b$ If the phase is H-disordered, a random proton arrangement obeying the ice rules with net-zero dipole moment was generated.
Figure 3.1: Lattice energies of the ice phases without vibrational corrections. The small markers denote the MP2 electronic energies calculated at each volume. The lines show the cubic polynomial fit in volume. The large circles show the experimental or DMC lattice energies of ices Ih, II, and IX at the equilibrium volume of each phase. These energies are extrapolated to 0 K and have zero-point vibrational contributions subtracted. The squares and diamonds show the same quantities computed with PBE and PBE+vdW\textsubscript{TS} DFT functionals, respectively.

below. We note that the $E(V)$ profiles are extremely congested for these solid phases of water; at MP2/CBS, the equilibrium energies of the six phases span only 30 meV, while the experimental lattice energies of Ih, II, and IX differ only by 4 meV (0.4 kJ mol\(^{-1}\)). Such a small energy scale poses a great challenge for electronic structure theory predictions of the lattice energy. The MP2/aug-cc-pVDZ energy ordering of II < V < VI < IX < III < Ih is almost backwards, as the high-pressure phases V and VI are more stable than the predominant phase Ih at ambient pressure. This is most likely due to the combination of MP2 theory, which intrinsically overbinds molecular crystals, and the smallness of the aug-cc-pVDZ basis set, which incurs a significant basis set superposition error (BSSE) in embedded-fragment calculations.

The lattice energy ordering of the phases is largely corrected in moving to MP2/CBS, where lower-densities ices I, II, and III/IX are significantly more stable than high density ices V and VI. Comparison to the DFT results shows that the semilocal PBE functional strongly overstabilizes ice Ih relative to II and IX, while the dispersion-corrected PBE+vdW\textsubscript{TS} functional performs more similarly to MP2 but still slightly overbinds ice Ih.

3.3.2 Inelastic neutron scattering intensity

The inelastic neutron scattering intensities are predicted from the hydrogen amplitude-weighted phonon density of states (DOS) generated by BIM-MP2/aug-cc-pVDZ. The spectra are divided into low energy translational and librational regions in Figure 3.2,
Figure 3.2: The INS intensity (arbitrary scale) of H$_2$O ices in the translational and librational regions. The intensity profile calculated by MP2/aug-cc-pVDZ is shown by the black curves, and the experimental spectrum is given by the filled red curves. The MP2 frequency axis is scaled by 0.79 to facilitate comparisons. The DFT(vdW-DF$^{PBE}$) result is from Pamuk.

and higher energy covalent bending and stretching regions in Figure 3.3.

The atomic embedding field and relatively small aug-cc-pVDZ basis cause BIM-MP2 to overestimate hydrogen bond strengths in ice, resulting in a blueshift of the intermolecular vibrational frequencies relative to a conventional calculation with larger basis set. There is also a corresponding redshift in the OH covalent stretching bands due to the well-known anticorrelation of OH OH-O bonding strengths. For this reason, the INS calculations are presented with MP2 translational and librational frequencies scaled by a factor of 0.79 to facilitate assignment of bands and comparisons between phases.

The simulated INS spectrum of ice Ih in Figure 3.2 accurately predicts the intensity ratio between the dominant peak at 60 cm$^{-1}$, assigned to the hydrogen bond bending modes (see section 3.3.3), and the hydrogen bond stretching peak at around 313 cm$^{-1}$.
The heavy right shoulder of the first peak is accurately reproduced as well, though its apparent width is reduced by the 0.79 frequency scale factor. The experimental stretching peaks at about 135 and 220 cm$^{-1}$ are apparent in the simulated spectrum but with lower intensity. The INS spectrum from Pamuk’s vdW-DF$^{PBE}$ calculation is relatively featureless in the hydrogen bond stretching region.

Comparing ice Ih to the proton-ordered phases II and IX, we find significantly more structure in the hydrogen bond stretching modes of the ordered phases. The simulated spectrum of ice IX, for instance, correctly reproduces the pronounced splitting of the three stretching peaks into doublets. On the other hand, dense, proton-disordered phases V and VI have significantly greater diversity in their local hydrogen bond environments and exhibit relatively broad peaks in the 0 - 350 cm$^{-1}$ range. The dominant features in this region are clearly assignable from the simulated INS spectra of phases V and VI.

The librational bands are less nuanced, but some discernable features in the experimental data can be identified in the simulated spectra as well. In ice II, a large drop in INS intensity between the peaks at 500 and 625 cm$^{-1}$ is clearly reproduced. It and the other proton-ordered phase IX exhibit significantly more structure than the disordered phases.

The covalent vibrational manifolds of ice Ih are shown in Figure 3.3. In order of increasing energy, the experimental spectrum has three bands corresponding to HOH bending, a lower-intensity bending-librational combination band, and the high intensity signal from OH stretching. The MP2 frequencies are unscaled for comparisons in this region of the INS spectrum. Pamuk’s DFT calculations show a pronounced splitting in the covalent stretching region, which may be due to their use of a proton-ordered ice Ih structure.\(^\text{13}\)

Interestingly, the observed bending peak is relatively stationary around 1700 cm$^{-1}$ for all phases but Ice Ih, where it is broadened and redshifted to 1600 cm$^{-1}$. The simulated bending INS intensity is centered on 1700 cm$^{-1}$ for all phases, giving good agreement with II, III, IX, V, and VI but not Ih. The vdW-DF$^{PBE}$ bending manifold is slightly redshifted with respect to the MP2 calculation, in closer agreement with the observed.

Moreover, the relative intensity of the bending and stretching manifolds is reproduced at least semi-quantitatively for all phases other than Ih. This is unlikely due to the proton disorder or finite 32-molecule unit cell used for ice Ih, because nearly the entire phonon DOS agrees quantitatively with one generated in a previous study\(^\text{26}\) from a random configuration of a larger 64-molecule cell.

### 3.3.3 Pseudotranslational mode analysis

The negative thermal expansion of ice Ih has been studied extensively\(^\text{26,41}\) and ascribed to the “reverse anharmonicity” of the hydrogen bond bending modes below 125 cm$^{-1}$. Unlike most normal modes, their frequencies decrease as the crystal volume shrinks; equivalently, they have negative Gr"{u}neisen parameters according to the definition in
Figure 3.3: The INS intensity (arbitrary scale) of H$_2$O ices in the water covalent bending and stretching regions. The intensity profile calculated by MP2/aug-cc-pVDZ is shown by the black curves, and the experimental spectrum$^9$ is given by the filled red curves. The DFT(vdW-DF$^{PBE}$) result is from Pamuk.$^{13}$
The $\text{H}_2\text{O}$ mode Grüneisen and stretching parameters computed by MP2/aug-cc-pVDZ are plotted on the left and right vertical axes, respectively, against frequency for each ice phase. Broken lines are due to gaps in the phonon DOS.

The relatively large Boltzmann population of these modes at low temperature skews the free energy surface to favor smaller volumes resulting in the observed volume contraction below 70 K.

All of the pseudo-translational modes may also be characterized by the mode stretching parameter defined in equation 3.1, which is 0 for modes which consist purely of hydrogen bond bending-type motion and 1 for purely hydrogen bond stretching motion. This parameter is plotted for all ice phases in Figure 3.4, together with the mode Grüneisen parameters from 0-350 cm$^{-1}$, at MP2/aug-cc-pVDZ.

For ice Ih, the abrupt shift from negative to positive Grüneisen parameters coincides exactly with a sharp increase of the mode parameter from below 0.1 to about 0.5. This makes clear that the anomalous modes are precisely the bending motions which alleviate crowding in compressed ice Ih by moving molecules into empty interstitial spaces. Does this effect vanish in higher-pressure forms of ice, which have significantly less interstitial space within their structures?

Figure 3.4 shows that for all ices, the mode stretching parameter does indeed increase with frequency from a relatively low value below 150 cm$^{-1}$. For none of the phases II-IX is the transition from bending to stretching character so pronounced as in phase Ih. Moreover, the Grüneisen parameters are most substantially negative for ice Ih, which has a significantly larger molar volume than the other ices. There is a clear trend that lower density phases like II, III, and IX have lower Grüneisen parameters than the high-density phases V and VI. We conclude that the pseudotranslational modes below 150 cm$^{-1}$ mostly retain their character in these six phases of ice, but the reverse anharmonicity is by far most significant for ice Ih. These modes are weakly
anharmonic in the other five phases, and the Grüneisen parameter is only significantly negative for ice IX. This leads to a prediction of very slight negative thermal expansion by the QHA for ice IX (only -0.005% volume contraction).

### 3.3.4 Thermal expansion

Having discussed the peculiar character of the hydrogen bond bending modes and their role in thermal contraction of Ice Ih, we note that the subsequent expansion is dominated by Boltzmann population of the hydrogen bond stretching modes up to about 250 K, while the librational modes remain mostly frozen out. With this in mind, the thermal expansion of the six H$_2$O ices at ambient pressure is presented in Figure 3.5. A significant negative thermal expansion is predicted for ice Ih. For the other phases, the volume shift is not appreciable until about 100 K due to the more harmonic behavior of the hydrogen bond bending modes compared to ice Ih.

When MP2/CBS is used for the electronic energy $E(V)$, the predicted volume shifts are significantly larger than from the aug-cc-pVDZ calculation. This may seem puzzling at first, since thermal expansion is due to volume modulation of vibrational frequencies, and in both calculations this effect is computed from the same aug-cc-pVDZ
force constants. The difference stems from the significantly smaller bulk modulus predicted by the CBS calculation, which is obvious even qualitatively from Figure 3.1 and reproduces experimental values much more accurately (section 3.3.5).

Detailed experimental measurements on the thermal expansivity of ice are only available for phases Ih and II. BIM calculations of the expansion of the D₂O isotopomer at ambient pressure are shown in Figure 3.6 alongside the observed data. By the same token that the lower bulk modulus of MP2/CBS improves agreement with the expansion of ice II, it also amplifies the volume shrinking of ice Ih and thereby worsens agreement with experiment in the temperature range of 50-200 K. The thermal expansivity of ice Ih predicted by the vdW-DF\textsubscript{PBE} DFT method is quite similar to that of MP2.

This nonsystematic improvement in ice Ih can be attributed to two factors: blue shift of the hydrogen bond stretching frequencies, which delays onset of expansion, and inaccurate estimation of the pseudotranslational mode anharmonicities by BIM-MP2 in general. As discussed previously,\textsuperscript{26} the neglect of higher-order anharmonicity is not a likely culprit, because similar calculations using fully anharmonic path integral molecular dynamics with DFT\textsuperscript{13} or empiricial potentials\textsuperscript{11} met the same difficulties in predicting the expansion of ice Ih quantitatively. Nonetheless, the basis set improvement in the prediction of ice II expansion is quite encouraging.

### 3.3.5 Equation of state

PV equations of state are equally accessible from $V_0(T, P)$ obtained in the QHA. Several isotherms are plotted against the available experimental data for the ice phases in Figure 3.7. It is immediately apparent that BIM-MP2 consistently predicts a density
that is too high by 5-10%. A significant improvement with basis set extension is observed for all phases but ice Ih. This occurs for precisely the same reasons discussed in section 3.3.4; after all, the results in this section are just a different slice of the same \( V_0 \) dataset in the \((T, P)\) domain.

A more significant improvement is seen in the isothermal bulk modulus, which is available analytically from the second-derivative of the free energy in cubic polynomial form. It is proportional to the reciprocal of the slope in Figure 3.7, where it is qualitatively apparent that MP2/CBS predicts softer crystals than BIM/aug-cc-pVDZ. This is quantified in Figures 3.8 and 3.9, which show the variation of bulk modulus with temperature at ambient pressure, and with respect to pressure at 238 K, respectively. At first glance, it's clear that MP2/CBS significantly improves the bulk moduli of MP2/aug-cc-pVDZ, bringing them to near-quantitative agreement with experimental measurements.

Figure 3.8 shows that the crystals soften slightly with temperature, while Figure 3.9 demonstrates hardening with increasing pressure for phases Ih, III, V, and VI. Interestingly, the MP2/CBS calculations predict a reduction in bulk modulus with increasing pressure for proton-ordered phases II and IX. This is likely related to the proximity of a mechanical instability, where the free energy surface no longer contains a local minimum. The bulk modulus should ultimately vanish on the spinodal curve where the second derivatives of free energy are zero. Interestingly, we also see foreshadowing of a spinodal at negative pressures for ices Ih, V, and VI at MP2/CBS. For ices V and VI, the vanishing of bulk modulus around -0.6 GPa suggests an instability with respect to one of the less dense phases of ice.

### 3.3.6 Phase diagram

Perhaps one of the most exciting prospects of embedded-fragment electron correlation theories is the first-principles prediction of a chemical’s phase diagram. Given the Gibbs free energy in the QHA for ice phases Ih, II, III, IX, V, and VI, it is natural to ask whether the experimental phase diagram, or at least coexistence lines for pairs of phases, can be approximated by BIM-MP2.

Figure 3.10 revisits the lattice energy predicted by BIM-MP2, this time corrected by vibrational zero-point energy at 0 K and 0 GPa, in comparison with experimental values for the ice phases. The minimum lattice energy of each curve in this figure can be equated with \( G(0 \text{ K}, 0 \text{ GPa}) \) and is therefore telling of the phase diagram in this region. It’s clear that the situation is rather hopeless with MP2/aug-cc-pVDZ, which predicts that the high density phase VI is more stable than ice Ih by about 7 kJ mol\(^{-1}\) even at ambient pressure. Since the predominant phase of water on Earth, ice Ih, does not occupy any region of stability with BIM-MP2/aug-cc-pVDZ, we do not analyze this calculation any further.

The situation looks difficult but more promising with MP2/CBS, which improves significantly on the lattice energies and correctly predicts ice I as being more stable than the high-density phases at ambient pressure. As in section 3.3.1, we note how
Figure 3.7: PV equations of state of H₂O ices Ih, II, III, V, and VI. Isotherms computed by MP2 in the aug-cc-pVDZ basis set and CBS limit are compared against experimental data\textsuperscript{7,70,77} for D₂O at 145 K and 225 K, or H₂O at 238 K. The isotope effect on volume is negligible on the scale of these comparisons.
Figure 3.8: Isothermal bulk modulus of each H$_2$O ice versus temperature at ambient pressure, computed by MP2 in the aug-cc-pVDZ basis set and CBS limit. Experimental data available for Ice Ih are shown as black dots.$^{59,70}$

Figure 3.9: Isothermal bulk modulus of each H$_2$O ice versus pressure at 238 K, computed by MP2 in the aug-cc-pVDZ basis set and CBS limit. The experimental data$^{77}$ are shown as black dots.
Figure 3.10: The vibrationally-zero-point-corrected lattice energy. As in Figure 3.1, the experimental data\textsuperscript{68} are shown by dots and the curves show the cubic interpolation through the MP2 data.

Tightly clustered the lattice energies and volumes of the higher-pressure phases are, the MP2/CBS lattice energies of II, III, IX, V, and VI all being in a window of 1 kJ mol\textsuperscript{-1}.

The phase diagram computed by MP2/CBS is illustrated in Figure 3.11. It was generated as a brute-force colormap, by explicit comparison of the QHA Gibbs energies for each phase and \((T, P)\) area element. From the overlaid experimental phase diagram, it is clear that the observed coexistence lines of the five phases occupy a small fraction of the computed diagram. At lower temperatures, ice Ih gives way to proton-ordered XI and ice VI to proton-ordered XV; these proton-ordered phases were not computed in the present study. The sixth computed phase, ice IX, is the proton-ordered form of III and metastable with respect to II in the real phase diagram.

Not surprisingly, the simulated phase diagram does not predict stability for some phases appearing in the experimental diagram. The predictions are not as disappointing as the diagram may first suggest, however. Ice III only appears at high temperatures near -1.0 GPa, where ice Ih becomes unstable due to its already low density. The slightest overestimation of ice II’s lattice energy relative to ice V dooms it to never appear in the phase diagram, despite the generally satisfactory accuracy of other thermochemical properties of ice II. A shift of -0.5 kJ mol\textsuperscript{-1} in its lattice energy relative to ice V may indeed change the diagram dramatically; it is conceivable that simply larger lattice truncation radii in BIM could achieve this. However, ice IX, which is metastable in the region of ice II stability, does appear between ices Ih and V at low temperature. Therefore, a slice of the phase diagram parallel to the pressure axis below about 40 K at least correctly predicts the order of transitions \(I \leftrightarrow II/III/IX \leftrightarrow V \leftrightarrow VI\). Moving horizontally parallel to the temperature axis, the \(IX \leftrightarrow III\) and \(IX \leftrightarrow V\) transitions correctly reflect the
Figure 3.11: The phase diagram of the studied H$_2$O ice phases, computed by MP2 in the CBS limit. The shaded regions and black labels indicate the phase with minimum computed Gibbs energy, while the blue lines and labels indicate the experimental phase boundaries.\textsuperscript{6}
order-disorder transitions that occur when moving from low temperature phases of ice to proton-disordered phases that border the liquid region.

Indeed, the slope of the coexistence lines is telling of the nature of the phase transition. It is defined by the Clausius-Clapeyron relation as the ratio of specific entropy change to specific volume change. Density-driven phase transitions such as V ↔ VI thus correspond to horizontal coexistence lines, while entropy-driven order-disorder transitions are characterized by vertical coexistence lines. The curves of Gibbs energy equivalence for each pair of ice phases were generated and are plotted in Figure 3.12. This is shown because not all phases appear in Figure 3.11 due to stiff competition between the phases, but their Gibbs energy surfaces and crossings thereof are still well-defined.

The simulated and observed coexistence lines are far from colinear. As discussed, very small shifts in relative lattice energies could suffice to move the transition pressures significantly. However, for all six experimentally-observed transitions, the value \((dP/dT)\) is reproduced reasonably well. Moreover, the regions of stability are correctly identified for each pair; that is, when the PT region is divided into stability regions of A and B, they are never swapped between MP2/CBS and experiment. The V ↔ VI coexistence line acquires some negative slope at higher temperatures due to the entropy...
increase, but resembles the experimental line below 100 K.

### 3.4 Conclusions

We have applied BIM-MP2 to the Gibbs free energy calculations of ices Ih, III, III, IX, V, and VI in the QHA. The equilibrium volumes predicted by MP2 are consistently low, but the lattice energies and bulk moduli predicted by MP2 in the complete basis set limit are reasonably close to experimental values. The INS spectra were computed from the phonon dispersion as the hydrogen amplitude-weighted phonon density of states. The covalent bending and stretching manifolds were reproduced quite accurately for all phases but ice Ih, while the pseudotranslational frequencies were significantly blueshifted compared with experiment. This shift was ascribed to the overbinding atomic embedding field and relatively small aug-cc-pVDZ basis used in the force constant calculations. Nevertheless, assignments of observed features to the BIM-MP2 spectra were relatively clear, especially the nuanced structure exhibited by proton-ordered phases.

From the QHA Gibbs energy calculations, the thermal expansivity and equations of state of the ice phases were simulated. We found that basis set extension from aug-cc-pVDZ towards the CBS limit significantly reduces the bulk modulus, giving good agreement between BIM-MP2 and experimental values. The softer crystals, in turn, exhibit a larger magnitude of thermal expansion. This significantly improves agreement with the observed thermal expansion of D$_2$O ice II. Vanishing of the bulk modulus at negative pressures was computationally observed for ices Ih, V, and VI, most likely corresponding to the proximity of a mechanical instability.

The phase diagram of the studied ices was not presented at BIM-MP2/aug-cc-pVDZ, because the lattice energy ordering of the phases is nearly backward. This is due to overbinding of BIM-MP2 exacerbated by the small aug-cc-pVDZ basis set with diffuse functions, which incurs significant BSSE and favors the high density phases V and VI too strongly. BIM-MP2 in the CBS limit predicts a reasonable slope of the coexistence line for each pair of phases, if not at the correction location in the PT plane. Moreover, relative stability of the phases on either side of each coexistence line is predicted correctly. Unfortunately, this method still falls short of an accurate prediction of the water ice phase diagram. Most significantly, ice II is everywhere metastable with respect to ices IX and V, but this failure is due to a relative energy difference of less than 1 kJ mol$^{-1}$ and could conceivably be overturned by changing parameters such as BIM truncation radii. In conclusion, an accurate computational phase diagram of the water ices requires sub-chemical accuracy in the electronic structure combined with anharmonic treatment of the lattice dynamics. It seems impossible to predict by fortuitous cancellation of errors and is therefore a stringent test for future embedded-fragment methods. The entire phase diagram of ice remains a great challenge for first-principles prediction by empirical potentials, DFT, and MO theories alike.
4 Molecular dynamics of liquid water

4.1 Introduction

Nearly all previous atomistic simulations of liquids have been performed with molecular mechanics (MM) force fields\textsuperscript{81–89} or with forces supplied by on-the-fly density-functional theory (DFT) calculations\textsuperscript{90–93} or by quantum mechanical/molecular mechanical (QM/MM) calculations.\textsuperscript{94} MM is limited to systems that are already well characterized, since force parameterization requires fitting to such \textit{a priori} information. DFT is a great improvement over MM because the former can be directly applied to any liquid systems. Yet, many of DFT functionals lack intermolecular dispersion interactions, despite their importance as a cohesive force in many liquids, or otherwise do not have sufficiently high accuracy. In fact, DFT simulations of bulk water have predicted a considerably elevated melting temperature ($T_m \approx 400$ K)\textsuperscript{95} and thus the existence of its liquid phase in rather unrealistic conditions.\textsuperscript{92} Although \textit{ad hoc} dispersion corrections have been shown to improve the description of liquid water,\textsuperscript{96, 97} such improvements are generally nonsystematic, and in these cases DFT falls short of being a predictive theory. One is thus interested in applying systematically improvable theories, such as \textit{ab initio} many-body perturbation (MP) and coupled-cluster (CC) theories, to the simulation of nature’s most important liquid, i.e., water, with the eventual goal of reproducing all its observed structural, dynamical, electronic, and spectroscopic features simultaneously, accurately, and from first principles.

An application of \textit{ab initio} MP or CC theory to a bulk liquid has not been possible because of its prohibitive cost. The situation has changed with the advent of the so-called embedded-fragment method,\textsuperscript{20} which enables a systematic, fast, and parallel-executable application of an \textit{ab initio} theory to large molecular clusters, crystals, and even liquids.\textsuperscript{98–100} The method divides a weakly interacting system into overlapping molecular dimers (“fragments”), which are embedded in the electrostatic environment of the whole system, and then applies well-developed molecular theories and software to these fragments to reconstruct an array of whole-system properties, with a computational cost scalable with respect to both system and computer sizes.

Fujita \textit{et al.}\textsuperscript{98} reported the first \textit{ab initio} Born–Oppenheimer molecular dynamics (BOMD) simulation of liquid water with the embedded-fragment scheme known as the fragment-molecular-orbital (FMO) method.\textsuperscript{101} They employed Hartree–Fock (HF) theory and the 6-31G* basis set as the \textit{ab initio} theory, which does not describe electron correlation and thus lacks the dispersion interaction. The effect of three-body interactions was assessed by Komeiji \textit{et al.}\textsuperscript{102} in a water cluster of ($H_2O)_{32}$ and that of electron
correlation by Mochizuki et al.\cite{mochizuki103} using second-order many-body perturbation (MP2) theory and the 6-31G* basis set on (H\textsubscript{2}O)\textsubscript{64}. Brorsen et al.\cite{brorsen99,brorsen100} repeated FMO-BOMD calculations of liquid water using more rigorous formulas for atomic forces at the HF and DFT level. All of these studies considered only the radial distribution functions and, owing to the smallness of the basis set employed, it is doubtful that the calculated results have much quantitative value. Perhaps the largest-scale electron-correlated calculation of liquid water performed thus far is an isobaric-isothermal Monte Carlo (MC) simulation at the MP2 level using a mixed Gaussian-planewave basis set.\cite{mp2104} Under the ambient conditions, it predicted a reasonable density of 1.02 g/cm\textsuperscript{3}, but did not provide any other meaningful experimental comparison such as detailed structures, diffusion rate, vibrational spectra, or dynamics of water molecule.

In the present Chapter, we perform BOMD simulations of liquid water using on-the-fly atomic forces obtained with thus-far the most accurate electronic structure theory for such simulations, i.e., the \textit{ab initio} spin-component-scaled MP2 (SCS-MP2) method\cite{bim106} with the aug-cc-pVDZ basis set. The necessary speedup of the \textit{ab initio} calculation is achieved by the embedded-fragment method known as the binary-interaction method (BIM)\cite{bim106} specifically extended for condensed-phase simulations. Refer to Chapter 1 for details about the BIM formalism and implementation. The choice of the theoretical level is based on our observation that it accurately reproduces the water-water interaction potentials of the coupled-cluster singles, doubles, and noniterative triples [CCSD(T)] with the aug-cc-pVQZ basis set, which is nearly exact. This theory is, therefore, expected to achieve an equally accurate quantum electronic description of the structural, dynamical, spectroscopic, and electronic properties of liquid water.

We apply this method to the radial distribution function, self-diffusion coefficient, monomer structure and its variation, coordination number and its fluctuation, intershell exchange mechanism, dipole moment and its distribution, as well as infrared (IR) and Raman spectra of liquid water (neat liquid H\textsubscript{2}O) under periodic boundary condition. We show that the simulations at the temperature (T) of 250 K and the density (\(\rho\)) of 1 g/cm\textsuperscript{3} predict the aforementioned properties in good agreement with those observed at ambient conditions. The results explain the shapes and widths of the OH stretching bands in the IR and Raman spectra, including the difference between the isotropic and anisotropic Raman components, which is known to reflect the different local hydrogen-bond environments surrounding the OH oscillators. They furthermore reveal intriguing dynamical features of the system, which are difficult to probe experimentally, such as a large fluctuation of the coordination number and the mechanism and time frame in which water molecules move across the first to second shell and thereby cause this fluctuation. They also inform us about the dipole moment and its distribution in the liquid phase as well as its breakdown into the permanent and induced components.

The properties calculated in this work range from the structural to dynamical to spectroscopic to electronic properties, which have hardly been studied simultaneously and with reasonable accuracy by any previously existing simulation methods. Therefore, the results in this Chapter can provide a useful benchmark against which some controversial issues on collective properties of liquid water can be reexamined.
4.2 Methods

4.2.1 Born-Oppenheimer molecular dynamics

BOMD simulations were carried out in the canonical (NVT) ensemble with the velocity Verlet algorithm in conjunction with the Nosé–Hoover chain method.\textsuperscript{107, 108} A time step of 1 fs was used. An initial liquid structure of $(\text{H}_2\text{O})_{32}$ in a cubic cell of side 9.858 Å was obtained from an equilibration run with the TTM3-F force field\textsuperscript{88} at $T = 300$ K. Using the same force field, we verified that the 32-water simulation yielded the calculated RDF which was unchanged from the result of a 512-water simulation and was, therefore, adequate. Using SCS-MP2/aug-cc-pVDZ or MP2/cc-pVDZ energies and atomic forces at $T = 250$ K, a 6-ps equilibration run was performed, after which the trajectory was sampled every 4 time steps during a 5-ps production run.

4.2.2 Self-diffusion coefficient

The self-diffusion coefficient ($D$) is related to the MSD by Einstein’s diffusion equation:

$$D = \frac{\text{MSD}}{6t} = \frac{1}{6t} \left\langle (\mathbf{r}(t) - \mathbf{r}(0))^2 \right\rangle. \quad (4.1)$$

In turn, MSD was obtained by the calculation of the squared relative displacement of the oxygen atoms as a function of time $t$ averaged over all water molecules in the unit cell.

4.2.3 IR and Raman spectra

The IR and Raman spectra were calculated by Fourier transformation of the time-correlation functions of the system’s dipole moment and polarizability, respectively. For a direct comparison with the experimental spectra, the harmonic quantum correction was applied to the classically computed spectra.\textsuperscript{109} IR intensity $I_{\text{IR}}(\omega)$ at frequency $\omega$ then becomes\textsuperscript{110}

$$I_{\text{IR}}(\omega) \propto \frac{\omega^2}{k_B T} \int_{-\infty}^{\infty} dt \ e^{-i\omega t} \left\langle \mathbf{u}(0) \cdot \mathbf{u}(t) \right\rangle, \quad (4.2)$$

where $\mathbf{u}(t)$ is the dipole moment of the system at time $t$, defined as the sum of the dipole moments of monomers, $\{\mathbf{u}_i\}$, in the unit cell,

$$\mathbf{u}(t) = \sum_i \mathbf{u}(t). \quad (4.3)$$

Here, the dipole moment of the $i$th monomer embedded in the electrostatic field, $\mathbf{u}_i$, is directly computed by the SCS-MP2 method.

The harmonic-quantum-corrected, Bose–Einstein-reduced isotropic and anisotropic
Raman spectra were calculated by\textsuperscript{110}

\begin{align}
I_{\text{Raman}}^{\text{iso}}(\omega) & \propto \frac{1}{9} \frac{\omega}{k_B T} \int_{-\infty}^{\infty} dt \, e^{-i\omega t} \langle \text{Tr}[\alpha(0)] \text{Tr}[\alpha(t)] \rangle, \\
I_{\text{Raman}}^{\text{aniso}}(\omega) & \propto \frac{2}{15} \frac{\omega}{k_B T} \int_{-\infty}^{\infty} dt \, e^{-i\omega t} \langle \text{Tr}[(\beta(0) \cdot \beta(t))] \rangle,
\end{align}

where \( \alpha \) is the polarizability tensor of the unit cell and \( \beta = \alpha - \text{Tr}[\alpha] I/3 \) with \( I \) being the identity tensor. Each element of \( \alpha \) is the sum of the corresponding elements of the embedded monomer polarizability tensors, \( \{\alpha_i\} \). The latter is evaluated as the ratio of the induced dipole moment \( \mu_{\text{ind}}^i \) to the applied electric field \( E \), e.g.,

\begin{equation}
(\alpha_{ij})_{xy} = \frac{(\mu_{\text{ind}}^i)_{x}}{E_y}.
\end{equation}

Induced dipole moments are, in turn, evaluated directly by the SCS-MP2 method for an embedded monomer with additional point charges of \( \pm 1 \) a.u. placed at \( r_{\text{cm}} \pm 100 \) bohr along each Cartesian axis. Here, \( r_{\text{cm}} \) denotes the center of mass of the monomer. The parallel-polarized (VV) and perpendicular-polarized (VH) Raman spectra, with each letter denoting the vertical (V) or horizontal (H) polarization direction of the incident and scattered light, are related to the isotropic and anisotropic components by

\begin{align}
I_{\text{Raman}}^{\text{VV}}(\omega) & = I_{\text{Raman}}^{\text{iso}}(\omega) + I_{\text{Raman}}^{\text{aniso}}(\omega), \\
I_{\text{Raman}}^{\text{VH}}(\omega) & = \frac{3}{4} I_{\text{Raman}}^{\text{aniso}}(\omega).
\end{align}

Using the classical MD with the TTM3-F force field,\textsuperscript{88} we verified that the simulation with a 1-fs time step yields the same IR spectra as the simulation with a 0.2-fs time step. The only exception is the positions of the O-H stretching bands (if not their widths and shape), in which a time-step error of up to 50 cm\(^{-1}\) (or \( \sim 1.5\% \)).

\subsection{4.2.4 Energy and gradient evaluation}

The potential energy, atomic gradients, and pressure of liquid water in each time step were calculated by the massively parallel implementation of the BIM using the SCS-MP2/aug-cc-pVDZ or MP2/cc-pVDZ method. The formalisms of BIM for condensed-phase simulations are described in Chapter 1. Briefly, it divides the liquid into monomers and overlapping dimers embedded in the electrostatic environment of the liquid. The potential energies and gradients of each of the monomers or dimers, treated quantum mechanically (QM), are evaluated with SCS-MP2 or MP2 implemented in a modified \textsc{nwchem}.\textsuperscript{111} They are used to reconstruct the potential energy, atomic forces, and pressure of the whole liquid under the periodic boundary conditions. The embedding field (BQ) consists of atomic point charges that approximate the electrostatic potential of each monomer and are self-consistently determined at the HF level. The calculation thus includes two-body electrostatic (i.e., induction or polarization), exchange, and dispersion interactions at the SCS-MP2 or
MP2 level and three-body and all higher many-body electrostatic interactions at the HF level. The truncation radii of the QM, BQ, and long-range (LR) regions, as defined by Equations 1.7–1.10 are as follows: \( R_{QM} = 7.5 \text{ Å} \), \( R_{BQ} = 9 \text{ Å} \), and \( R_{LR} = 12 \text{ Å} \).

### 4.3 Results

#### 4.3.1 Potential energy surface

The fidelity of BOMD simulations for bulk systems largely hinges on the reliability of the potential energy surface (PES) obtained from the electronic structure method, hence verifying that the latter for intermolecular interactions is a crucial prerequisite for such simulations. Figure 4.1 plots the PES’s of the water dimer in three geometrical configurations (having \( C_s \), \( C_i \), or \( C_{2v} \) symmetry) as a function of the O–O separation (\( R_{OO} \)).

Compared with the accurate reference data of MP2/aug-cc-pVQZ or CCSD(T)/aug-cc-pVQZ, the MP2/cc-pVDZ (top panel of Figure 4.1) overestimates the intermolecular attractions by as much as 50%, whereas the PES’s of MP2/aug-cc-pVDZ (middle panel) and SCS-MP2/aug-cc-pVDZ (bottom panel) are reasonable. This underscores the importance of diffuse basis functions in describing intermolecular forces; namely, basis sets without diffuse basis functions such as 6-31G* are inadequate for BOMD simulations of liquid water and water clusters.\(^{98,103}\) On the other hand, DFT calculations with the Becke–Lee–Yang–Parr (BLYP) functional and diffuse basis sets tend to underestimate \( \Delta E \).

SCS-MP2/aug-cc-pVDZ (bottom panel) describes the attractive tail of the PES slightly better than MP2/aug-cc-pVDZ. Its repulsive parts are a little too strong, making the PES’s somewhat shallower and equilibrium \( R_{OO} \) longer than the reference data. Nonetheless, the value of \( \Delta E \) predicted by SCS-MP2/aug-cc-pVDZ (4.86 kcal/mol) is closer to the complete-basis-set (CBS) limit of CCSD(T) of 5.02 kcal/mol (dashed lines) than the value obtained at MP2/aug-cc-pVDZ (5.26 kcal/mol), while the MP2-calculated \( \Delta E \) seems to converge more accurately at the correct limit with a further basis-set extension.\(^{112–114}\)

Since zero-point vibrational energy (ZPVE) corrections are not at present considered in our BOMD simulations, we find SCS-MP2/aug-cc-pVDZ convenient because its \( \Delta E \) (4.86 kcal/mol) is smaller and its \( R_{OO}^{min} \) (2.942 Å) is larger than those of CCSD(T)/CBS (5.02 kcal/mol and 2.910 Å) and thus expected to be closer to experimental values that include the ZPVE effect.\(^{115,116}\) Needless to say, the use of a large basis set such as aug-cc-pVQZ in \textit{ab initio} electron-correlated BOMD simulations takes a prohibitive computational cost and is out of the question.

In this study, therefore, we primarily used SCS-MP2/aug-cc-pVDZ to generate atomic forces. We additionally used MP2/cc-pVDZ as the secondary method to assess the effect of overestimated \( \Delta E \) and underestimated \( R_{OO}^{min} \) on the simulation results.
Figure 4.1: The potential energy surfaces (PES’s) of water dimers in the $C_s$, $C_1$, and $C_2$ geometries as a function of the oxygen-oxygen separation ($R_{OO}$) were obtained at various theoretical levels (solid curves). The dimer geometries were generated by varying $R_{OO}$, with all other geometrical parameters frozen at their MP2/aug-cc-pVTZ optimized values under the appropriate symmetry constraint. The accurate results from MP2/aug-cc-pVQZ (dashed curves) and CCSD(T)/aug-cc-pVQZ (dots) are superimposed. The binding energy at CCSD(T)/CBS is indicated as dashed black lines.
Figure 4.2: Oxygen-oxygen radial distribution function. The oxygen-oxygen radial distribution functions (RDF’s) of liquid water were obtained from the BOMD simulations at $T = 250$ K and $\rho = 1$ g/cm$^3$ and were compared with the experimental result$^{117}$ at $T = 25{^\circ}$C.

4.3.2 Radial distribution function

No a priori information exists about the phase diagram of water for the $ab\ initio$ potentials. The first step is, therefore, to identify where the liquid phase boundaries exist. One approach is to determine the melting temperature computationally.$^{95,96}$ Another is to vary the temperature such that the BOMD simulation reproduces the experimental liquid structure at ambient conditions.$^{92}$ The liquid structure is inferred from the radial distribution function (RDF), which can be observed by neutron or X-ray scattering experiment$^{117,118}$ and serves as a fingerprint of local structures and thus phases. We first performed BOMD simulations at $T = 300$ K, which produced a gas-phase-like RDF, and then reduced the temperature until the calculated RDF coincided with that of the observed result for the liquid phase at ambient conditions.

Figure 4.2 shows the oxygen-oxygen RDF obtained from the SCS-MP2/aug-cc-pVDZ simulation at $T = 250$ K. It reproduces the experimental RDF of the liquid phase observed at $T = 25$ °C obtained from the X-ray diffraction technique.$^{117}$ The calculated position of the first peak is close to the experimental value ($\sim 2.8$ Å) and the agreement of RDF in the region of $R_{OO} > 3.2$ Å is striking. The slight difference in the calculated peak intensity from the experimental one at $R_{OO} = 2.8$ Å can be diminished by taking into account the nuclear quantum effect, as noted from the previous path-integral molecular dynamics simulation.$^{93}$ The average coordination numbers of 4.7 ($R_{OO} \leq 3.4$ Å) or 4.4 ($R_{OO} \leq 3.3$ Å) are in excellent agreement with the experimental values.$^{117}$ We therefore conclude that the SCS-MP2/aug-cc-pVDZ simulation at $T = 250$ K and $\rho = 1$ g/cm$^3$ provides an accurate description of the structural features of liquid water at ambient conditions.

In contrast, the RDF of MP2/cc-pVDZ at $T = 250$ K exhibits an excessively broad first peak, which is a manifestation of an over-coordinated local structure with an av-
verage coordination number of as large as 8 ($R_{OO} < 3.7 \text{ Å}$). This, in turn, is due to too strong an attractive intermolecular interaction predicted by MP2/cc-pVDZ, as evidenced by Figure 4.1. This first peak is followed by the second peak, which is further displaced at longer $R_{OO}$, indicating that the second shell is pushed outward by the overcrowded first shell. Hence, the MP2/cc-pVDZ description of liquid water structure is nonphysical and this is largely due to the smallness of the basis set.

The isobaric-isothermal MC simulations at the MP2 level with a mixed Gaussian-planewave basis set was performed by Del Ben et al. They yielded the liquid density ($\rho = 1.02 \text{ g/cm}^3$) slightly larger than the correct value. This is in qualitative agreement with our simulation, in which the average pressure is found to be $-0.6 \pm 0.5 \text{ GPa}$ at $T = 250 \text{ K}$ and $\rho = 1.0 \text{ g/cm}^3$; the larger calculated liquid density at 1 bar means that a negative pressure is required to maintain it at 1 g/cm$^3$. The RDF calculated at $T = 295 \text{ K}$ by Del Ben et al. was, however, overstructured, suggesting that the liquid is super-cooled, whereas the RDF obtained from our simulation is much closer to experiment. The source of the quantitative difference is hard to pinpoint because there are many approximations involved in both studies. We speculate, however, that the most likely primary source is the different choices of the basis set, in particular, the use of a pseudopotential in the preceding HF calculation in Del Ben et al.

### 4.3.3 Self-diffusion coefficient

Whereas the RDF characterizes time-averaged local structures, the self-diffusion coefficient ($D$) measures dynamical properties of liquids. Its value is related to the mean square displacement (MSD) by Einstein’s diffusion equation. The calculated value of $D \approx 0.27 \text{ Å}^2/\text{ps}$ from our BOMD simulation at $T = 250 \text{ K}$ compares favorably with the experimental value of $0.23 \text{ Å}^2/\text{ps}$ at $T = 25 \degree \text{C}$. Hence, we argue that the present simulation properly describes the dynamical properties of liquid water at ambient conditions also.

### 4.3.4 Water molecular geometry in the liquid phase

Each water molecule undergoes a significant geometry change on going from the gas to liquid phase. The water molecule in the gas phase has the OH bond length ($R_{OH}$) of 0.966 Å and HOH angle ($\angle \text{HOH}$) of 103.9 $\degree$ at the SCS-MP2/aug-cc-pVDZ level, which are in good agreement with the experimental values of 0.957 Å and 104.5 $\degree$. In the liquid phase, the calculated bond lengths and angles form distributions (Figure 4.3) with $R_{OH} = 0.980 \pm 0.019 \text{ Å}$ and $\angle \text{HOH} = 104.7 \pm 4.6 \degree$. Two experimental investigations of liquid water yielded measurements of ($R_{OH} = 0.970 \pm 0.005 \text{ Å}$; $\angle \text{HOH} = 106.1 \pm 0.9 \degree$) and ($R_{OH} = 0.983 \pm 0.008 \text{ Å}$; $\angle \text{HOH} = 104.1 \pm 1.9 \degree$). Remarkably, the BOMD simulation predicts an average OH bond elongation (relative to the gas phase) of 0.014 Å, which is in excellent agreement with the measured elongation of 0.013 Å. It is notable that the bond angle in liquid water ranges from 90$\degree$ to 120$\degree$ with a large standard deviation of 4.6 $\degree$. 

60
Figure 4.3: Bond lengths and angles. Distributions of bond lengths (top) and angles (bottom) of water molecules in the liquid phase, compared to the optimized geometry of the water molecule in the gas phase. The blue dots and bars indicate the averages and standard deviations.

4.3.5 Coordination numbers and intershell exchange

Figure 4.4 plots the distribution of the coordination number, that is, the number of water molecules in the first shell, as well as the distribution of the hydrogen-bond number. According to our simulation, a water molecule has an average coordination number of 4.7 and an average hydrogen-bond number of 3.8. The former value is in good agreement with experiment. In most cases (65%), each molecule is hydrogen-bonded to four neighbors with two serving as a H-donor (hydrogen donating water) and two as a H-acceptor (hydrogen accepting water), while the remaining neighbors in the first shell stay non-hydrogen-bonded or weakly hydrogen-bonded, as shown in Figure 4.5.

The coordination number displays a surprisingly large fluctuation (4.7 ± 0.9), ranging from mono-coordination to octa-coordination. Overall, tetra-coordination is the most common (41%), followed by penta-coordination (34%) and hexa-coordination (15%). Tri-coordination (6%) and hepta-coordination (3%) are also significant, while bi-coordination (0.6%), octa-coordination (0.1%), and mono-coordination (0.01%) occur much less frequently. A majority of the penta- and higher-order coordinations have one or more non-hydrogen-bonded or weakly hydrogen-bonded water molecules in the first shell, judging from the rapid falloff of the hydrogen-bond number after four.

Figure 4.6 illustrates the time-evolution of the coordination numbers of three randomly selected water molecules. The large and rapid variation in each case may be
Figure 4.4: Coordination number and hydrogen-bond number. The distribution of the number (CN) of water molecules in the first shell \((R_{OO} \leq 3.36 \, \text{Å})\) excluding the central molecule and the distribution of the number (HN) of hydrogen bonds with the central molecule (with a hydrogen bond being considered to exist between two water molecules when \(R_{OO} \leq 3.36 \, \text{Å} \) and \(\angle \text{HOO} \leq 40^\circ\)).

Figure 4.5: Snapshots of local structures in liquid water. The snapshots of the penta-coordinated water molecule are drawn. It is, on average, hydrogen-bonded to two H-acceptors and two H-donors. The fifth water molecule in the first shell stays non-hydrogen-bonded or weakly hydrogen-bonded. It can be (a) near the D zone or (b) near the A zone in the contour maps in Figure 4.7.
Figure 4.6: Fluctuation of the coordination number. The number of water molecules in the first shell ($R_{OO} \leq 3.36 \text{ Å}$) as a function of time is tracked down for three selected water molecules.
Figure 4.7: Oxygen-atom population. The \((R_{OO}, \theta)\) and \((R_{OO}, \varphi)\) contour maps of the oxygen atom population within the second-shell RDF peak \((R_{OO} \leq 4.5\,\text{Å})\) and the \((\theta, \varphi)\) contour map within the first shell \((R_{OO} \leq 3.4\,\text{Å})\). \(R_{OO}, \theta,\) and \(\varphi\) are spherical coordinates as defined above. \(D\) and \(A\) stand for the hydrogen-donating (H-donor) and accepting (H-acceptor) water molecules, respectively.

related to the significant density fluctuations observed even in a nm-sized droplet of liquid water, as discussed in a recent study\(^{123}\).

Figure 4.7 shows the \((R_{OO}, \theta)\) and \((R_{OO}, \varphi)\) population contour maps of the oxygen atoms for \(R_{OO}\) up to the second-shell RDF peak \((R_{OO} \leq 4.5\,\text{Å})\) as well as the \((\theta, \varphi)\) population contour map for the first shell \((R_{OO} \leq 3.4\,\text{Å})\). They reveal that H-acceptors (the \(A\) zone) are found around \(R_{OO} = 2.8\,\text{Å}, \theta = \pm 50^\circ,\) and \(\varphi = 0^\circ,\) while H-donors (the \(D\) zone) populate around \(R_{OO} = 2.8\,\text{Å}, \theta = 120^\circ,\) and \(\varphi = \pm 90^\circ.\) A significant population density is observed at \(\theta = 100^\circ\) around \(R_{OO} = 3.4\,\text{Å}\) in the \((R_{OO}, \theta)\) contour. This is the likely path through which a non-hydrogen-bonded or weakly hydrogen-bonded neighbor in the \(D\) zone migrates from the first shell to the second shell and vice versa, causing the fluctuation in the coordination number.

Our BOMD trajectory data such as those shown in Figure 4.6 provide even more detailed information about key events responsible for the coordination number fluctuation. Let us define a 5-4-5 hetero-exchange as the process by which a penta-coordinated water molecule loses its non-hydrogen-bonded neighbor to become tetra-coordinated, and a different water molecule takes its place (hetero-exchange), restoring the penta-coordinated state. This event is frequently observed in our trajectories, lasting an average of \(ca.\) 100 fs and occurring to a penta-coordinated species approximately every 1.0 ps. Similarly, we use the term the 4-5-4 self-exchange to describe the transient attachment of a non-hydrogen-bonded water molecule to the first shell of a tetra-coordinated water, followed by the detachment of the same molecule. This process occurs, on aver-
age, approximately every 1.5 ps with a duration of ca. 80 fs. The 5-4-5 self-exchange and 5-6-5 hetero-exchange are also noted at an average frequency of every few ps. The high frequency and long duration of these exchange events involving tetra- and penta-coordinated species are due to their greater stability and thus longer average lifetime.

### 4.3.6 Dipole moments

The dipole moment of a water molecule in the liquid phase has been under debate for a long time, partly because there is no experimental means of measuring it directly. Accurate theoretical predictions are, therefore, extremely valuable, but previously reported computed values ranged widely from 2.6 to 3.0 D. An earlier DFT-MD study placed it at 2.66 D, but later revised it to 2.95 D. The latter, larger predicted value may be traced to the significantly overestimated polarizability characteristic of DFT. Subsequently, Badyal et al. determined the partial charge on the hydrogen atom as \( q_H = 0.5 \pm 0.1 \) a.u. by X-ray diffraction. On this basis, they suggested the dipole moment of 2.9 ± 0.6 D, which had such a large error bar that it encompassed nearly all computationally suggested values.

The dipole moment of the water molecule in the gas phase calculated by SCS-MP2/aug-cc-pVDZ is 1.878 D, which agrees well with the experimental value of 1.855 D. Figure 4.8 plots the distributions of the permanent dipole moments \( \{ \mu_0^i \} \) (obtained with the embedding field turned off) and the total dipole moments \( \{ \mu_i \} \) (calculated with the embedding field within 9 Å and thus including the induced dipole moments) in the liquid phase. The permanent dipole moment \( \mu_0^i \) at SCS-MP2/aug-cc-pVDZ is 1.872 ± 0.07 D, whereas the total dipole moment \( \mu_i \) is 2.67 ± 0.2 D. Hence, the induced dipole moment is significant and amounts to as much as 0.80 D on average. It could further increase slightly by the nuclear quantum effect, but perhaps by not more than 0.05 D.

Furthermore, the total dipole moment is broadly distributed, showing large fluctuations from 1.8 to 3.5 D, with about 5% exceeding 3.0 D, owing more to the diversity of electrostatic fields than to the monomer geometry fluctuation. The partial charge of the hydrogen atom, \( q_H = 0.48 \pm 0.03 \) a.u. (Figure 4.8), is in good agreement with the mean of the experimental value of Badyal et al., \( q_H = 0.5 \pm 0.1 \) a.u. The difference in the total dipole moment between our study and Badyal et al. may be traced to the difference in the water structure: the OH bond length proposed by one of the authors of Ref. 126 is 1.045 Å and is much greater than our predicted (0.980 Å) or experimental values (0.970–0.983 Å). Our predicted dipole moment (\( \mu_i = 2.67 \) D) is in excellent agreement with other independent experimental expectation (\( \sim 2.6 \) D) and theoretical calculation (\( \sim 2.65 \) D). In short, the dipole moment suggested by Badyal et al. (2.9 D) is, in view of this and other accurate theoretical calculations, questionable and likely overestimated.

### 4.3.7 IR and Raman spectra

The SCS-MP2/aug-cc-pVDZ BOMD simulation predicts the vibrational frequencies of the water molecule in the gas phase at 1676, 3800, and 3900 cm\(^{-1}\). Though BOMD
Figure 4.8: Dipole moments and hydrogen partial charge. The distribution of the partial charges of the hydrogen atoms (top) and the distribution of the permanent ($\mu^0_i$) and total ($\mu_i$) dipole moments (bottom) of the water molecules in the liquid phase. The total dipole moment is obtained with the embedding field within the cut-off distance $R_{BO} = 9$ Å, whereas the permanent dipole moment with the field turned off.
Figure 4.9: IR and Raman spectra. The IR and Raman spectra of liquid water were calculated with the SCS-MP2/aug-cc-pVDZ simulation at $T = 250$ K and $\rho = 1$ g/cm$^3$. The experimental IR spectrum at $T = 25$ °C is taken from Ref. 135 and the experimental VV and VH Raman spectra at $T = 23$ °C from Ref. 134. The dashed curves are obtained by scaling the frequencies of the calculated spectra by 0.93 to empirically take into account the ZPVE (see the text).

simulations sample the fully anharmonic PES, owing to the classical treatment of vibrations, these frequencies are too high relative to the experimental values (1595, 3657, and 3756 cm$^{-1}$), for reasons that are well understood. To correct such deviations rigorously, one should use path-integral molecular dynamics, which can take into account the ZPVE or quantum nuclear effects. A more empirical correction is to simply multiply the calculated frequencies by a scale factor (0.96 for the gas phase and 0.93 for the liquid phase) that can bring the calculated peak positions in better agreement with those observed experimentally.

The IR and Raman spectra of liquid water computed from the SCS-MP2/aug-cc-pVDZ BOMD simulation at $T = 250$ K are compared with the corresponding experimental results at ambient conditions in Figure 4.9. Unlike most calculated spectra, each band in our simulated spectra is not convoluted with an assumed band shape, but is instead associated with a physically meaningful width that reflects the diversity of local environments in which the vibrations occur.

The BOMD simulation reproduces all three major features of the observed IR spectrum (top panel): the broad, intense OH stretching band above 3000 cm$^{-1}$, the narrower peak due to the bending mode at around 1600 cm$^{-1}$, and the manifold of intermolecular
vibrations below 1000 cm$^{-1}$.

The OH stretching band in the observed IR spectrum peaks at about 3400 cm$^{-1}$ with a weak shoulder around 3250 cm$^{-1}$ and a full width at half maximum (FWHM) of 375 cm$^{-1}$.\textsuperscript{135,136} Remarkably, the BOMD simulation predicts the OH stretching band with the correct band width (FWHM of about 375 cm$^{-1}$) and slightly asymmetric band shape, which also agrees with the observed shape, though the calculated peak positions of 3550 and 3700 cm$^{-1}$ are blue-shifted. It is striking that the classical “continuum” picture,\textsuperscript{131} in which diverse instantaneous hydrogen-bond environments surrounding the OH stretching vibrations give rise to the band width, is borne out by our simulation. Our time-domain computation of the spectra should also naturally account for the motional narrowing of this band, which is estimated to be as much as 30% of the width.\textsuperscript{137}

The OH stretching band of the observed VV Raman spectrum has bimodal peaks at about 3250 and 3400 cm$^{-1}$ and a FWHM of about 425 cm$^{-1}$, while the observed VH spectrum shows the OH stretching band around 3460 cm$^{-1}$ with an asymmetric line shape and a FWHM of about 300 cm$^{-1}$.\textsuperscript{134} The calculated VV and VH Raman spectra are also in agreement with the observed spectra. It explains the asymmetric line shape of the OH stretching band in the VH spectrum and the double-peaked nature of the band in the VV spectrum, including the difference in peak positions between the VH and VV components, known as the Raman noncoincidence effect.\textsuperscript{138} With the aforementioned frequency scaling by a factor of 0.93, the simulated VV and VH spectra are again in excellent agreement with their experimental counterparts both in peak positions and shapes.

It could be emphasized that our spectral simulation does not involve any adjustable parameter (which is not to say that there is no approximation) in the anharmonic force field, mode-mode coupling, polarizability tensor, etc., unlike previous analyses on the same bands.\textsuperscript{131} Note that polarizable model potentials of water\textsuperscript{88,110} currently fall short of simultaneously accounting for polarization effects on intermolecular interactions and optical response, such as the Raman effect. Our \textit{ab initio} method can achieve just that, although the peak positions are systematically overestimated because of the classical treatment of vibrations.

### 4.4 Conclusions

We have performed \textit{ab initio} BOMD simulations of liquid water at the SCS-MP2 and MP2 levels of theory and computationally reproduced the structural properties (the radial distribution function, monomer geometry and its fluctuation) and structure-driven collective properties (the self-diffusion coefficient, coordination number and its variation, intershell exchange) as well as the electronic and response properties (permanent and induced dipole moments, IR spectra, and isotropic and anisotropic Raman spectra) simultaneously, accurately, and from first principles. The level of theory has been chosen to reproduce the CCSD(T)/aug-cc-pVQZ results of the water dimer, allowing us to study the electronic details of liquid water with unprecedented accuracy.
The calculated properties of liquid water are in good agreement with the experimental results, when the latter are available. The simulated IR and Raman spectra correctly predict the band shapes and widths of the OH stretching bands and the difference between their isotropic and anisotropic Raman components, which are exceedingly difficult with classical MD with parameterized force fields, even if they are polarizable.

For properties that are difficult to probe experimentally, our simulations provide unique insights. The average bond length and angle of the water molecule in the liquid phase are found to increase only minimally, but range from 0.92 to 1.05 Å and from 90° to 120°. The coordination number fluctuates even more greatly from monocoordination to octa-coordination, though the tetra- (41%) and peta-coordinations (34%) are most common. Such large fluctuation in the coordination number might lead to significant density fluctuations. We have identified 5-4-5 hetero-exchange and 4-5-4 self/hetero-exchange as the key events, occurring approximately every 1–2 ps with an average duration of ca. 0.1 ps and causing the large fluctuation in the coordination number. Such exchanges take place as a non-hydrogen-bonded water molecule in the D zone migrates between the first and second shells. The total dipole moment ($\mu = 2.67 \pm 0.2$ D) of the water molecule in the liquid phase also shows a large fluctuation, ranging from 1.8 to 3.5 D, reflecting more the diversity of the electrostatic field surrounding the water molecule than the geometry fluctuation.

Hence, the present method allows us to assess the validity of uncertain results or conclusions about a whole range of liquid water properties. It can also be exploited to study other unknown collective properties of liquid water and various solvents as well as colligative properties of solute-solvents generally.
5 Embedded-fragment valence bond charge transfer method

5.1 Introduction

The previous chapters have illustrated the power of the binary interaction method (BIM) in applying accurate ab initio molecular orbital (MO) theories to the Born-Oppenheimer potential energy surface (PES) of water. Otherwise prohibitively expensive post-Hartree Fock (HF) methods, such as Møller-Plesset perturbation (MP) or coupled-cluster (CC) theory, are now routinely used for predictive crystal thermochemistry in general or as force fields for ab initio molecular dynamics simulation of liquids. However, BIM and most other embedded-fragment methods can only treat insulating systems with electron density localized on weakly interacting molecular fragments. Because the methods rely on the premise of a rapidly convergent many-body expansion of the total energy, systems with significant partial charge transfer (CT), or non-integer electron counts per molecule, are ill-suited for fragmentation. Moreover, molecular integrals between fragment orbitals are not considered explicitly, and thus desirable quantities like transfer integrals and valence band structures in molecular organic semiconductors are not accessible. Fragment methods that can handle CT systems would find broad applicability and are therefore an important target for method development.

Relevant efforts in treating CT systems can be divided into two categories. The first includes valence bond (VB) or configuration interaction (CI) calculations of the total energy in a nonorthogonal basis of localized or diabatic states. Generally speaking, these are multiconfigurational methods designed to handle static correlation, which can be defined as the exact correlation energy of the valence electrons in a minimal basis. To illustrate the relevance in CT systems, first consider how the static correlation energy would vanish for a cluster of helium atoms having no virtual orbitals in MO theory. It can become significant upon ionization, as the electron hole may occupy one of several near-degenerate valence orbitals. The same idea applies to electron or hole transfer in other molecular systems with nondegenerate frontier MOs: the static correlation is small, albeit nonzero, in the neutral system, but it can become large upon addition of a charge carrier. This multireference scenario complicates DFT and unrestricted HF (UHF) calculations, which tend to respond by over-delocalizing charge or breaking symmetry, respectively.

Van Voorhis developed a constrained density functional theory (CDFT) CI method...
to deal with static correlation in a DFT framework.\textsuperscript{27,139} CDFT is first used to build localized diabatic states, and then a CI matrix is diagonalized in this basis to recover the static correlation energy and adiabatic ground state. It is easy to see the formal similarity of this approach to VB theory. It was extended by another group and applied directly to molecular cluster cations to model charge resonance with tight-binding DFT.\textsuperscript{141} Other electronic structure methods in the VB/CI category include a block-localized wave function VB method\textsuperscript{142} for CT systems and a fragment exciton model\textsuperscript{143} for CI singles-type vertical excitation energies (not CT \textit{per se} but a formally similar idea).

Methods in the second category take a different one-electron approach and target the delocalized orbitals involved in CT directly. Here, the total energy and related quantities (equilibrium structures, vibrational frequencies, binding energies) of the cluster are not available. Instead, transfer integrals and ionization potentials (or band structures) are computed to gain insight into electron transfer kinetics and energetics. Valeev’s superposition-of-fragment-states (SFS) method\textsuperscript{28,144} constructs an effective one-electron Hamiltonian in a diabatic basis of fragment ionized states. The diagonal elements are computed with equation-of-motion ionization-potential CC singles and doubles (EOM-IP-CCSD), plus an embedding correction from a full-system HF/Koopmans’ calculation. The basis includes low-lying ionized monomer states and gives accurate results for several ionization energies of the benzene dimer, for example. Similarly, the fragment MO-linear combination of MOs (FMO-LCMO) method produces a one-electron cluster Hamiltonian in the basis of monomer valence HF MOs.\textsuperscript{145} The Hamiltonian is built up from FMO HF calculations requiring only monomers and dimers. Again, total energy is not available from the Hamiltonian; instead, a spectrum of approximate HF cluster orbital energies is used in combination with Koopmans’ theorem to approximate transfer integrals or ionization potentials. A similar approach, using MOs as the basis for crystalline orbital calculations, was employed six years earlier as an efficient route to the band structure of four pentacene polymorphs.\textsuperscript{146}

In this Chapter, we introduce an embedded-fragment valence bond method for CT systems (VBCT) that fits with the first category of CI/VB-type approaches to calculating total energy. VBCT is closer in spirit to conventional fragment methods, however, because it aims to reproduce the non-fragmented energy at a given level of theory. We consider only finite molecular cluster cations with a single positive charge and evaluate the performance of VBCT based on its fidelity in reproducing the conventional UHF energy or second-order MP (MP2) energy based on the former reference. In this view, the role of the diabatic basis is to assign integer electron counts and enable fragmentation when CT is significant.

VBCT shares many desirable attributes of the embedded-fragment scheme: it is extensible in terms of electron correlation theory, in that any post-UHF MP or CC theory can be applied. Moreover, the cost scaling is low and fixed regardless of correlation method, and the procedure consists of independent fragment calculations that are trivially run in parallel. By including many-body polarization effects through the embedding field and all two-body electrostatic, exchange, and dispersion interactions, the method can reproduce total cluster energies quantitatively. Most importantly, it is
easily implemented by interfacing with unmodified quantum chemistry software, and there is no need for explicit evaluation of molecular integrals of any type.

The remainder of this chapter is structured as follows: in Section 5.2, we introduce the formalisms and procedure used to perform VBCT calculations on cluster cations. We then analyze the accuracy and behavior of the method for a variety of molecular clusters in Section 5.3. Finally, we conclude in Section 5.4 with a summary of VBCT’s strengths, weaknesses, and the next steps toward improving and applying this method to complex CT processes in condensed phase systems.

5.2 Computational methods

5.2.1 Diagonal elements

The only type of system considered in this work is a finite molecular cluster cation with unit positive charge. The delocalization of the positive charge can be regarded as a resonance between charge-localized valence bond configurations. Let us introduce an eective Hamiltonian in the VBCT basis, with the goal of approximating the cluster UHF or post-UHF energy using only fragment calculations performed by conventional quantum chemistry software.

The adiabatic ground state of an \( N \)-molecule cluster cation, \( |\Psi_0\rangle \), is proposed to be a linear combination of diabatic VBCT states:

\[
|\Psi_0\rangle = \sum_{i=1}^{N} c_{i0} |\Phi_i\rangle
\]

(5.1)

where \( |\Phi_i\rangle \) corresponds to the many-electron state with the positive charge localized on fragment \( i \). The total electronic energy \( E_0 \) and coefficients \( c_{i0} \) are determined from the lowest eigenvalue and corresponding eigenvector of the effective Hamiltonian

\[
H = \begin{pmatrix}
\langle \Phi_1 | H | \Phi_1 \rangle & \cdots & \langle \Phi_1 | H | \Phi_N \rangle \\
\vdots & \ddots & \vdots \\
\langle \Phi_N | H | \Phi_1 \rangle & \cdots & \langle \Phi_N | H | \Phi_N \rangle
\end{pmatrix}
\]

(5.2)

where \( \langle \Phi_i | H | \Phi_j \rangle \) is the Hamiltonian matrix element between orthonormal charge local states \( |\Phi_i\rangle \) and \( |\Phi_j\rangle \). A number of related approaches\(^{27,141} \) (mentioned in Section 5.1) posit explicit forms for \( |\Phi_i\rangle \) and proceed by rigorous evaluation of the Hamiltonian matrix elements. Here, we bypass the explicit construction of \( |\Phi_i\rangle \) and use a series of embedded-fragment MO calculations on the monomers and overlapping dimers comprising the cluster to directly define the cluster electronic Hamiltonian matrix elements in the VBCT basis. The rationalization of our approach is as follows.

Suppose a neutral cluster contains one fragment with a significantly lower ionization potential than all the other fragments. One example considered in this Chapter is
Na(H$_2$O)$_3$. For the ionized form of this cluster, the ground state $|Ψ_0⟩$ should be dominated by one configuration $|Φ_{Na}⟩$ corresponding to Na$^+$(H$_2$O)$_3$ in the VBCT framework.

If $|Ψ_0⟩ ≈ |Φ_{Na}⟩$ and is normalized, then it follows that $⟨Φ_{Na}|H|Φ_{Na}⟩$ should be a good approximation to the total cluster ground state electronic energy $E_0$. This idea guides our scheme for computing diagonal elements of $H$. The scheme is similar to BIM, where energy is calculated by the embedded many-body expansion truncated at second order (refer to Chapter 1). We follow this idea because it is known to give accurate HF and post-HF energies for weakly interacting molecular clusters.

The crucial difference from BIM is that, for the calculation of $⟨Φ_i|H|Φ_i⟩$, all $(N-1)$ dimers $ij$ containing the fragment $i$ must not allow any positive CT from $i$ to $j$. Otherwise, SCF relaxation of these dimers would make it impossible to construct VBCT basis states where charge is unfavorably localized. We refer to these as charge-local dimers; their HF wavefunctions are not stationary, but such calculations can be formulated as constrained HF procedures such as SCF-MI.

The diagonal element is computed as

$$⟨Φ_i|H|Φ_i⟩ = E_i^+ + \sum_{j \neq i} E_j + \left(\text{HF}E_{ij}^{(i+)} - \text{HF}E_i^+ - \text{HF}E_j\right)$$

$$+ \sum_{j < k \neq i} \left(E_{jk} - E_j - E_k\right),$$

(5.3)

where $E_i^+$, $E_j$, and $E_{jk}$ are the conventional UHF or electron-correlated post-UHF energies of cation monomer $i$, neutral monomer $j$, and neutral dimer $jk$, respectively. $\text{HF}E_{ij}^{(i+)}$ is the non-stationary UHF energy of the charge-local dimer $ij$ with charge localized on $i$. $\text{HF}E_i^+$ and $\text{HF}E_j$ are the conventional UHF parts of the energy from the first and second terms of equation 5.3.

All of the fragment energies entering $⟨Φ_i|H|Φ_i⟩$ are computed in an embedding field of atomic charges (ESP) fit to reproduce the monomer electrostatic potentials. This field is determined from a self-consistent cycle of monomer UHF calculations where fragment $i$ carries a positive charge and all other fragments $j$ are neutral. At the end of the cycle, the canonical UHF MO coefficients of each fragment associated with VBCT state $|Φ_i⟩$ are stored.

$\text{HF}E_{ij}^{(i+)}$ is computed by concatenating the canonical occupied MOs of $i$ and $j$, forming the dimer Fock matrix from the resulting block diagonal density matrix, and evaluating the non-self-consistent energy. This is accomplished in NWChem with the fragment initial guess or in Gaussian09 with the Guess(Fragment=2, Only) keyword. Because the Fock matrix is diagonalized once, if there is overlap between the fragment AOs, a small degree of CT from $i$ to $j$ may occur. This could be rigorously corrected with a constrained HF method, but the described method is more widely available, easily implemented, and works quite well for a variety of systems, as demonstrated in the Results.

The charge-local dimer energy $\text{HF}E_{ij}^{(i+)}$ can be robustly calculated by conventional
quantum chemistry programs, unless the HOMO of fragment $i$ is degenerate, which is the case with molecules such as Ne$^+$. As shown by our results, the method can remain reasonably accurate for such systems, but unfortunately a precise value of $E_0$ loses meaning because $HF E^{|i+\rangle}$ acquires a small degree of arbitrariness. This is a manifestation of the degeneracy of the singly occupied MO and consequent arbitrariness of the electron hole orientation. It has no effect in the symmetric monomer cation calculation, but this symmetry is broken in the charge-local dimer, and $j$ experiences an unpredictably-oriented dipole from $i$.

Let us illustrate this situation with the Ne$^+ UHF$ calculation behavior. At 1.54 Å separation, UHF easily converges to a symmetric wave function with equal spin density on both atoms. As expected, the charge-local dimer spin densities equal 1.0/0.0 or 0.0/1.0, but the energy is not reproducible; it varies by 0.03 $E_h$ with a simple translation of the dimer! When the bond length is increased to 2.54 Å, the charge-local dimer energies converge to a less erratic value that varies by only 0.0002 $E_h$ with translation. This variation can be tracked to rotations of the Ne$^+$ singly-occupied MO which create a varying dipole-induced-dipole contribution to $HF E^{|i+\rangle}$, which becomes small compared to the monopole interaction as the internuclear distance increases.

The computation of $\langle \Phi_i | H | \Phi_i \rangle$ proceeds in the following order:

1. Calculate self-consistent ESP embedding field with charge localized on fragment $i$.
2. Store the canonical fragment MOs, which are polarized by the field computed in the previous step.
3. Enumerate all monomers $j$ and dimers $jk$. Perform the appropriate energy calculation according to the values of $i$, $j$, and $k$.
4. Evaluate $\langle \Phi_i | H | \Phi_i \rangle$ via equation 5.3.

Note that all of the above steps, including the embedding field calculation, must be repeated $N$ times for each of the charge local states. Since there are $N(N - 1)/2$ dimers, the number of fragment calculations scales as $O(N^3)$. If the embedding field is disabled, the number of fragment calculations is reduced to $O(N^2)$. Because the diagonal element calculations are completely independent, the method remains trivially parallelizable. Nested parallelism works best here; the processors are divided into $N$ groups which work concurrently for all values of $i$. This accomplished with the split communicator functionality of MPI when the number of processors exceeds $N$.

### 5.2.2 Off-diagonal elements

The off-diagonal elements $\langle \Phi_i | H | \Phi_j \rangle$ are defined such that the conventional UHF or post-UHF energy of a dimer cation, $E^{|i+\rangle}_j$, is reproduced exactly by the lowest eigenvalue of the 2x2 dimer Hamiltonian in the basis of $|\Phi_i\rangle$ and $|\Phi_j\rangle$. To arrive at the formula, consider the diagonal element expression in a cluster cation containing only the two
molecules \( i \) and \( j \):

\[
\langle \Phi_i | H | \Phi_j \rangle_{\text{dimer}} = E_i^+ + E_j^+ - HF E_{ij} + cor\ E_i^+ + cor\ E_j
\]

where \( cor\ E \) denotes only the correlation (post-UHF) energy and vanishes in calculations that simply target the UHF energy. Given the quantities \( \epsilon_i, \epsilon_j \), and the stationary dimer cation energy \( E_{ij}^+ \), the coupling is reverse-engineered from the 2x2 secular eigenvalue equation:

\[
\langle \Phi_i | H | \Phi_j \rangle = -\sqrt{(E_{ij}^+ - \epsilon_i)(E_{ij}^+ - \epsilon_j)}
\]

A subtle but important detail is that \( \epsilon_i \) is computed with a two molecule embedding field, regardless of cluster size, with the positive charge localized on \( i \), neutral \( j \), and no charges from the other molecules \( k \). The diagonal element \( \langle \Phi_i | H | \Phi_i \rangle \) of the full Hamiltonian cannot be substituted for \( \epsilon_i \).

Moreover, the calculation of \( E_{ij}^+ \) requires some care, especially for symmetric dimers, because the UHF can have difficulty converging when multiple closely-spaced or degenerate UHF solutions exist. Also, a symmetric initial guess (which is usually the default in quantum chemistry software) leads to a symmetric UHF wave function, even when a lower energy symmetry-broken solution exists.

This is handled robustly by repeating the UHF part of each \( E_{ij}^+ \) calculation three times: once with the default (superposition of atomic densities) guess, once from the converged monomer MOs of cation \( i \) and neutral \( j \), and once from the MOs of neutral \( i \) and cation \( j \). Calculations that fail to converge are terminated, and the minimum energy UHF wavefunction is used directly for \( E_{ij}^+ \) or as the reference for the post-UHF calculation.

Referring back to our example of the Ne\(^+\) \(_2\) dimer at 2.54 Å separation, the UHF with default initial guess converges rapidly but to a high-energy state with symmetric spin density. The charge-local guesses converge to a lower energy solution with 0.99/0.01 or 0.01/0.99 spin density, albeit with greater difficulty due to the degeneracy of the two symmetry-broken states.

### 5.2.3 Total energy and charge distribution

Regardless of the overlap between fragment AOs, the method remains exact for any dimer by construction. Hence, we interpret that overlap is absorbed by the off-diagonal element calculation and \( H \) is constructed in an effectively orthonormal VBCT basis.

The overlap matrix element \( \langle \Phi_i | \Phi_j \rangle \) is thus taken as \( \delta_{ij} \). The ground state VBCT energy \( E_0 \) and wave function \( |\Psi_0\rangle \) are obtained by diagonalizing \( H \) and taking the lowest eigenvalue and corresponding eigenvector.

The squared eigenvector coefficient \( c_{i0}^2 \) can be interpreted as amount of positive
charge on fragment $i$ in the cluster. The cluster atomic charges are approximated as a linear combination of the $N$ atomic charge ESP vectors (from step 1 of the diagonal element procedure), weighted by the coefficients $c_{2i}^2$.

### 5.2.4 Rationale

Alternate versions of VBCT can be formulated by redefining $\langle \Phi_i | H | \Phi_i \rangle$ and computing off-diagonal elements so as to preserve exactness for a 2 molecule cluster. One version truncates equation 5.3 after the second term; that is, all two-body interactions are omitted from the diagonal element and contained in the coupling calculation. This calculation, with or without embedding, fails to reproduce a trimer PES even qualitatively and can become ill-defined at repulsive regions of the PES where $\epsilon_i < E_{ij}^-$. Another version defines $\langle \Phi_i | H | \Phi_i \rangle$ by the neutral cluster BIM energy plus the $\Delta$-UHF or -MP2 ionization potential of embedded monomer $i$. While this approach is appealing because it does not rely on non-canonical dimer MOs, it too exhibits serious flaws that are exposed in larger cluster calculations.

Overall, experience shows that the diagonal elements must contain the one- and two-body electrostatic and electron-correlation interactions in the cluster, but the two-body calculations must not allow CT relaxation to betray the charge-local character of the state. In equation 5.3, the two-body electron correlation energy is computed between all neutral dimers but not charge-local dimers, where non-canonical MOs complicate post-HF theories. Instead, the correlation in charged dimers is captured by the off-diagonal element calculation, along with the pair relaxation due to charge delocalization. When fragments are close together, the overlap tends to promote formation of bonding MOs and charge delocalization, while the proximity of electron densities increases the magnitude of the correlation energy; both of these factors contribute to the magnitude of the coupling.

### 5.3 Results

#### 5.3.1 He$^+_2$

To check the correctness of the implementation, VBCT was first applied to the UHF and MP2 energies of the Helium dimer cation in the aug-cc-pVDZ basis. Figure 5.1 shows that VBCT reproduces the energy of a dimer exactly, as required. Let us compare the charge distributions, represented by the atomic charges fit to the molecular electrostatic potential (ESP), between the conventional calculation and VBCT. The calculation method for the VBCT atomic charges is explained in section 5.2.3. We find in Figure 5.2 that UHF spontaneously breaks $D_{\infty h}$ symmetry when the He$^+_2$ bond is stretched to about 1.8 Å. While VBCT tracks the conventional calculation energy exactly, it restores the symmetric charge distribution at long separations.
Figure 5.1: UHF and MP2/aug-cc-pVDZ energies of He$_2^+$, computed with conventional methods (Exact) and VBCT (Fragment).

Figure 5.2: ESP charge distributions from the calculations in Figure 5.1. The atoms are labeled ‘1’ and ‘2’ in the legend. Where only one color is visible, the two lines coincide exactly.
Figure 5.3: Structure of He$_3^+$ when the dimerization reaction coordinate is 1.0. The dimer has 1.1 Å bond length, and the third atom is located along the bond axis 2.0 Å away.

Figure 5.4: UHF and MP2/aug-cc-pVDZ energies of He$_3^+$ along the symmetric stretching coordinate, computed with conventional methods (Exact) and VBCT (Fragment).

5.3.2 He$_3^+$

He$_3^+$ is perhaps the smallest nontrivial system by which the performance of VBCT can be judged. We evaluate its UHF and MP2/aug-cc-pVDZ PESs along three coordinates. The first is a symmetric stretching of the linear structure. The second is referred to as a “dimerization” reaction coordinate, wherein the symmetric trimer becomes a dimer and lone atom (Figure 5.3). The third is simply an excursion of the third lone atom away from the global minimum He$_2^+$ structure, along the dimer bond axis. Let us first consider the symmetric stretching coordinate (Figures 5.4 and 5.5). VBCT reproduces both the UHF and MP2 energies quantitatively. It correctly predicts the greater localization of charge on the inner atom. The CT behavior is essentially the same for the HF and MP2 VBCT calculations (which have the same ESP charge vectors for each basis state but differ in the coefficients $c_{0i}$; refer back to the methods section). The central atom charge in VBCT increases monotonically to a He-He spacing of 2.5 Å, while the conventional UHF calculation exhibits a minimum around 1.2 Å before localizing strongly. UHF localizes the charge on the inner atom more completely than the VBCT calculations.

Next, we analyze the dimerization reaction coordinate (Figures 5.6 and 5.7). On an absolute scale, VBCT reproduces the UHF and MP2/aug-cc-pVDZ energies quantitatively to within 1.5 mE$_h$. Note that the PES is relatively flat along this coordinate and becomes even flatter with MP2, where the difference between the PES maximum and minimum along the entire coordinate is only 1 mE$_h$. VBCT still manages to track the curvature of the PES on this tiny scale qualitatively.

In Figure 5.7 it’s clear that the charge distribution at reaction coordinate 0 resembles that of the symmetric trimer in Figure 5.5. As the symmetric structure gives way to a
dimer (atoms 1 and 2) plus a lone He (atom 3), the dimer takes almost the entire positive charge, while atom 3 becomes almost neutral. The exact UHF method shows slightly more CT to the lone atom at the reaction coordinate value 1.0, and there is a slight disproportionation of charge towards atom 1 which is not predicted by VBCT.

Finally, the excursion coordinate data is presented in Figures 5.8 and 5.9. The performance of VBCT and nature of the CT along this coordinate are very similar to the dimerization. Again, the fragment UHF and MP2 energies are quantitatively accurate, and the localization of charge on the dimer is predicted robustly. VBCT also correctly shows that as the lone atom moves away from the dimer, it becomes increasingly neutral.

5.3.3 He$_4^+$

The linear He$_4^+$ PES was computed along symmetric and asymmetric stretching coordinates. Figures 5.10 and 5.12 show that the UHF and MP2/aug-cc-pVDZ energies are accurately reproduced by VBCT. A larger error is incurred in the highly repulsive region of the symmetric stretching PES.

For the symmetric stretch in Figure 5.11, conventional UHF predicts that initially, about 70% of the charge is equally shared between the inner atoms (2 and 3) of the tetramer, while the remaining 30% is shared equally between the outer atoms (1 and 4). At around 1.7 Å, there is a spontaneous symmetry breaking, and UHF once again localizes the charge on atom 2. While VBCT reproduces the UHF charge distribution reasonably well up to this point, it never breaks symmetry. Instead, up to 90% of the charge becomes localized on the inner dimer.

At a value of -0.2 along the asymmetric stretching coordinate, the inner two atoms (labeled 2-3) form a dimer with distant lone atoms (1-4) on either side (see Figure 5.12). When the coordinate reaches 1.0, the inner dimer bond has broken to form a symmetric pair of two separated dimers (1-2 and 3-4). In Figure 5.13, both UHF and VBCT agree that the charge is initially delocalized over the inner 2-3 dimer. As the coordinate passes 0.6, VBCT predicts that the charge becomes equally resonant over
Figure 5.6: UHF and MP2/aug-cc-pVDZ energies of He$_3^+$ along the dimerization coordinate, computed with conventional methods (Exact) and VBCT (Fragment). The upper panel is less flattering and shows UHF and MP2 energies on separate scales spanning a few mE$_h$. The lower panel shows the same data on a shared scale.

Figure 5.7: ESP charge distributions from the calculations in Figure 5.6. The legend labels the atoms 1,2,3 in order from left to right.
Figure 5.8: UHF and MP2/aug-cc-pVDZ energies of He$^+_4$ along the excursion coordinate, computed with conventional methods (Exact) and VBCT (Fragment).

Figure 5.9: ESP charge distributions from the calculations in Figure 5.8. The legend labels the atoms 1,2,3 in order from left to right.

Figure 5.10: UHF and MP2/aug-cc-pVDZ energies of He$^+_4$ along the symmetric stretching coordinate, computed with conventional methods (exact) and VBCT (fragment).
all four atoms of the dimers 1-2 and 3-4. For UHF on the other hand, the 2-3 and 1-4 charge distributions bifurcate at a reaction coordinate of 0.6, once again breaking symmetry. UHF ultimately localizes 90% of the charge on the 1-2 dimer alone.

5.3.4 He$_8^+$

The PES of He$_8^+$ is shown in Figure 5.14 on separate scales for the UHF and MP2 energies at aug-cc-pVDZ. The reaction coordinate involves a He pentamer ring shuttling over a cationic trimer core. VBCT tracks the energy almost quantitatively; the total octamer energy error does not exceed about 1.5 mE$_h$, which means the error-per-He is less than 0.12 kcal/mol.

The reason for the good agreement is that this structure is relatively "easy" for VBCT; the He$_3$ core is fixed at the global minimum geometry for an isolated trimer cation, which consists of a helium dimer (1.1 Å bond) and a third atom 2 Å away. It is encouraging that both HF and MP2 errors remain small for a system of this size and that the strong localization of charge on the dimer is properly encoded in the 8-dimensional Hamiltonian with no a priori information.

The charge distributions in Figure 5.15 show less agreement between VBCT and the conventional calculation. Both unambiguously assign the majority of charge correctly to atoms 1 and 2. However, the distribution is markedly more asymmetric in conventional UHF, and this skewness increases as the ring displacement coordinate passes 0.0 Å. Meanwhile, about 25% of the charge transfers into the solvent ring. VBCT predicts almost no CT from the trimer core into the ring, and a very slight disproportionation of charge between atoms 1 and 2 of the dimer.

5.3.5 Ar$_4^+$

As discussed in the Methods section, the precise VBCT energy for a given level of theory loses meaning when some fragments in the system contain a degenerate HOMO, due to a limitation of the charge-local dimer calculations. However, the method is still executable and will produce results that will vary slightly for arbitrary shifts of the
Figure 5.12: Top: The He$_4^+$ asymmetric stretching coordinate at -0.2 and 1.0. Bottom: UHF and MP2/aug-cc-pVDZ energies of He$_4^+$ along the asymmetric stretching coordinate, computed with conventional methods (exact) and VBCT (fragment).

Figure 5.13: ESP charge distributions from the calculations in Figure 5.12. The legend labels the atoms in order from left to right.
Figure 5.14: Top: The He$_8^+$ structure used. Bottom: UHF and MP2/aug-cc-pVDZ energies of He$_8^+$ along the pentamer shuttling coordinate, computed with conventional methods (Exact) and VBCT (Fragment).

Figure 5.15: ESP charge distributions from the calculations in Figure 5.14. The legend labels the atoms; and 1-2 belong to the dimer core, 3 is the third atom on the dimer axis, and 4-8 form the solvating pentamer ring.
5.3.6 \([\text{Na(H}_2\text{O)}_3]\)^+ 

Two heterogeneous cluster cations structures are tested, containing one sodium atom and three water molecules. The PESs are plotted with a description of the reaction coordinate in Figures 5.18 and 5.20.

One can see that both the UHF and MP2/6-31+G* energies are reproduced quantitatively. This is not surprising, as these cations have a strongly single-reference char-
Figure 5.18: Top: The $[\text{Na(H}_2\text{O)}_3]\text{]^+}$ linear structure. The leftmost water (with bond to sodium) undergoes excursions along the O-Na axis. Bottom: UHF and MP2/6-31+G* energies of $[\text{Na(H}_2\text{O)}_3]\text{]^+}$ along the linear water excursion coordinate, computed with conventional methods (exact) and VBCT (fragment).

character in the VBCT basis; that is, the charge is unambiguously localized on the Na atom, and VBCT is not really even needed over BIM.

The charge distribution of both conventional UHF and VBCT places nearly 100% of the positive charge on the Na atom. A closer look at the partial CT to the water molecules in Figures 5.19 and 5.21 shows very little deviation away from 0.0 in VBCT. The small trends in the UHF calculation may be an artifact of the finite grid used in fitting the ESP charges to the cluster electrostatic potential.

5.3.7 $(\text{C}_2\text{H}_4)_3^+$

Figures 5.22 and 5.23 show the VBCT approximations to the UHF and MP2/6-31+G* PES of the $(\text{C}_2\text{H}_4)_3^+$ stacked cation with respect to rotation of the inner molecule. The details of the geometry are in Figure 5.22. At both levels of theory, the VBCT PES is within 1 m$E_a$ of the exact trimer PES. Interestingly, MP2 predicts a qualitatively different profile near 90° with a local minimum instead of a saddle point. Moreover, Figure 5.23 shows that VBCT correctly predicts localization of charge on the central molecule (2), while the remaining charge is shared equally between the outer two molecules (1 and 3).
Figure 5.19: ESP charge distributions from the calculations in Figure 5.18. The legend labels the molecules, and the Na line is intentionally outside the scale of this figure.

Figure 5.20: Top: The [Na(H₂O)_3]⁺ triangle structure. The sodium atom is displaced vertically along the axis normal to the O-O-O plane. Bottom: UHF and MP2/6-31+G* energies of [Na(H₂O)_3]⁺ along the triangle sodium displacement coordinate, computed with conventional methods (exact) and VBCT (fragment).
Figure 5.21: ESP charge distributions from the calculations in Figure 5.20. The legend labels the molecules, and the Na line is intentionally outside the scale of this figure.

Figure 5.22: Top: The \((\text{C}_2\text{H}_4)_3^+\) structure used. At 0 degrees, three ethylene molecular planes are parallel and spaced by 4.0 Å. At 90 degrees, the central molecule is rotated about the axis perpendicular to the stacking direction and C-C bond axis. Bottom: UHF and MP2/6-31+G* energies of \((\text{C}_2\text{H}_4)_3^+\) along the central molecule rotation coordinate, computed with conventional methods (Exact) and VBCT (Fragment).
5.3.8 \((\text{C}_2\text{H}_4)_4^+\)

The \((\text{C}_2\text{H}_4)_4^+\) tetramer is studied in a parallel stacked geometry with planes separated by 4.0 Å and 6.0 Å. The CT reaction coordinate is described and illustrated in Figure 5.24. Briefly, the inner two ethylene geometries interpolate between the neutral and cation-optimized geometries, while the outer ethylenes remain fixed at the neutral geometry.

Let us first consider the geometry with 4.0 Å intermolecular spacing. Near the reactants and products configurations (CT coordinate close to 0 or 1), the VBCT method performs very well in predicting both UHF and MP2/6-31+G* energies (Figure 5.24). It also predicts the energy quantitatively at the D$_{2h}$ symmetric structure where the CT coordinate is 0.5. However, in the neighborhood of this partial CT at coordinate values of 0.4 and 0.6, a pathology likely related to the multireference character manifests as an oscillation in the VBCT PES away from the conventional calculation energy. Note in Figure 5.25 that this erratic behavior is accompanied by an oscillation of charge between the reacting fragments (2 and 3). Still, the absolute errors do not exceed 2 m$E_h$ per molecule or about 1 kcal/mol, which is encouraging.

When the CT coordinate is 0.5, it is very interesting to contrast the charge distribution of VBCT versus conventional UHF in the 4.0 Å-spaced geometry (Figure 5.25). Notice that the UHF charge distribution breaks D$_{2h}$ symmetry, and the charge discontinuously jumps from molecule 3 to 2. This nonphysical behavior and the cusp in the PES are related and well-known limitations of UHF and post-UHF theories for doublets. In contrast, VBCT predicts a symmetric charge distribution at 0.5 CT, and its energy also agrees quantitatively with the exact HF and MP2 energies at this point. The oscillation of charge and erratic PES in the vicinity of this symmetric structure might be somehow related to an inconsistency between the symmetric VBCT-UHF wavefunction and the cusped UHF PES it is trying to mimic. This issue should be revisited with methods that do not depend on the UHF doublet reference, e.g. EOM-IP-CC theories.

Notice the charge is less strongly localized in the VBCT-MP2 calculation. For clarity, the charges on the outermost spectator molecules are omitted from Figure 5.25. In the range of 0.4 to 0.6, VBCT predicts a significant oscillation of charge into and out
Figure 5.24: Top: The \((\text{C}_2\text{H}_4)_4^+\) CT reaction coordinate.\(^{148}\) Four ethylene molecular planes are parallel with a 4.0 Å separation. The outer molecules are fixed at the neutral molecular geometry. The inner two molecule geometries are a linear interpolation between the cation and neutral geometries. At CT coordinate 0.0, ethylene 2 has the cation geometry (shorter C-C bond) and 3 has the neutral geometry. At CT coordinate 0.5, the system has \(D_{2h}\) symmetry. Bottom: The UHF and MP2/6-31+G* energies of \((\text{C}_2\text{H}_4)_4^+\) with 4.0 Å intermolecular spacing along the CT coordinate, computed with conventional methods (exact) and VBCT (fragment).
of these spectators which may or may not be physical. The only point of comparison here is the UHF charge distribution, which is known to bias strong localization.

The energy profile of this \((\text{C}_2\text{H}_4)_4^+\) system changes drastically when the intermolecular spacing increases to 6.0 Å. Figures 5.26 and 5.27 show the same calculation results as above, but for the slightly expanded system. Now there is no cusp, and VBCT reproduces the PES of both UHF and MP2/6-31+G* without erratic behavior. Still, all calculations agree that the charge ultimately hops from molecule 3 to 2 as the CT coordinate changes from 0 to 1. It is intriguing that at the larger spacing, though, the charge first hops onto the outer molecules 4 and 1 before settling on 2. Again, VBCT predicts a totally symmetric charge distribution at the halfway point of CT, unlike UHF. Finally, notice how the VBCT-MP2 calculation predicts a more gradual, smooth, and monotonic decay of the positive charge on 3 and corresponding growth of the charge on 2.
5.3.9 \((\text{H}_2\text{O})_4^+\)

The water tetramer cation PES and charge distribution at UHF and MP2/6-31+G* are shown in Figures 5.28 and 5.29. The reaction coordinate is a scale factor of the side length of the square formed by the four oxygen atoms, while the molecular internal coordinates remain fixed. The performance of VBCT in predicting this system’s energy is significantly worse than for any of the previous clusters. When the scale factor is 0.9, the errors become as large as 8 kcal/mol. Considering the charge distributions, UHF strongly localizes the hole on one of two molecules, depending on the cluster perimeter, but it is evenly distributed over all four molecules of the tetramer in VBCT. With this symmetric geometry, the charge distribution in VBCT is inevitably also symmetric by the construction of the Hamiltonian; there is nothing in the diagonal or off-diagonal element calculation to stabilize or destabilize one of the four sites. It seems that there are four nearly degenerate UHF solutions and this is a very difficult system, in general.
Figure 5.28: Top: The \((\text{H}_2\text{O})_4^+\) structure used. The four O atoms form a square whose side length is scaled from an equilibrium value by the reaction coordinate while the O-H bonds remain fixed. Bottom: UHF and MP2/6-31+G* energies of \((\text{H}_2\text{O})_4^+\) along the O-O distance scaling coordinate, computed with conventional methods (Exact) and VBCT (Fragment).

Figure 5.29: ESP charge distributions from the calculations in Figure 5.28. The legend labels the molecules 1-4.
5.4 Conclusions

As a first step toward extending embedded-fragment methods to handle molecular systems with significant electron delocalization, we developed the valence bond charge transfer (VBCT) scheme and applied it to a variety of singly ionized molecular cluster cations. Even without fragmentation, the problem of treating molecular ion doublet states is often complicated due to the multireference character of the ground state.\textsuperscript{148} The VBCT secular equation is defined entirely in terms of monomer and dimer fragment energies, and as an energy-based scheme, it works quite well in reproducing the cluster ion energies of Helium. In stable hydrated ion clusters, where the adiabatic ground state closely resembles one diabatic charge-local state (single-reference), the method almost reduces to BIM and predicts the electron-correlated energy quantitatively.

A technical detail related to the construction of the diabatic states is the use of so-called charge-local dimer energies. Without a more involved calculation, these quantities are ill-defined for fragments with a degenerate HOMO. Nevertheless, the method is still applicable to such systems, which include the noble gas cluster cations besides Helium, with the caveat that a certain (not necessarily large) degree of uncontrollable error is incurred. In principle, the method could be rigorously corrected by an orbital rotation procedure or the use of special HF iterations that enforce charge local constraints from the outset.\textsuperscript{147}

In two systems with significant multireference character, we observed erratic or poor performance of the VBCT energy prediction. These were the closely-stacked ethylene tetramer cation and symmetric water tetramer. As mentioned, these systems are intrinsically challenging for UHF, which is known to exhibit nonphysical behaviors like a cusped PES. Consequently, UHF and less sophisticated post-UHF methods may be inappropriate for evaluating VBCT in these systems.

The VBCT state couplings are constructed to reproduce total dimer cation energies; molecular Hamiltonian and overlap integrals are never explicitly considered, though they are certainly implicit in the dimer relaxation. Therefore, it remains unclear how faithfully VBCT couplings reproduce rigorous ET calculations of off-diagonal Hamiltonian matrix elements in a diabatic basis. Future work will address this and the applied question of electron transfer rate calculations.

The full spectrum of VBCT Hamiltonian eigenvalues should be considered and compared against charge transfer excitation energies computed by EOM-CC or related theories. It is conceivable that VBCT couplings and overlaps should be reverse-engineered from ground and excited-state energies of charged dimers. Whether this more involved procedure improves the accuracy of the CT energy spectrum must be tested numerically. Transition dipole moments between VBCT eigenstates could be approximated from the ESP charge vectors and used to model the optical absorption bands related to CT. Once analytical (or semi-analytical) gradients of the VBCT energy are implemented, the method may also be used to optimize complex cluster cation geometries and compute vibrational frequencies and spectra. Trends in cluster ion

94
binding energies, for instance, are of great experimental interest and could be targeted directly by the method.
Conclusions

In Chapter 2, BIM was used to calculate the Gibbs energy of proton-disordered ice Ih, using MP2 in the aug-cc-pVDZ and aug-cc-pVTZ basis sets. Lattice energies were additionally computed at the CCSD level of theory. One question was the effect of embedding field on the observed properties. It was found that the atomic embedding charges best fit to the water molecule electrostatic potentially somewhat overestimate the molecular dipole moment and contribute to overbinding of the crystal at MP2/aug-cc-pVDZ. This is manifest both in the lattice energy and the phonon density of states, where pseudotranslational and librational modes are blueshifted relative to the dipolar embedding calculations. The anharmonic variation of mode frequencies with volume was computed explicitly, without any approximation based on the mode Grüneisen parameters. The QHA enabled predictions of the H_2O and D_2O ice Ih volumes at finite temperature and pressure. BIM-MP2 with the atomic embedding field reproduced the anomalous volume isotope effect, wherein the heavier isotopomer possesses a slightly larger molar volume than H_2O ice. This effect was traced to a near-zero cancellation of volume-expanding and contracting effects between all phonon groups with competing anharmonicities. The dominant contributors to volume expansion and contraction at 0 K are the librational and OH stretching phonon groups, respectively. The experimental thermal expansivity was reproduced accurately up to 80 K, at which point the BIM-MP2 ices did not expand as fast the observed. This can be attributed to a number of effects, including the greatly overestimated bulk modulus of ice Ih by MP2/aug-cc-pVDZ. Geometry optimizations at high pressure led to a flattening of acoustic phonon branches and eventual loss of symmetry and volume collapse in ice Ih. This was taken as computational detection of a mechanical instability causing the pressure-induced amorphization to high density amorphous ice. An immediate volume collapse and loss of symmetry was also observed in a short constant-temperature, constant-pressure BOMD simulation at 2.5 GPa and 75 K using on-the-fly MP2/aug-cc-pVDZ, supporting the findings from static geometry optimization and force constant analyses.

The study of ice was extended to other crystalline phases in Chapter 3, where BIM MP2 calculations in the aug-cc-pVDZ basis set or CBS limit were performed on ices Ih, II, III, IX, V, and VI. In agreement with the basis set comparisons of Chapter 2, it was found that MP2/aug-cc-pVDZ overbinds and overestimates the bulk modulus to a much greater extent than the CBS calculation. Part of this can be attributed to basis set superposition errors, which vanish in the CBS limit but are exacerbated with incomplete basis sets and diffuse functions. The phonon dispersion was computed for each phase at MP2/aug-cc-pVDZ, and INS intensities were simulated as the hydro-
gen amplitude-weighted phonon density of states. The covalent vibrational bands were accurately aligned with experimental measurements for all phases, but the pseudo-translational bands in MP2 were significantly blueshifted due to the overbinding effect. Still, an assignment of peaks between the simulated and observed spectra was mostly clear. While the volumes were consistently overestimated, BIM-MP2/CBS predicted bulk moduli of each phase with far greater accuracy than BIM-MP2/aug-cc-pVDZ. The Gibbs free energy surfaces in temperature and pressure were compared to estimate phase boundaries between the ices. While calculations in the aug-cc-pVDZ basis erroneously favored the high-pressure phases V and VI over Ih even at ambient pressure, the MP2/CBS calculations at least predicted a reasonable stability region for ice Ih. The slopes of coexistence lines were qualitatively correct, but transition pressures and temperatures were still rather far from experiment. We concluded that the determination of the ice phase diagram remains exceedingly difficult and requires sub-chemical accuracy to discriminate between the nearly isoenergetic phases II, III, IX, and V.

In Chapter 4, BIM was used to produce on-the-fly SCS-MP2/aug-cc-pVDZ forces for Born–Oppenheimer molecular dynamics simulation of liquid water. The spin component scaling of MP2 energies was used instead of conventional MP2, because it more faithfully reproduces the attractive tail of the water dimer PES at CCSD(T)/aug-cc-pVQZ. The radial distribution function was found to be in excellent agreement with experiment, given the density fixed at experimental value and temperature of 250 K. At the experimental density, the mean simulation pressure was found to be -0.6 GPa, reflecting once more MP2/aug-cc-pVDZ’s tendency to overestimate the cohesive dispersion interactions in water. A broad distribution of molecular geometries and dipole moments in the liquid phase is indicative of the diverse local environments experienced by water in the liquid phase. The IR and Raman spectra were found to be in reasonable agreement with the observed, especially in relative intensity and width of the OH stretching and HOH bending bands. The OH stretching frequencies were somewhat overestimated due to the lack of zero point vibrational effects in the purely classical motion of the nuclei. Finally, the local coordination of water molecules was analyzed, and penta-coordinate structures was found to exist with almost the same probability as the more intuitive tetra-coordinated water. This fifth water occupying a non-hydrogen-bonded region relative to the tetramer plays a large role in the fluctuation of coordination number and exchange of water molecules between coordination sites.

The focus was shifted away from condensed phase water in Chapter 5, where a new method was proposed to handle clusters with significant charge resonance of an electron hole by fragmentation. The molecular cluster cation system was chosen as one of the simplest prototypes for charge transfer to guide method development. The proposed VBCT scheme constructs an effective Hamiltonian in the basis of valence bond-like charge-local configurations. Rather than defining configurations explicitly and evaluating matrix elements in terms of such configurations, the effective Hamiltonian was directly constructed with a physically-motivated BIM-type energy expression for the diagonal elements. Central to the method is the evaluation of non-stationary Hartree–Fock energies of dimers, which are needed to preserve charge-locality for each
state. The coupling matrix elements were computed to reproduce the energy of a dimer exactly by construction. In doing so, we arrived at a scheme requiring only monomer and dimer calculations with an embedding-field, much like the conventional BIM. It was tested for its ability to predict conventional UHF and UMP2 energies of molecular cluster cations of helium, argon, ethylene, water, and water-sodium. The preliminary results for PES’s of these systems are encouraging, except for the water tetramer PES which is significantly overbound by VBCT. We proposed that the origin of this difficulty is related to the strongly multireference character of the symmetric tetramer cation ground state. In general, UHF tends to break symmetry and favor charge localization, which may necessitate reference methods other than UHF/UMP2 to benchmark VBCT.
References


