THE UNIVERSITY OF CHICAGO

FIRST PRINCIPLES SIMULATIONS OF VIBRATIONAL SPECTRA OF AQUEOUS SYSTEMS

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE INSTITUTE FOR MOLECULAR ENGINEERING
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

BY
QUAN WAN

CHICAGO, ILLINOIS
AUGUST 2015

For Jing

TABLE OF CONTENTS

LIST OF FIGURES vii

LIST OF TABLES xii

ACKNOWLEDGMENTS xiii

ABSTRACT xiv

1 INTRODUCTION 1

2 FIRST PRINCIPLES CALCULATIONS OF GROUND STATE PROPERTIES . 5
2.1 Density functional theory 5
2.1.1 The Schrödinger Equation 5
2.1.2 Density Functional Theory and the Kohn-Sham Ansatz 6
2.1.3 Approximations to the Exchange-Correlation Energy Functional . . . 9
2.1.4 Planewave Pseudopotential Scheme 11
2.1.5 First Principles Molecular Dynamics 12

3 FIRST-PRINCIPLES CALCULATIONS IN THE PRESENCE OF EXTERNAL ELECTRIC FIELDS 14
3.1 Density functional perturbation theory for Homogeneous Electric Field Perturbations 14
3.1.1 Linear Response within DFPT 14
3.1.2 Homogeneous Electric Field Perturbations 15
3.1.3 Implementation of DFPT in the Qbox Code 17
3.2 Finite Field Methods for Homogeneous Electric Field 20
3.2.1 Finite Field Methods 21
3.2.2 Computational Details of Finite Field Implementations 25
3.2.3 Results for Single Water Molecules and Liquid Water 27
3.2.4 Conclusions 32

4 VIBRATIONAL SIGNATURES OF CHARGE FLUCTUATIONS IN THE HYDROGEN BOND NETWORK OF LIQUID WATER 34
4.1 Introduction to Vibrational Spectroscopy 34
4.1.1 Normal Mode and TCF Approaches 35
4.1.2 Infrared Spectroscopy 36
4.1.3 Raman Spectroscopy 37
4.2 Raman Spectroscopy for Liquid Water 39
4.3 Theoretical Methods 41
4.3.1 Simulation Details 41
4.3.2 Effective molecular polarizabilities 42
4.3.3 Molecular Polarizabilities 43
4.4 Results and Discussion 45

4.4.1 Calculated Raman Spectra: Comparison with Experiments 45
4.4.2 Low-Frequency Bands 46
4.4.3 Molecular Polarizabilities 49
4.5 Summary and Conclusions 53

5 A FIRST-PRINCIPLES FRAMEWORK TO COMPUTE SUM-FREQUENCY GENERATION VIBRATIONAL SPECTRA OF SEMICONDUCTORS AND INSULATORS 55
5.1 Introduction 55
5.2 Basic Theories of SFG Spectroscopy 58
5.2.1 Interfacial and Bulk Contributions 59
5.2.2 The Calculation of the linear and nonlinear susceptibilities 62
5.2.3 The Origin Dependence Problem in the Calculation of Quadrupole Moments 65
5.2.4 Polarization Combinations 66
5.3 Computational Details 67
5.4 Tests and Validation of the Method 69
5.4.1 Local Dielectric Constant Profile 69
5.4.2 Electrostatic Correction for Slab Models 70
5.4.3 Effect of Incident Angles Used in Different Experiments 71
5.4.4 Convergence of SFG spectra with respect to the slab thickness 71
5.4.5 SFG spectra from FPMD simulations 73
5.4.6 The Origin Dependency of Quadrupole Contributions 73
5.5 Results and Discussion 74
5.6 Conclusion 78

6 SOLVATION PROPERTIES OF MICROHYDRATED SULFATE ANION CLUSTERS: INSIGHTS FROM *AB INITIO* CALCULATIONS 80
6.1 Introduction 80
6.2 Theoretical Methods 82
6.3 Results and Discussion 84
6.3.1 Structural Properties of 12-Water Sulfate Clusters 84
6.3.2 Structural Properties of 13-Water Sulfate Clusters 87
6.3.3 Electronic Properties 91
6.4 Conclusions 95

7 ELECTRONIC STRUCTURE OF AQUEOUS SULFURIC ACID FROM FIRST PRINCIPLES SIMULATIONS WITH HYBRID FUNCTIONALS 97
7.1 Introduction 97
7.2 Computational Methods 99
7.3 Results and Discussion 100
7.3.1 Structural Properties 100
7.3.2 Dissociation of Sulfuric Acid 101
7.3.3 Electronic Properties 105
7.4 Conclusions 107

8 CONCLUSIONS 109

REFERENCES 112

LIST OF FIGURES

3.1 Schematic representation of the refinement approach described in Section 3.2 for a one-dimensional periodic system with periodicity L. $w_n(r)$ is nth Maximally localized Wannier Function (MLWF) localized in real space at r_n^0 and $X_n(r)$ is the corresponding saw-tooth function (see text). 24

3.2 Dipole moment (a), spread of lone pair and bond pair MLWFs (b) and isotropic polarizability (c) of a water molecule computed at different cell size using Berry phase (black), MLWF (red), refined MLWF (blue), DFPT (green) and all electron (orange) approaches. The Berry phase, MLWF (Maximally localized Wannier Function) and refined MLWF methods are described in Section 3.2.1, and the DFPT (Density Functional Perturbation Theory) approach is described in Section 3.1. d-aug-cc-pVQZ denotes the basis set [1] adopted in All Electron (AE) calculations. Four MLWFs are associated to each water molecule (2 lone pairs and 2 bond pairs) constructed as linear combinations of occupied KS eigenstates. 28

3.3 Diagonal elements of the traceless quadrupole moment (a) and the quadrupole polarizability (b) of a water molecule computed using refined MLWF method (see Section 3.2.1) at different cell sizes. Here X, Y and Z denote the axis in the molecular plane, the axis perpendicular to the molecular plane and the dipole axis of the water molecule, respectively. 29

3.4 (a) Distribution of dipole moments of water molecules computed using the MLWF method and PBE functional (black), the refined MLWF method and PBE functional (red) and the refined MLWF and PBE0 functional (blue) for 200 snapshots extracted from an FPMD simulation using PBE functional and 64 molecule liquid water sample [2]. (b) Distribution of traceless quadrupole moments of water molecules computed using the refined MLWF method in combination with PBE (black) and PBE0 functionals (red) for the same snapshots. The blue dots and vertical lines indicate values computed for an isolated water molecule using the refined MLWF method and PBE functional. 31

4.1 (a) Calculated isotropic (upper panel) and anisotropic (lower panel) Raman spectra (black lines) compared with several experiments. Red, blue and green curves are experimental Raman spectra of heavy water measured at 283K [3], 293K [4] and 303K [3] respectively. Orange and purple curves show experimental Raman spectra of hydrogenated water at 278K and 308K [5]. (b) Calculated isotropic (upper panel) and anisotropic (lower panel) Raman spectra (black lines) decomposed into intra- (red lines) and inter- (green lines) molecular contributions. . . 46

4.2 (a) The spectrum (black line) calculated from the spread of the MLWFs (see text) is decomposed into intra- (red line) and inter- (green line) molecular contributions. The vibrational density of states of the relative speed between oxygen atoms of HB water molecules is also shown (orange line). (b) Distribution of different components of the molecular polarizabilities of water molecules obtained in the simulation, compared with the polarizability of an isolated water molecule in different directions (vertical bars): black, red and blue lines represent polarizabilities along the dipole axis, the axis perpendicular to and the axis within the molecular plane, respectively. 48
4.3 (a) Isotropic (upper panel) and anisotropic (lower panel) Raman spectra (scaled; black lines) compared with the spectra computed from molecular polarizabilities $\boldsymbol{\alpha}_i$ (blue lines) and their intra- (red lines) and inter- (green lines) molecular contributions (see text). (b) Raman spectra arising from the $\boldsymbol{\alpha}_i$ (black lines) and $\boldsymbol{\alpha}'_i$ (polarizability of a single, isolated water molecule at the geometry of the molecule in the liquid (see text); red lines) 51

5.1 Representative geometry of of a Sum Frequency Generation (SFG) experiment (left panel) where an interface between two media α and β is illuminated by two light beams of frequency ω_1 and ω_2 (the corresponding wavevectors $\mathbf{k}$ are also indicated); I, R and T indicate incident, reflected and transmitted light, respectively. The emitted SF light has frequency $\omega_s = \omega_1 + \omega_2$ (right panel). . . 56
5.2 Computed z component and x and y components of the high-frequency dielectric constant tensor ϵ_∞ as a function of the distance z in the direction perpendicular to the surface (a); local frequency-dependent dielectric constant, $\epsilon_z(\omega)$, averaged over the first (1st), second (2nd) and third (3rd) BLs (b) in one of the ice slab models studies in this work. The latter is compared with experimental measurement of $\epsilon(\omega)$ for bulk ice I_h [6]. 67
5.3 Imaginary part of the $\boldsymbol{\chi}^{\mathrm{ID}}_{\mathrm{eff}}$ contribution to SFG spectra [the *ssp* (left) and *ppp* (right) polarization combinations] computed for one proton-disordered surface with (black) and without (red and blue) correction for the long-range interaction between slabs. We used the optical geometry (incident angles) of Ref. [7] for black and red curves and the optical geometry of Ref. [8] for the blue curve. 72
5.4 Imaginary part of the computed $\boldsymbol{\chi}^{\mathrm{ID}}_{\mathrm{eff}}$ contribution to the SFG spectra of *ssp* and *ppp* polarization combinations, as obtained from AIMD (black) and finite difference (FD, blue) calculations. The upper two panels show spectra computed for two proton disordered surfaces; the lower two panels for two proton ordered surfaces. All AIMD and FD spectra were obtained by including three surface BLs in the calculation and setting the Fresnel coefficients to the identity tensor. . . . 74
5.5 The computed IQB contribution $|\chi^{\mathrm{IQB}}_{\mathrm{eff}}(\omega_2)|^2$ of one bulk ice sample. The black and red curves were computed using (0,0,0) and (5,5,5), in a.u., as the origin (see text). 75

5.6 Real (top panels) and imaginary (middle panels) part of $\chi_{\rm eff}^{\rm ID}$ spectra computed for the proton ordered (black line) and proton disordered (red line) ice I_h basal surfaces. The corresponding $\chi_{\rm eff}^{\rm ID} + \chi_{\rm eff}^{\rm IQB}$ spectra for the two types surfaces (blue and green dashed lines) are shown as well. Imaginary part of the $\chi_{\rm eff}^{\rm IQB}$, $\chi_{\rm eff}^{\rm BQ}$ and $\chi_{\rm eff}^{\rm IQI}$ are shown in the bottom panels. The spectra for *ssp* and *ppp* polarization combinations are shown in left and right panels, respectively. The computed spectra were red-shifted by 100 cm^{-1} to align them with experiments at 3100 cm^{-1}. The discrepancy in peak positions is ascribed to a combined effect of the neglect of quantum effects [9] and the use of the PBE functional [10]. Anharmonic effects were not found to be significant for this system (Fig. 5.4). 76

5.7 SFG intensities for *ssp* (left panels) and *ppp* (right panels) polarization combinations for ice I_h basal surfaces. We report measured (at 173 K, Ref. [7], and at 100 K, Ref. [8]; top panels) and computed results (as averages for proton disordered surface models in the middle panels, and proton ordered surface models in the bottom panels). The spectra were computed by including the ID contribution originating from the top one (green), two (blue) or three (black) surface BLs (see text). Violet curves show only the ID contribution from three surface BLs obtained by setting $\epsilon(\omega, z) = 1$ in the calculation. The total spectra $|\chi_{\rm eff}^{(2)}|^2$, including ID contribution from three surface BLs as well as IQI, IQB and BQ contributions, are shown by the red curve. The computed spectra were red-shifted by 100 cm^{-1} to align them with experiments at 3100 cm^{-1}. 77

6.1 Internally (12A) and surface (12B) solvated configurations of the SO_4^{2-} $(H_2O)_{12}$ cluster. Oxygen, hydrogen and sulfur are represented in red, white and yellow, respectively. Red dotted lines indicate hydrogen bonds. 85

6.2 Infrared spectra of 12-water sulfate clusters $SO_4^{2-}(H_2O)_{12}$. The uppermost panel shows the experimental IRMPD spectra[11]. The middle and lower panels show computed spectra for the 12A and 12B clusters, respectively (see Fig. 6.1), obtained using finite differences (FD) and molecular dynamics (MD) simulations, with either the PBE or PBE0 functional. 86

6.3 Internally (13A and 13C) and surface (13B) solvated configurations of the SO_4^{2-} $(H_2O)_{13}$ cluster. Oxygen, hydrogen and sulfur are represented in red, white and yellow, respectively. Red dotted lines indicate hydrogen bonds. 88

6.4 Distance between the sulfur atom and three oxygen atoms (denoted as O3, O10 and O13) belonging to solvating water molecules, during a ∼30ps *ab initio* molecular dynamics simulation at 100K. The starting (13A) and final (13C) configurations are shown as insets. The bond lengths between S and oxygen atoms 3, 10 and 13 are denoted by black, red and blue lines respectively. A jump of O10 from the first to the second solvation shell can be observed from the change in sulfur-oxygen distance: the affected hydrogen bonds are highlighted in orange in both the starting and final configurations. 89

6.5 Infrared spectra of 13-water sulfate anion clusters $SO_4^{2-}(H_2O)_{13}$. The uppermost panel shows the experimental IRMPD spectra[11]. The remaining panels show computed spectra for the 13A, 13B and 13C clusters, respectively (see Fig. 6.3), obtained using finite differences (FD) and molecular dynamics (MD) simulations, with either the PBE or PBE0 functional. 90
6.6 Structural changes observed during the first 2 ps of an *ab initio* MD simulation of a 13-water hydrated sulfate anion cluster (see text). 91
6.7 Charge density differences between the 12 water hydrated SO_4^{2-} and SO_4^- clusters in the 12 A (two upper panels) and 12B (two lower panels) configurations (see Fig. 6.1), obtained using the PBE (left panels) and PBE0 (right panels) functionals. Note the difference in charge localization obtained with semilocal and hybrid functionals. The same value of charge density is plotted in all cases. 93
6.8 Electronic density of states (EDOS) of the 12A and 12B (Fig. 6.1) computed using the PBE (upper panel) and PBE0 (lower panel) exchange-correlation functionals. Green and black arrows indicate the first three orbitals belonging to sulfate ion, water and sulfate ion respectively (see text). 94

7.1 Oxygen-oxygen (a) , oxygen-hydrogen (b) and sulfur-oxygen (c) radial distribution functions (RDF), $g_{O-O}(r)$, $g_{O-H}(r)$ and $g_{S-O}(r)$; O_W denotes oxygen atoms belonging to the water molecules in the liquid. Red and black solid curves denote our results for a 0.87 mol/L sulfuric acid solution obtained with simulations at 376 K using the PBE0 and PBE functionals, respectively. We compare our RDFs with those reported in Ref. [12] (brown curve). We also report the $g_{O-O}(r)$ of pure heavy water from a PBE simulation at 378 K [2] (blue curve) and the experimental water correlation function obtained from recent X-ray diffraction experiments at 295 K [13] (dashed green curve). 102
7.2 Computed oxygen-hydrogen (O_S-H) (a) and oxygen-water oxygen (O_S-O_W) radial distribution functions (RDFs), where O_S and O_W denote oxygen atoms belonging to sulfuric acid and to water molecules, respectively. Red and black solid curves denote our results for a 0.87 mol/L solution obtained with the PBE0 and PBE functionals, respectively. We compare our O_S-H RDF with that reported in Ref. [12] (brown curve). 103
7.3 Distances (top panel) and single particle energies (middle and bottom panels) of a sulfuric acid aqueous solution (0.87 mol/L) computed from trajectories obtained at the PBE (left panels) and PBE0 (right panels) level of theory. In the upper panel the red and blue curves show the distance between the sulfur atom and the oxygen atom belonging to its first and second hydronium ion neighbor, respectively (the blue curve is zero when only one hydronium ion is present in the simulation). The middle and bottom panels show Kohn-Sham eigenvalues of the highest eight occupied orbitals (black circles), with states localized on the anion (HSO_4^- or SO_4^{2-}) shown in red. The electronic structure was computed using the PBE (middle panels) and PBE0 (bottom panels) functionals on samples extracted from PBE (left) and PBE0 (right) trajectories. The orange and green boxes highlight several typical portions of our simulations where the HOMO of the anion is significantly below or above the water VBM, respectively. 104

7.4 Relative positions of the highest occupied molecular orbital (HOMO) of the HSO_4^- (green) and SO_4^{2-} (blue) anions, with respect to the valence band maximum (VBM) of water (black), as obtained in our simulations with the PBE and PBE0 functional. 105

LIST OF TABLES

3.1 Diagonal elements of the traceless quadrupole moment, in Buckingham, of an isolated water molecule 30

3.2 High-frequency dielectric constant, ϵ_∞, of a 64-molecule liquid water sample extracted from FPMD trajectories obtained in the NVE ensemble and the PBE functional, with the Qbox code. Calculations of ϵ_∞ were carried out with both GGA (PBE) and hybrid (PBE0) functionals. The Berry phase and refined MLWF methods are described in Section 3.2.1, and the DFPT approach is described in Section 3.1. 30

3.3 Average diagonal elements of the traceless quadrupole moment tensor of water molecules in liquid computed using the refined MLWF method and the PBE or PBE0 functionals and 200 snapshots extracted from FPMD simulations using PBE [2] and PBE0 [14] functionals. 32

6.1 Energy difference (eV) between cluster geometries represented in Fig. 6.1 and 6.3 for the 12 and 13 water hydrated sulfate dianion clusters. Calculations were carried out using semilocal (PBE) and hybrid (PBE0) functionals, and adding the harmonic zero point energy contribution (ZPE) computed from finite difference calculations. 87

6.2 Average bond lengths (Å) and average bond angles (°) for water-ion and water-water hydrogen bonds, computed over 20 ps *ab initio* MD trajectories. No results are shown for 13A,which was found to be unstable in our simulations at finite temperature. 88

6.3 Calculated vertical ionization potential (eV) of 12A and 12B sulfate dianion clusters (see Fig. 6.1), obtained with semilocal (PBE) and hybrid (PBE0) functionals. 92

ACKNOWLEDGMENTS

I am deeply indebted to my advisor Professor Giulia Galli. She lead me to an exciting and promising research area, and motivated and encouraged me through this five-year journey. What I learned from her is not only problem solving skills in research, but also diligent and attentive working ethics. She worked very hard to maintain a highly talented and friendly research group (Galli group, fomerly known as Angstrom group) and showed genuine care for each member of the group. Without her, This dissertation would be impossible.

I am sincerely grateful to Dr Leonardo Spanu, who was a research scientist in Angstrom group and my mentor when I joined the group. His encouragement was an important reason that I chose the this research direction. He was very patient and thorough when he taught me the basics in scientific research and scientific programming. These skills turned out to be a very important asset for me.

I am very thankful to Professor François Gygi, who is very knowledgeable in the technical issues in scientific computing. We have had many productive discussions on my projects.

I would like to thank everyone in Galli group and Professor Gygi's research group for numerous inspiring discussions, as well as friendly support on various aspects. I want to thank my colleagues and friends from UC Davis and the University of Chicago.

ABSTRACT

Vibrational spectroscopy is an ideal tool to probe the complex structure of ice, water and other aqueous systems. However, the interpretation of experimental spectra is usually not straightforward, due to complex spectral features associated with different bonding configurations present in these systems. Therefore, accurate theoretical predictions are required to assign spectral signatures to specific structural properties and hence to fully exploit the potential of vibrational spectroscopies. My dissertation focused on the development and applications of first-principles electronic structure methods for the simulation of vibrational spectra of water and aqueous systems, as well as of their basic electronic properties.

In particular, I focused on the calculation of response properties of aqueous systems in the presence of external electric fields, including the computation of dipole and quadrupole moments and polarizabilities, which were then used to simulate vibrational spectra, e.g. Raman and sum frequency generation (SFG) spectra. I developed linear response and finite field methods based on electronic structure calculations within density functional theory (DFT), which were applied to accurate and efficient evaluations of the electric field response, and coupled to large-scale first-principles electronic structure and molecular dynamics (MD) simulations. Our implementation enabled on-the-fly calculations of polarizabilities in first-principles MD (FPMD) simulations, which are necessary to compute vibrational spectra using time correlation functions (TCF) formulations. In addition, in this dissertation I provide the first *ab initio* implementation of SFG calculations, inclusive of quadrupole contributions as well as of electric field gradients at the interface.

I present the first calculation of the Raman spectra of liquid water using FPMD simulations. Interesting signatures were found in the low frequency region of the spectra, indicating intermolecular charge fluctuations that accompany hydrogen bond stretching vibrations. Furthermore I applied the newly developed method for the calculation of SFG spectra, a surface specific spectroscopic probe, to the investigation of the ice I_h basal surfaces. Note that all the methods developed here, although only applied to aqueous systems, are of general

applicability to semiconductors and insulators.

Finally I present investigations of the vibrational and electronic properties of aqueous sulfuric acid systems, including sulfate-water clusters and sulfuric acid solutions. Our results on the energy alignment between sulfate and water states bear important implications on the relative reactivities of ions and water in electrochemical environment used in water splitting reactions.

CHAPTER 1

INTRODUCTION

Despite a long history in the study of water and other aqueous solutions, several of their properties are not yet well understood because of the subtle and complex interactions between water molecules and between water and solutes [15]. These interactions are determined by the interplay of hydrogen bonding [16], nuclear quantum effects [17–19], van der Waals interactions [20, 21], charge transfer effects [22–24] and overall many-body electronic and nuclear interactions.

Vibrational spectroscopies, notably infrared and Raman, have been widely used to investigate water and aqueous solutions and to unravel their complex structure through analyses of vibrational signatures [16, 25, 26]. However infrared and Raman signals are not necessarily surface sensitive and hence may not be useful in understanding and determining the surface and interface properties of aqueous systems. Nonlinear sum-frequency generation (SFG) vibrational spectroscopy [27] is instead surface sensitive and has recently enabled a series of studies of aqueous surfaces [7, 28–30]. However, due to the complexity of vibrational spectral features, in-depth theoretical studies and simulations are necessary to help interpret SFG spectra in particular, and in general of IR and Raman signals [26].

First-principles molecular dynamics (FPMD) simulations provide a straightforward way to simulate vibrational spectra of disordered systems, e.g. liquid water, through the computation of time correlation function (TCF) [31]. The FPMD approach combines the molecular dynamics (MD) technique to obtain atomic trajectories with density functional theory (DFT) calculations of forces acting on atoms at each MD step [32]. For aqueous systems, this approach has proved very promising in tackling a wide range of problems, including covalent bond breaking and formation [33, 34], infrared vibrational excitations [23, 35] and dielectric response [36, 37] in water, as well as water at high temperature and high pressure [37–39].

Most FPMD carried out in the last two decades were based on the calculation of ground state properties of isolated systems, e.g. in the absence of external fields. FPMD simulations

for Raman and SFG spectra, which require electric field response calculations, have been sparse so far [40–42], due to the lack of accurate and efficient *ab initio* methods to predict the response to electric fields of complex, disordered systems. It is therefore highly desirable to develop such methods and seamlessly integrate them with FPMD simulation software, so as to efficiently compute Raman and SFG spectra from first principles.

Two formulations may be used to simulate condensed systems in the presence of the electric field from first principles: linear response and finite field methods. The linear response approach based on density functional perturbation theory (DFPT) [43] treats the electric field as a perturbation and response properties, e.g. polarizabilities, are computed by solving the Sternheimer equation [44]. DFPT is general and may in principle be applied using any energy functional. However in practice its applicability is limited to local or semilocal energy functionals, due to algorithmic and technical difficulties in evaluating functional derivatives of exchange and correlation potentials in the case of orbital-dependent density functional approximations, i.e. hybrid functionals [43]. Within the finite field approach, the presence of an external electric field is treated in a non-perturbative manner, by introducing an electric enthalpy functional [45], which is easily defined irrespective of the exchange correlation functional used, whether semilocal or hybrid. However, the results of this approach suffer from slow convergence with respect to the supercell size used in first-principles calculations, due to computational difficulties involved in evaluating the polarization of condensed systems.

In this dissertation, I will present efficient formulations of DFPT and finite field approaches in conjunction with FPMD simulations, and their implementation in the massively parallel, open source Qbox code [46, 47]. To improve the accuracy of the finite field approach, a simple and accurate method for polarization calculations was developed, which is based on maximally localized Wannier functions [48] and a so called refinement procedure originally proposed by Stengel and Spaldin [49]. The method was also generalized to enable the accurate evaluation of electric multipole moments, e.g. quadrupole moments, of molecules in the condensed phase.

The development and implementation DFPT and finite field methods have enabled the first FPMD simulation of Raman spectra of liquid water, presented in this dissertation, where I observed signatures of intermolecular charge fluctuations in the liquid, at low frequencies. Building upon these methods, I then formulated a robust first-principles framework for the simulation of SFG spectra, and performed the first *ab initio* simulation of SFG spectra of ice surfaces. The capability of computing multipole moments in the condensed phase has made possible the evaluation of higher-order quadruple contributions to the spectra, which were not previously predicted from first principles. Our results provided guidance for the interpretation of experimental spectra of surfaces and interfaces. In addition, our formulation of FPMD simulations in an external field was used to study the dielectric properties of water under high pressure and temperature, providing results with important implications for carbon transport in the Earth's mantle [37].

In addition to the vibrational properties of pure water, this dissertation presents studies of the electronic properties of aqueous solutions, and of their implications for understanding the reactivity of ions and water, whose properties are of interest in most chemical and electrochemical reactions [50, 51]. I carried out FPMD simulations of sulfuric acid solutions, using highly accurate hybrid functionals. In particular, I focused on the energy alignment between sulfate and water states, and on its implication for the mechanism of oxygen evolution reactions in water splitting experiments [50, 51].

The rest of this dissertation is organized as follows: in Chapter 2, I summarize the basic concepts of DFT and FPMD techniques and their application to the study of ground state properties; the development and implementation of linear response and finite field methods for *ab initio* calculations in the presence of electric fields are described in Chapter 3; results on Raman spectra of liquid water are presented in Chapter 4; in Chapter 5, I present the theoretical framework for SFG spectra calculations developed in this dissertation, as well as its application to ice surfaces; results on structural, vibrational and electronic properties of small sulfate water clusters and aqueous sulfuric acid solutions are presented in Chapter 6

and Chapter 7, respectively; Chapter 8 summarizes and concludes the dissertation.

CHAPTER 2
FIRST PRINCIPLES CALCULATIONS OF GROUND STATE PROPERTIES

In this chapter, I will introduce the basic principles of density functional theory (DFT) and first principles molecular dynamics (FPMD).

2.1 Density functional theory

Density functional theory (DFT) provides an exact description of the properties of an interacting many-electron system through the Hohenberg-Kohn (HK) theorems [52]. The latter states that ground and excited state properties of a system of interacting electrons under the influence of an external potential can be derived from the ground state charge density of the system, hence eliminating the need to evaluate the many-body wavefunction; the energy of the interacting electronic system is a unique functional of its charge density. However, the explicit form of such density functional is unknown, and approximate functionals are used in practical implementations of DFT. In addition the minimization of the functional is recast into the solution of a set of Schrödinger like equations for single particle orbitals using the Kohn-Sham formulation of DFT [53], on which I will focus in the rest of this chapter.

2.1.1 The Schrödinger Equation

According to quantum mechanics, the ground state energy and wavefunction of an electronic system is obtained by solving the time independent Schrödinger equation:

$$\hat{H}\Psi = \mathcal{E}\Psi, \tag{2.1}$$

where $\hat{H}$ is the Hamiltonian operator, $\mathcal{E}$ is the electronic energy and Ψ is the wavefunction. The Hamiltonian depends on the electronic and nuclear coordinates and spin variables. In

most problems one adopts the so-called Born-Oppenheimer (BO) approximation to separate the nuclear degrees of freedom from the electronic ones, given that the electronic mass is much smaller than that of the nuclei and the nuclei may be considered at rest when one computes the electronic wavefunction. Hereafter we adopt the BO approximation and we only consider the electronic coordinate in Eq. 2.1.

For systems consisting of light elements for which relativistic effects can be neglected, the Hamiltonian $\hat{H}$ can be written as:

$$\hat{H} = \hat{T} + \hat{V} + \hat{U}, \tag{2.2}$$

where $\hat{T}$ is the kinetic energy of the electrons, $\hat{V}$ is the external potential exerted by the nuclei on the electrons and $\hat{U}$ is the electron-electron interaction energy. The explicit expressions of these three operators are:

$$\hat{T} = -\frac{1}{2}\sum_{n}^{N_E}\nabla^2, \tag{2.3}$$

$$\hat{V} = -\sum_{n=1}^{N_E}\sum_{i=1}^{N_N}\frac{Z_i}{|\mathbf{r}_n - \mathbf{R}_i|}, \tag{2.4}$$

$$\hat{U} = \frac{1}{2}\sum_{n<m}\frac{1}{|\mathbf{r}_n - \mathbf{r}_m|}, \tag{2.5}$$

where N_E and N_N are the number of electrons and nuclei in the system, ∇^2 is the Laplacian operator, $\mathbf{r}_n$ is the position of the nth electron and $\mathbf{R}_i$ and Z_i are the position and charge of the ith nucleus, respectively. Solving Eq. 2.1 with $\hat{H}$ given by Eq. 2.2 is prohibitively difficult for systems with more than two electrons [54].

2.1.2 Density Functional Theory and the Kohn-Sham Ansatz

Density functional theory seeks to find the ground state of interacting electrons by minimizing an energy functional instead of solving Eq. 2.1. It is based on two theorems put forward in

a seminal paper by Hohenberg and Kohn [52], which connect all the properties of a many-electron system to its electron density $\rho(\mathbf{r})$. The HK theorems state:

1. The external potential $v_{ext}(\mathbf{r})$ acting on an interacting electronic gas is a unique functional, within a constant, of the ground state electron density $\rho(\mathbf{r})$. Thus $\rho(\mathbf{r})$ uniquely determines all properties of the system, including the many-electron wavefunction.

2. The total energy of an interacting electronic gas is a unique functional of the electron density. The ground state energy is the variational global minimum of this unique functional $E[\rho]$.

Practical implementations of the HK theorems use the Kohn-Sham Ansatz [53], where a system of N_E interacting electrons is mapped onto a fictitious system of non-interacting electrons with the same electron density. The wavefunction of this non-interacting system is a Slater determinant constructed from the orbitals of the individual non-interacting electrons $\{\psi_n\}$, which may be obtained by self-consistently solving a set of Schrödinger like Equations called the Kohn-Sham equations.

The total electron density of the non-interacting system $\rho(\mathbf{r})$ is:

$$\rho(\mathbf{r}) = \sum_{n}^{N_E} \psi_n^*(\mathbf{r})\psi_n(\mathbf{r}). \tag{2.6}$$

The total energy functional in terms of the electron density is:

$$E_{KS}[\rho] = T[\rho] + V_{ext}[\rho] + E_H[\rho] + E_{xc}[\rho] + E_N, \tag{2.7}$$

where the first four terms on the rhs are the kinetic, external potential, Hartree and exchange-correlation (XC) energy functionals. The explicit form of the kinetic energy functional is

$$T[\rho] = -\frac{1}{2}\sum_{n}^{N_E} \int d\mathbf{r}\psi_n^*(\mathbf{r})\nabla^2\psi_n(\mathbf{r}), \tag{2.8}$$

which is an implicit functional of the electron density $\rho(\mathbf{r})$ since $\psi_n(\mathbf{r})$ is a functional of $\rho(\mathbf{r})$, according to the first HK theorem. The external potential energy functional is:

$$V_{ext}[\rho] = -\sum_{i=1}^{N_N} \int d\mathbf{r}\rho(\mathbf{r})\frac{Z_i}{|\mathbf{r}-\mathbf{R}_i|}. \tag{2.9}$$

The Hartree energy functional accounts for the Coulomb interaction between electrons:

$$E_H[\rho] = \frac{1}{2}\int d\mathbf{r}d\mathbf{r}'\frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}. \tag{2.10}$$

The only unknown term in Eq. 2.7 is the XC functional $E_{xc}[\rho]$, which will be discussed in the next section. The term E_N is the Coulomb interaction between nuclei:

$$E_N = \frac{1}{2}\sum_{i<j}\frac{Z_i Z_j}{|\mathbf{R}_i - \mathbf{R}_j|} \tag{2.11}$$

Using the second HK theorem, the ground state energy of the interacting electrons can be obtained by minimizing the energy functional $E_{KS}[\rho]$ with respect to ρ. This requires the functional derivatives of $E_{KS}[\rho]$ with respect to ψ_n^* to be zero.

$$\begin{aligned}\frac{\delta\Big(E_{KS}[\rho] - \varepsilon_n\int\psi_n^*\psi_n d\mathbf{r}\Big)}{\delta\psi_n^*} &= \left(\frac{\delta T[\rho]}{\delta\rho} + \frac{\delta V_{ext}[\rho]}{\delta\rho} + \frac{\delta E_H[\rho]}{\delta\rho} + \frac{\delta E_{xc}[\rho]}{\delta\rho} - \epsilon_n\right)\frac{\delta\rho}{\delta\psi_n^*} \\ &= \left(-\frac{1}{2}\nabla^2 + v_{ext} + v_H + v_{xc} - \epsilon_n\right)\psi_n \\ &= 0.\end{aligned} \tag{2.12}$$

Here ε_n is a Lagrange multiplier used to ensure normalization of the wavefunctions and v_{ext}, v_H and v_{xc} are the external, Hartree and XC potentials, respectively. The minimization of $E_{KS}[\rho]$ can be recast into the solution of Schrödinger-like equations for non-interacting electrons:

$$\hat{H}_{KS}\psi_n = \varepsilon_n\psi_n, \tag{2.13}$$

where

$$\hat{H}_{KS} = -\frac{1}{2}\nabla^2 + v_{ext} + v_H + v_{xc}. \tag{2.14}$$

Eq. 2.13 is to be solved self-consistently since v_{ext}, v_H and v_{xc} are functionals of ρ. The resulting ψ_n and ε_n are KS eigenfunctions and KS eigenvalues, respectively. Here all the many-body electron interactions are included in the XC potential v_{xc} within a mean-field approach.

2.1.3 Approximations to the Exchange-Correlation Energy Functional

As mentioned above, the exact form of the XC functional and hence of its functional derivatives are unknown. Approximate forms of this functional are often used in practical calculations. One often separates the exchange and correlation functionals:

$$E_{xc} = E_x + E_c. \tag{2.15}$$

Within the local density approximation (LDA) [55, 56], the XC functional is expressed in terms of the exchange and correlation energy of the homogeneous electron gas:

$$E_{xc}^{LDA} = \int d\mathbf{r}\rho(\mathbf{r})\epsilon_{xc}^{LDA}(\rho) \tag{2.16}$$

where $\epsilon_{xc}^{LDA}(\rho)$ is the exchange-correlation energy per electron of a homogeneous electron gas with density ρ; the correlation part of Eq. 2.15 is obtained by fitting the results of highly accurate quantum Monte Carlo simulations of a homogeneous electron gas for different values of ρ [55]. More sophisticated approximations, such as the generalized gradient approximation (GGA), include the gradient of the electron density in the XC functional:

$$E_{xc}^{GGA} = \int d\mathbf{r}\rho(\mathbf{r})\epsilon_{xc}^{GGA}\big(\rho(\mathbf{r}), \nabla\rho(\mathbf{r})\big). \tag{2.17}$$

Commonly used GGA functionals, sometimes referred to as semilocal functionals, include BLYP [57, 58] and PBE [59, 60]. These two functionals have been widely used, e.g. for simulations of water and aqueous solutions and provide a reasonable description of their structural and vibrational properties [23, 61–63]. However at ambient conditions, GGA functionals often predict an overstructured liquid and underestimate its vibrational frequencies [64, 65]. In addition, GGA functionals are often not sufficiently accurate to treat excitation properties of materials and molecules [66, 67]. These disadvantages of the GGA functionals arise in part from the so-called delocalization error [66, 67], i.e. from the incorrect partial cancellation of the Hartree and exchange energy for each electron in the system, which in turn leads to some bonds, e.g. the OH bond in water, being characterized by an excessively delocalized charge density.

One way to improve the accuracy of GGA functionals is to define an exchange correlation functional as a linear combination of E_{xc}^{GGA} and the Hartree-Fock (HF) exchange, which is sometimes referred to as exact exchange (EXX). The HF approximation has been widely used in quantum chemistry community and the HF exchange energy is given by:

$$E_x^{HF} = -\frac{1}{2}\sum_{m\neq n}\int d\mathbf{r}d\mathbf{r}'\frac{\psi_n^*(\mathbf{r})\psi_m(\mathbf{r})\psi_m^*(\mathbf{r}')\psi_n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}. \tag{2.18}$$

Among the most commonly used hybrid functionals, the PBE0 functional [68] is the one used in this dissertation. It is defined as:

$$E_{xc}^{PBE0} = 0.25E_x^{HF} + 0.75E_x^{PBE} + E_c^{PBE}, \tag{2.19}$$

where the parameters 0.25 and 0.75 in Eq. 2.18 are not empirical and they are derived by sum rules of the homogeneous electrons gas. The PBE0 functional has been shown to considerably improve upon PBE in the description of the structural and vibrational properties of several systems, in particular water [10, 65, 69] and ice [14, 70].

2.1.4 Planewave Pseudopotential Scheme

Numerical solutions of the KS equations are obtained by expanding the KS orbitals into appropriate basis sets. For condensed phase systems, periodic boundary conditions and planewave basis sets have been a popular and computationally efficient choice in the literature of the last three decades [44]. KS orbitals that satisfy the Bloch theorem [44] can be expanded in planewaves:

$$\psi_n(\mathbf{r}) = \sum_{\mathbf{k}} \sum_{\mathbf{G}} c_{n,\mathbf{k},\mathbf{G}} e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}}, \tag{2.20}$$

where $\mathbf{k}$ is a k-point vector in the first Brillouin zone of the unit cell or supercell used in the calculation and $\mathbf{G}$ is a reciprocal lattice vector. For disordered solids and liquids, one often considers only the Γ point ($\mathbf{k} = \{0, 0, 0\}$). In this dissertation I will focus on methods using only the Γ point because our primary interest is in disordered systems. The number of $\mathbf{G}$ vectors included in Eq. 2.20 is usually expressed in terms of a cutoff E_{cut}, where $\mathbf{G}$ vectors with $\frac{1}{2}|\mathbf{G}|^2 < E_{cut}$ are included in the summation.

Describing KS orbitals using planewaves may result in slow convergence with respect to basis set size if core electrons are included, since $\psi_n(\mathbf{r})$ is rapidly varying near atomic nuclei. Therefore the use of planewaves is always acompanied by that of pseduopotentials. A psedupotential [44] is defined for each atomic species and describes the effective interaction between chemically inert core electrons and valence electrons, i.e. those electrons participating in chemical bonds. The partition between core and valence electrons depends on the elements and, to some extent, on the problem. Replacing the Coulomb potential in Eq. 2.9 by sum of pseudopotentials and solving for pseduopwavefunction $\psi_n(\mathbf{r})$ allows one to greatly reduce the number of planewaves needed to represent KS orbitals. In our studies we choose to use norm-conserving pseudopotentials of the Hamann-Schlüter-Chiang-Vanderbilt (HSCV) type [71, 72], which were previously tested in the literature [10, 73] in the case of water and ice.

The combination of planewave basis sets and psudopotentials permits effective use of

Fast Fourier transform (FFT) algorithms that enable efficient transformation between the real space wavefunction $\psi_n(\mathbf{r})$ and its reciprocal space Fourier components $c_{n,\mathbf{k},\mathbf{G}}$. The efficiency of this approach comes from the fact that KS Hamiltonian matrix $\hat{H}_{KS}$ can be written as the sum of two sparse matrices, one corresponding to the kinetic energy and the other to the potential energy operator, respectively (see Eq. 2.14). The kinetic energy and potential energy matrices are diagonal in Fourier and real spaces, respectively, and are computed efficiently in the these spaces [74].

The LDA and GGA XC functionals are efficiently computed in the real space. The calculation of hybrid XC functionals is much more demanding than that of LDA or GGA functionals and cannot be straightforwardly carried out in real space [10]. In particular, the evaluation of E_x^{HF} requires the computation of a large number ($\frac{1}{2}N^2$) of integrals over KS orbital pairs (see Eq. 2.18). Its computational cost may be greatly reduced by using a recursive subspace bisection algorithm, which localizes KS orbitals onto different domains of the supercell and thus greatly decrease the number of required integrals [75, 76]. The recursive subspace bisection algorithm has made possible a number of accurate simulations of aqueous systems using hybrid functionals [39, 69, 77].

2.1.5 First Principles Molecular Dynamics

Molecular dynamics (MD) is a simulation technique to model the time evolution of a system consisting of interacting atoms and molecules. The position and velocity of a given atom at a given time is obtained by numerical integration of the Newton's equation of motion, e.g. using the Verlet algorithm [78]:

$$-\frac{\partial E\big(\{\mathbf{R}_j\}\big)}{\partial \mathbf{R}_i} = m_i \ddot{\mathbf{R}}_i, \tag{2.21}$$

where m_i and $\ddot{\mathbf{R}}_i$ are the mass and the acceleration of the ith nuclei, respectively. $E\big(\{\mathbf{R}_j\}\big)$ is the total energy of the system.

Based on the method used to compute $E\big(\{\mathbf{R}_j\}\big)$, the MD technique can be categorized into classical and first principles MD (FPMD) schemes. The classical MD scheme uses empirical potentials to approximate the total energy of the system $E\big(\{\mathbf{R}_j\}\big)$, resulting in low computational costs (compared to FPMD), but often sacrificing accuracy and transferability from different bonding configurations. On the other hand, the FPMD scheme provides an accurate and non-empirical description of the potential energy surface $E\big(\{\mathbf{R}_j\}\big)$ [32], and it is the method of choice in this dissertation. In an FPMD simulation, one computes $E_{KS}\big(\{\mathbf{R}_j\}\big)$ by solving the KS equations (Eq. 2.13), at each time step. Evaluation of the force $-\partial E_{KS}\big(\{\mathbf{R}_j\}\big)/\partial\mathbf{R}_i$ is carried out efficiently using the Hellmann-Feynman theorem [79]. Compared to classical MD, the accuracy of FPMD is superior for a wide range of systems and configurations, including those with bond breaking and formation [34], charge transfer [23] and at high pressure and temperature conditions [39]. The disadvantage of FPMD is the high computational cost which limits its applicability to systems much smaller (one to several order of magnitudes) than those tractable with classical MD and to shorter simulation times (typically in tens of ps range, not exceeding ~ 100 ps).

CHAPTER 3

FIRST-PRINCIPLES CALCULATIONS IN THE PRESENCE OF EXTERNAL ELECTRIC FIELDS

In this chapter, I focus on first principles methods to compute the properties of ordered and disordered systems in the presence of external electric fields. Two types of methods will be discussed: linear response and finite field methods.

3.1 Density functional perturbation theory for Homogeneous Electric Field Perturbations

As a linear response method, density functional perturbation theory (DFPT) is a powerful tool for studying the response of a system to external perturbations. The detailed formulation of DFPT is described by Baroni *et al.* [43]. Here we only discuss DFPT for homogeneous electric field perturbations.

3.1.1 Linear Response within DFPT

The study of a system's response to homogeneous electric field perturbations is one of the main focuses of this dissertation, as such response is needed to compute vibrational spectra such as Raman and sum frequency generation (SFG) spectra. After summarizing density functional perturbation theory (DFPT) in this section, we will focus on its application to describe homogeneous electric field perturbation in Section 3.1.2. In DFPT, the change of the KS orbitals in response to a external perturbational potential $\Delta v(\mathbf{r})$ can be obtained by solving the Sternheimer equation [43]:

$$\big(\hat{H}_{KS} - \varepsilon_n\big)|\Delta\psi_n\rangle = P_c\big(\Delta v + \Delta v^{lf}\big)|\psi_n\rangle, \tag{3.1}$$

where $\hat{H}_{KS}$ is the KS Hamiltonian of the unperturbed system defined in Eq. 2.14, P_c is the projector of the conduction band manifold: $P_c = 1 - \sum_n^N |\psi_n\rangle\langle\psi_n|$; Δv^{lf}, often referred to as the local field correction, is by definition:

$$\Delta v^{lf} = \int \frac{\Delta\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}' + \frac{\delta v_{xc}(\rho, |\nabla\rho|)}{\delta\rho}\Delta\rho + \frac{\delta v_{xc}(\rho, |\nabla\rho|)}{\delta|\nabla\rho|}\Delta|\nabla\rho|; \tag{3.2}$$

it accounts for the change of the Hartree and XC potential due to the change of the charge density $\Delta\rho(\mathbf{r}) = 4\sum_n^N \psi_n^*(\mathbf{r})\Delta\psi_n(\mathbf{r})$, where N is the number of doubly occupied KS orbitals. Note that here we only consider systems with zero total electronic spins, and we show the expression of Δv^{lf} for GGA XC potentials, which are functionals of the charge density ρ and its gradient $|\nabla\rho|$. For LDA XC potentials, the third term on the rhs is zero. In the case of hybrid functionals, the evaluation of Δv^{lf} is computationally demanding since Δv^{lf} depends explicitly on the wavefunctions ψ_n, through the expression of the HF exchange functional (see Eq. 2.18). Therefore, DFPT has been mainly implemented for LDA and GGA functionals. For hybrid functionals, it is computationally more efficient to use finite field methods instead of DFPT, as explained in Section 3.2.

We note that by using P_c in Eq. 3.1, one has $\langle\psi_n|\Delta\psi_m\rangle = 0$ for all n, m in the valence manifold. This property permits to solve the linear system of Eq. 3.1 using efficient iterative algorithms such as the conjugate gradient (CG) method [80].

3.1.2 Homogeneous Electric Field Perturbations

We now turn to the discussion of homogeneous electric field perturbations. In isolated systems, $\Delta v = \mathbf{E}\cdot\mathbf{r}$, where $\mathbf{E}$ is the external electric field. In condensed systems with periodic boundary conditions, the position operator $\mathbf{r}$ is ill defined and hence Δv is not straightforward to obtain. Within DFPT, the matrix element of the position operator is

written in terms of a commutator between the KS Hamiltonian and the position operator:

$$\langle\psi_m|\mathbf{r}|\psi_n\rangle = \frac{\langle\psi_m|[\hat{H}_{KS},\mathbf{r}]|\psi_n\rangle}{\varepsilon_m - \varepsilon_n}, \tag{3.3}$$

where ψ_n and ψ_m are KS eigenfunctions and ε_n and ε_m the corresponding eigenvalues, and:

$$[\hat{H}_{KS},\mathbf{r}] = -\nabla. \tag{3.4}$$

In the absence of nonlocal pseudopotentials [43, 81], only the kinetic energy part of the KS Hamiltonian does not commute with $\mathbf{r}$. When using nonlocal pseudopotentials, an extra term should be included in Eq. 3.4 [82]. For simplicity we assume using only local pseudopotentials in our formulation. We obtain $r_\nu|\psi_n\rangle$ by multiplying $|\psi_m\rangle$ on both sides of Eq. 3.3 and summing over m:

$$r_\nu|\psi_n\rangle = \sum_{m\neq n} |\psi_m\rangle \frac{\langle\psi_m|[\hat{H}_{KS}, r_\nu]|\psi_n\rangle}{\varepsilon_m - \varepsilon_n}, \tag{3.5}$$

where r_ν is the νth Cartesian component of $\mathbf{r}$ and we now define $r_\nu|\psi_n\rangle \equiv |\bar{\psi}_n^\nu\rangle$. The latter may be obtained by solving a linear system similar to Eq. 3.1:

$$\big(\hat{H}_{KS} - \varepsilon_n\big)|\bar{\psi}_n^\nu\rangle = P_c[\hat{H}_{KS}, r_\nu]|\psi_n\rangle. \tag{3.6}$$

Replacing $\Delta v|\psi_n\rangle$ in Eq. 3.1 by $-\sum_\nu E_\nu|\bar{\psi}_n^\nu\rangle$, where E_ν is the νth Cartesian component of the external electric field $\mathbf{E}$, we obtain the basic equation of DFPT in the case of an applied homogeneous electric field perturbation $\mathbf{E}$:

$$\big(\hat{H}_{KS} - \varepsilon_n\big)|\Delta^{\mathbf{E}}\psi_n\rangle = -\sum_\nu E_\nu|\bar{\psi}_n^\nu\rangle + P_c\Delta v^{lf}|\psi_n\rangle. \tag{3.7}$$

This equation can be solved self-consistently by updating Δv^{lf} using Eq. 3.2, where the response density $\Delta^{\mathbf{E}}\rho(\mathbf{r})$ is obtained by

$$\Delta^{\mathbf{E}}\rho(\mathbf{r}) = -4\sum_{n}^{N}\psi_n^*(\mathbf{r})\Delta^{\mathbf{E}}\psi_n(\mathbf{r}). \tag{3.8}$$

For crystalline solids, the self consistent solution of Eq. 3.7 may be obtained by assuming a constant applied electric field perturbation $\mathbf{E}$. However for slab geometries modeling surfaces or interfaces, this assumption is inappropriate since the electric field in the direction perpendicular to the surface is not constant, due to the polarization charge at the interface. In this cases, the applied perturbation to be used in Eq. 3.7 is the electric displacement field $\mathbf{D} = \mathbf{E} + 4\pi\mathbf{P}$ (see Section 5.4.2).

Once the wavefunction response is obtained, we compute the νth Cartesian component of polarization change using:

$$\begin{aligned}\Delta^{\mathbf{E}}P_\nu &= -\frac{4}{V}\sum_{n}^{N}\langle\psi_n|r_\nu|\Delta^{\mathbf{E}}\psi_n\rangle \\ &= -\frac{4}{V}\sum_{n}^{N}\langle\bar{\psi}_n^\nu|\Delta^{\mathbf{E}}\psi_n\rangle.\end{aligned} \tag{3.9}$$

The solutions of Eq. 3.9 may then be used to compute the macroscopic high-frequency dielectric constant:

$$\epsilon_\infty^{\mu\nu} = \delta_{\mu\nu} + 4\pi\frac{\Delta^{E_\mu}P_\nu}{E_\mu}, \tag{3.10}$$

where μ and ν are components of the Cartesian coordinates and $\delta_{\mu\nu}$ is the Kronecker delta function.

3.1.3 Implementation of DFPT in the Qbox Code

In this subsection, I will discuss the implementation of DFPT in the Qbox Code [46], including the calculation of Δv^{lf}, and the conjugate gradient (CG) linear solver.

The calculation of Δv^{lf} in Eq. 3.2 requires the functional derivative of the XC potential v_{xc}. For LDA functionals, an analytical form of v_{xc} is available. But for GGA functionals analytical functional derivatives are not available, and hence numerical functional derivative are computed. The XC potential for GGA functionals is expressed as a sum of two terms:

$$v_{xc} = v_1 + \nabla \cdot (v_2 \nabla \rho), \tag{3.11}$$

where v_1 and v_2 are the functional derivative of the GGA XC energy functional with respect to ρ and $\nabla\rho$, respectively [59]. The numerical derivatives of v_a $(a = 1, 2)$ with respect to the charge density ρ and the gradient $|\nabla\rho|$ are:

$$\frac{\partial v_a(\rho, |\nabla\rho|, \mathbf{r})}{\partial \rho(\mathbf{r})} = \frac{v_a(\rho + \delta\rho, |\nabla\rho|, \mathbf{r}) - v_a(\rho - \delta\rho, |\nabla\rho|, \mathbf{r})}{2\delta\rho}, \tag{3.12}$$

$$\frac{\partial v_a(\rho, |\nabla\rho|, \mathbf{r})}{\partial |\nabla\rho(\mathbf{r})|} = \frac{v_a(\rho, |\nabla\rho| + \delta|\nabla\rho|, \mathbf{r}) - v_a(\rho, |\nabla\rho| - \delta|\nabla\rho|, \mathbf{r})}{2\delta|\nabla\rho|}, \tag{3.13}$$

where $\delta\rho$ and $\delta|\nabla\rho|$ are chosen to be:

$$\delta\rho = min(10^{-4}, 0.01n), \delta|\nabla\rho| = min(10^{-4}, 0.01|\nabla\rho|).$$

The choices above help reduce the numerical instability for small values of the charge density. Then Δv_{xc} is given by:

$$\Delta v_{xc} = \left(\frac{\partial v_1}{\partial \rho} + \frac{1}{|\nabla\rho|} \frac{\partial v_1}{\partial |\nabla\rho|} \sum_\nu \partial_\nu \rho \partial_\nu \right) \Delta\rho + \sum_\nu \partial_\nu h_\nu, \tag{3.14}$$

where $\partial_\nu \equiv \partial / \partial r_\nu$ and

$$h_\nu = v_2 \partial_\nu \Delta\rho + \frac{\partial v_2}{\partial \rho} \Delta\rho \partial_\nu \rho + \frac{1}{|\nabla\rho|} \frac{\partial v_2}{\partial |\nabla\rho|} \left(\sum_\mu \partial_\mu \rho \partial_\mu \Delta\rho \right) \partial_\nu \rho. \tag{3.15}$$

These are all the equations necessary to compute Δv_{xc}.

We used the preconditioned CG (PCG) linear solver [80] to solve Eq. 3.5 and Eq. 3.6. Here we describe in detail the implementation of the PCG solver in the Qbox Code. We seek to solve an equation of the type $Ax = b$, where A is a square matrix and x and b are column vectors. In our case $A = \hat{H}_{KS} - \varepsilon_n$, and x and b denote wavefunctions represented in Fourier space. First, we define column vectors f, z, p as the residual, the preconditioned residual and the search direction, respectively. We denote the preconditioner as P_{prec}. For initial conditions:

$$f_0 = b - Ax_0 \tag{3.16}$$

$$z_0 = P_{prec} f_0 \tag{3.17}$$

$$p_0 = z_0 \tag{3.18}$$

where the subscript indicate the iteration number. The wavefunctions at the $(k+1)$ th iteration are updated as follows:

$$\alpha_{k+1} = \frac{r_k \cdot z_k}{p_k \cdot (Ap_k)} \tag{3.19}$$

$$f_{k+1} = r_k - \alpha_k A p_k \tag{3.20}$$

$$x_{k+1} = x_k + \alpha_k A p_k \tag{3.21}$$

$$z_{k+1} = P_{prec} f_{k+1} \tag{3.22}$$

$$\beta_{k+1} = \frac{r_{k+1} \cdot z_{k+1}}{r_k \cdot z_k} \tag{3.23}$$

$$p_{k+1} = r_k + \beta_k p_k \tag{3.24}$$

We use is a diagonal matrix as the preconditioner, with diagonal elements defined as:

$$P_{prec}(\mathbf{G}) = \begin{cases} 0.5 E_{cut}^{prec} & |\mathbf{G}|^2 < 2E_{cut}^{prec} \\ 1/|\mathbf{G}|^2 & |\mathbf{G}|^2 > 2E_{cut}^{prec} \end{cases} \tag{3.25}$$

where $\mathbf{G}$ is the index of the planewave basis defined in Eq. 2.20 and E_{cut}^{prec} is the preconditioner cutoff. We note that as E_{cut}^{prec} decreases, P_{prec} becomes similar to A, and this results in faster convergence of PCG algorithms. However, as E_{cut}^{prec} approaches 0, large $1/|\mathbf{G}|^2$ terms in the P_{prec} matrix may cause numerical instabilities. In our implementation, we set $E_{cut}^{prec} = 0.01E_{cut}$, where E_{cut} is the planewave kinetic energy cutoff used in KS DFT calculations.

It is well known that the CG algorithm diverges if A is a non-positive definite matrix [80]. In our case, in order to apply the CG algorithm to the non-positive definite matrix $(\hat{H}_{KS} - \varepsilon_n)$, we apply P_c on search directions p at each CG iteration so as to limit the solution x to the conduction band manifold. As the latter is spanned by the eigenvectors of $(\hat{H}_{KS} - \varepsilon_n)$ with positive eigenvalues, the convergence of the CG algorithm is guaranteed. The use of CG algorithms also requires attaining good convergence in solving KS equations prior to DFPT calculation, otherwise P_c contains valence band components which may lead to divergence of the CG algorithm.

3.2 Finite Field Methods for Homogeneous Electric Field

As mentioned in Section 3.1, density functional perturbation theory provides a way to treat external electric fields within the linear response regime. But the use of DFPT to obtain nonlinear responses and tackle orbital-dependent functionals (i.e. hybrid functional) is not straightforward and can be computationally cumbersome and expensive [43]. The Berry phase approach (or the modern theory of polarization) [83–85] provides a rigorous way to define the position operator in condensed systems, enabling first principles calculation of the change in polarizations. In a similar fashion, maximally localized Wannier functions (MLWF), or localized linear combinations of Bloch orbitals, provides a simple way to compute the position operator.

The Berry phase and MLWF approaches can be used to simulate insulating systems under applied external electric fields [48, 85], by introducing an electric enthalpy functional

[45]. However, we found that when using the Berry phase and MLWF approaches to evaluate the polarizability or dielectric properties of a system, the convergence of response function calculations with respect to the size of the supercell may be slow, especially with only the Γ point to sample the supercell Brillouin zone [86]. A method proposed by Stengel and Spaldin is a promising route to overcome the slow convergence [49]. This method makes use of a computationally inexpensive refinement procedure to estimate the position of MLWF centers. One can use similar procedures to calculate higher multipole moments which can be useful in parameterization of classical force field models and in simulations of nonlinear spectroscopies such as sum frequency generation (SFG). In this chapter, I will summarize the basic formulations of the Berry phase, MLWF and the refinement approaches, and their pactical implementation in the Qbox code [46, 47].

3.2.1 Finite Field Methods

Following Ref. [45] we define an electric enthalpy functional F:

$$F = E_{KS} - V\mathbf{E}\cdot\mathbf{P}, \tag{3.26}$$

where E_{KS} is the ground state KS energy functional, $\mathbf{E}$ is the applied total electric field, $\mathbf{P}$ is the polarization and V is the total volume of the system. The polarization $\mathbf{P}$ can be written as a sum of an ionic and an electronic contributions: $\mathbf{P} = \mathbf{P}^{ion} + \mathbf{P}^{elec}$. Obtaining the ionic contribution is trivial: $\mathbf{P}^{ion} = \sum_j Z_j\mathbf{R}_j$, where Z_j and $\mathbf{R}_j$ are the charge and position of the jth nucleus; however the calculation of $\mathbf{P}^{elec}$ is not straightforward. In analogy to the calculation of ground state KS orbitals via a minimization of the KS energy functional (see Eq. 2.13), the wavefunction and the polarization of the system in the presence of an applied electric field $\mathbf{E}$ may be obtained by minimizing F.

In order to minimize F, its functional derivative with respect the nth ground occupied

state KS orbital $\langle\psi_n|$ is needed:

$$\frac{\delta F}{\delta\langle\psi_n|} = \hat{H}_{KS}|\psi_n\rangle - V\mathbf{E}\cdot\frac{\delta\mathbf{P}}{\delta\langle\psi_n|}, \tag{3.27}$$

where $\hat{H}_{KS}$ is the KS Hamiltonian (Eq. 2.14). Since the nuclei are kept fixed within the BO approximation, $\delta\mathbf{P}^{ion}/\delta\langle\psi_n| = 0$, and thus $\delta\mathbf{P}/\delta\langle\psi_n| = \delta\mathbf{P}^{elec}/\delta\langle\psi_n|$. In the following I will discuss different methods to compute $\mathbf{P}$ and $\delta\mathbf{P}^{elec}/\delta\langle\psi_n|$.

The Berry phase approach was first introduced by several authors [83–85] using a formulation employing dense grids of k-points. Later, a specialized Γ point formulation [87] was used to treat perturbative electric fields [88, 89]. Using this approach the νth Cartesian component of polarization, $\mathbf{P}^{elec}$, are:

$$P_\nu^{elec} = -2\frac{L_\nu}{2\pi V}\,\mathrm{Im}\ln\det\mathbf{S}_\nu, \tag{3.28}$$

where the factor -2 accounts for the two negatively charged electrons in one occupied orbital, L_ν is the νth cell dimension; the mnth matrix elements of the $N\times N$ matrix $\mathbf{S}_\nu$, where N is number of occupied states, are:

$$S_{\nu,mn} = \langle\psi_m|e^{i\frac{2\pi}{L_\nu}r_\nu}|\psi_n\rangle. \tag{3.29}$$

The functional derivative of the electric enthalpy functional F is therefore:

$$\frac{\delta F}{\delta\langle\psi_n|} = \hat{H}|\psi_n\rangle + \sum_\nu E_\nu\frac{L_\nu}{2\pi}\,\mathrm{Im}\sum_{m=1}^{N} e^{i\frac{2\pi}{L_\nu}r_\nu}|\psi_m\rangle S^{-1}_{\nu,mn}, \tag{3.30}$$

where $S^{-1}_{\nu,mn}$ is the mnth element of the inverse of $\mathbf{S}_\nu$.

The MLWF approach is a formulation very closely related to the Berry phase [48, 90] one and its Γ point formulation [61] has been widely used to compute the dipole moment of molecules in condensed phases [23, 62, 91]. Within this approach, the electronic polarization

$\mathbf{P}^{elec}$ is:

$$\mathbf{P}^{elec} = \frac{-2}{V}\sum_n \mathbf{r}_n^0. \tag{3.31}$$

In Eq. 3.31, $\mathbf{r}_n^0$ is the Wannier center of the nth MLWF, w_n, defined as:

$$r_{n,\nu}^0 = \frac{L_\nu}{2\pi}\,\mathrm{Im}\ln\langle w_n|e^{i\frac{2\pi}{L_\nu}r_\nu}|w_n\rangle, \tag{3.32}$$

where w_n is obtained by applying a unitary transformation to the occupied Kohn-Sham (KS) eigenstates (see Eq. 2.13) so as to minimize the total spread: $(s^0)^2 = \sum_\nu (L_\nu/2\pi)^2 \sum_{n\neq m} \left|\langle w_n|e^{i\frac{2\pi}{L_\nu}r_\nu}|w_m\rangle\right|^2$; the latter is an approximation to the quadratic spread: $\sum_n \langle r^2\rangle_n - \langle\mathbf{r}\rangle_n^2$ [61]. The spread of individual MLWF is $(s_n^0)^2 = \sum_\nu (L_\nu/2\pi)^2\big(1 - \left|\langle w_n|e^{i\frac{2\pi}{L_\nu}r_\nu}|w_n\rangle\right|^2\big)$, and their second moment [91, 92] is given by:

$$\langle r_\nu r_\mu\rangle_n - \langle r_\nu\rangle_n\langle r_\mu\rangle_n = \frac{L^2}{16\pi^2}\left\{\ln\left|\langle w_n|e^{i\frac{2\pi}{L_\nu}r_\nu}e^{-i\frac{2\pi}{L_\mu}r_\mu}|w_n\rangle\right|^2 - \ln\left|\langle w_n|e^{i\frac{2\pi}{L_\nu}r_\nu}e^{i\frac{2\pi}{L_\mu}r_\mu}|w_n\rangle\right|^2\right\} \tag{3.33}$$

Within the MLWF approach, the functional derivative of F is:

$$\frac{\delta F^{mlwf}}{\delta\langle w_n|} = \hat{H}|\psi_n\rangle + \sum_\nu E_\nu \frac{L_\nu}{2\pi V}\,\mathrm{Im}\, e^{i\frac{2\pi}{L_\nu}r_\nu}|w_n\rangle/\langle w_n|e^{i\frac{2\pi}{L_\nu}r_\nu}|w_n\rangle. \tag{3.34}$$

As shown above, Berry phase and MLWF methods use the same kernel function, $\exp(i\frac{2\pi}{L_\nu}r_\nu)$, to evaluate the polarization in Γ-point only calculations. In the thermodynamic limit, the two methods yield the same, exact result [87]. However in practical calculations, this is often difficult or impossible to obtain. As pointed out by Stengel and Spaldin (see Fig. 1 of Ref. [49]), $\exp(i\frac{2\pi}{L_\nu}r_\nu)$ is an approximation to the position operator $\hat{\mathbf{r}}$ in periodic boundary conditions (PBC); such approximation is not justified when the unit cell is small compared to MLWF spreads [87], and it may lead to large errors in computing the polarization and other related physical properties, e.g. Born effective charge [49], polarizabilities [86] and high-frequency dielectric constants ϵ_∞ [86].

Stengel and Spaldin [49] proposed to compute the polarization in real space using saw-tooth functions instead of $\exp(i\frac{2\pi}{L_\nu}r_\nu)$. The calculation is done as a refinement step after the MLWFs w_n are computed, e.g. using the algorithm described in Ref. [93]. Since MLWFs decay exponentially in real space, the discontinuity of the saw-tooth function within PBC does not have any significant impact on the numerical stablity of the refinement procedure and on the final results. In the following, we describe in detail the refinement scheme and derive the functional derivatives required to optimize of the electric enthalpy functional (Eq. 3.27).

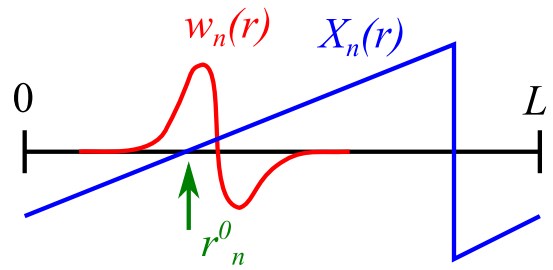


Figure 3.1: Schematic representation of the refinement approach described in Section 3.2 for a one-dimensional periodic system with periodicity L. $w_n(r)$ is nth Maximally localized Wannier Function (MLWF) localized in real space at r_n^0 and $X_n(r)$ is the corresponding saw-tooth function (see text).

The refinement procedure for a 1D system is illustrated in Fig. 3.1. The nth refined MLWF center $\mathbf{r}_n$ is computed as $\mathbf{r}_n = \mathbf{r}_n^0 + \Delta\mathbf{r}_n$, where $\mathbf{r}_n^0$ is defined by Eq. 3.32 and $\Delta\mathbf{r}_n$ is the refinement we seek to compute. The νth Cartesian component of $\Delta\mathbf{r}_n$, $\Delta r_{n,\nu}$, can be computed in real space using a periodic saw-tooth function, $X_{n,\nu}(r_\nu)$ (see Fig. 3.1), centered at the MLWF center $\mathbf{r}_n^0$:

$$X_{n,\nu}(r_\nu) = \begin{cases} r_\nu - r^0_{n,\nu} + L_\nu & \text{if } r_\nu - r^0_{n,\nu} <= -L_\nu/2 \\ r_\nu - r^0_{n,\nu} & \text{if } r_\nu - r^0_{n,\nu} \in (-L_\nu/2, L_\nu/2] \\ r_\nu - r^0_{n,\nu} - L_\nu & \text{if } r_\nu - r^0_{n,\nu} > L_\nu/2. \end{cases} \tag{3.35}$$

After multiplying the MLWF by the corresponding saw-tooth function in real space, we

define:

$$u_n^\nu(\mathbf{r}) \equiv X_{n,\nu}(r_\nu) w_n(\mathbf{r}), \tag{3.36}$$

and the refinement is computed as:

$$\Delta r_{n,\nu} = \langle w_n | u_n^\nu \rangle. \tag{3.37}$$

The total electronic contribution to the polarization becomes:

$$\begin{aligned} P_\nu^{elec} &= -\frac{2}{V} \sum_n r_{n,\nu}^0 + \Delta r_{n,\nu} \\ &= -\frac{2}{V} \sum_n r_{n,\nu}^0 + \langle w_n | u_n^\nu \rangle. \end{aligned} \tag{3.38}$$

The derivative of the electric enthalpy functional using the refinement scheme is:

$$\frac{\delta F}{\delta \langle w_n |} = \hat{H} | w_n \rangle + \sum_\nu E_\nu | u_n^\nu \rangle. \tag{3.39}$$

It is then straightforward to extend Eqs. 3.36 and 3.37 to compute the refined MLWF spread:

$$(s_n)^2 = \sum_\nu \langle u_n^\nu | u_n^\nu \rangle - (\Delta r_{n,\nu})^2. \tag{3.40}$$

Similarly, the refined second moment of the nth MLWF can be obtained as:

$$\langle r_\nu r_\mu \rangle_n - \langle r_\nu \rangle_n \langle r_\mu \rangle_n = \langle u_n^\nu | u_n^\mu \rangle - \Delta r_\nu \Delta r_\mu. \tag{3.41}$$

In the following, we will call this method the refined MLWF method.

3.2.2 Computational Details of Finite Field Implementations

We implemented the above three methods (Berry phase, MLWF and refined MLWF methods) in a massively parallel planewave FPMD code, the Qbox Code [46]. We verified our

implementation for single water molecules at the experimental geometry [94, 95] and for 64-water liquid water samples with a cell size of 23.46 a.u.; the latter were obtained from previous FPMD simulations [2, 14] with PBE [59, 60] and PBE0 [68] functionals. We used norm-conserving pseudopotentials [71, 72] and a planewave basis set with a kinetic energy cutoff of 85 Ry. We computed several quantities, including the dipole moment $\mathbf{M}$, quadrupole moment $\mathbf{Q}$ and the corresponding dipole and quadrupole polarizabilities; the latter is defined as the derivative of the quadrupole moment with respect to the applied electric field: $A_{\nu,\mu\xi} = dQ_{\mu\xi}/dE_\nu$. As reference results, we calculated dipole polarizabilities with linear response DFPT, as implemented in the Qbox Code (see Section 3.1). Additional reference results for polarizabilities were obtained from all-electron calculations as implemented in the NWChem software package [96], with a d-aug-cc-pVQZ basis set [1] so as to ensure convergence.

For a single molecule, the dipole moment $\mathbf{M} = V\mathbf{P}$. The quadrupole moment is often defined as a traceless tensor [97] with elements:

$$Q_{\nu\mu}^{traceless} = \frac{1}{2}\int d\mathbf{r}\rho(\mathbf{r})\big(3r_\nu r_\mu - \delta_{\nu\mu}r^2\big), \tag{3.42}$$

where $\rho(\mathbf{r})$ is the charge density and $\delta_{\nu\mu}$ is the Kronecker delta. Alternatively, the quadrupole moment can be defined as:

$$Q_{\nu\mu} = \frac{1}{2}\int d\mathbf{r}\rho(\mathbf{r})r_\nu r_\mu, \tag{3.43}$$

and it equals the derivative of the total energy with respect to the electric field gradient; such expression can be used to compute nonlinear vibrational spectra such as the SFG spectra [98] (see Chapter 5). $Q_{\nu\mu}^{traceless}$ and $Q_{\nu\mu}$ may be obtained using the second moment:

$$\int d\mathbf{r}\rho(\mathbf{r})r_\nu r_\mu = \sum_k Z_k R_{k,\nu}R_{k,\mu} - 2\sum_n \langle r_\nu r_\mu\rangle_n, \tag{3.44}$$

where the first and second terms on the rhs account for the ionic and electronic contributions,

respectively, and $\langle r_\nu r_\mu \rangle_n$ are computed using MLWF or refined MLWF methods, using Eq. 3.33 or Eq. 3.41, respectively. We computed dipole and quadrupole polarizabilities, using finite differences:

$$\alpha_{\nu\mu} = \frac{M_\mu(\delta E_\nu) - M_\mu(-\delta E_\nu)}{2\delta E_\nu}, \tag{3.45}$$

$$A_{\nu,\mu\xi} = \frac{Q_{\mu\xi}(\delta E_\nu) - Q_{\mu\xi}(-\delta E_\nu)}{2\delta E_\nu}, \tag{3.46}$$

where $\mathbf{M}$ is the dipole moment. δE was chosen to be 0.001 a.u.; it was chosen to be small enough to fall in the linear response regime and large enough to avoid numerical instabilities caused by small denominators.

We computed the high-frequency dielectric constant, $\boldsymbol{\epsilon}_\infty$, for liquid water samples, using the expression:

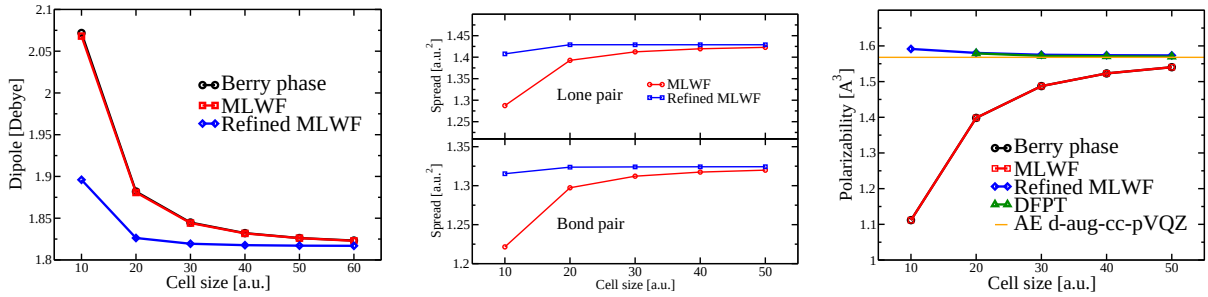
$$\epsilon_{\infty,\nu\mu} = \delta_{\nu\mu} + 4\pi \frac{\delta P_\mu}{\delta E_\nu}. \tag{3.47}$$

For isotropic systems such as liquids, $\boldsymbol{\epsilon}_\infty$ is diagonal with all equal elements.

3.2.3 Results for Single Water Molecules and Liquid Water

We first discuss our computed dipole moment, MLWF spread and polarizability for a single water molecule. We tested the convergence of these quantities with respect to the supercell size, where all calculations were carried out for the same molecular geometry. As shown in Fig. 3.2a, the dipole moment computed using Berry phase or MLWF aproaches are not converged even for cell edge as large as 60 a.u. Instead results from the refined MLWF method show a more favorable convergence behavior. The converged value obtained at 30 a.u. is 1.82 Debye, in agreement with all electron calculations [99] and the experimental value of 1.85 Debye [100]. The spread of the two types of MLWFs present in the water molecule, bond pair and lone pair MLWFs, is shown in Fig. 3.2b. Also in this case, the refined MLWF method shows faster convergence than the MLWF approach (see Fig. 3.2b).

The computed isotropic polarizability, $\alpha = (\alpha_{xx} + \alpha_{yy} + \alpha_{zz})/3$, is shown in Fig. 3.2c.



(a) (b) (c)

Figure 3.2: Dipole moment (a), spread of lone pair and bond pair MLWFs (b) and isotropic polarizability (c) of a water molecule computed at different cell size using Berry phase (black), MLWF (red), refined MLWF (blue), DFPT (green) and all electron (orange) approaches. The Berry phase, MLWF (Maximally localized Wannier Function) and refined MLWF methods are described in Section 3.2.1, and the DFPT (Density Functional Perturbation Theory) approach is described in Section 3.1. d-aug-cc-pVQZ denotes the basis set [1] adopted in All Electron (AE) calculations. Four MLWFs are associated to each water molecule (2 lone pairs and 2 bond pairs) constructed as linear combinations of occupied KS eigenstates.

As observed in previous studies [86], we found that the MLWF and Berry phase approaches show a rather slow convergence. Interestingly, both methods yield almost the same results at each cell size, probably because the same kernel function $\exp(i\frac{2\pi}{L_\nu}r_\nu)$ is used to compute the polarization. The polarizability computed with the refinement method converges to a value of 1.574 Å^3 at a cell size of 30 a.u. The convergence behavior is almost the same as that observed using DFPT [43], and the converged result is almost identical to the DFPT one (1.571 Å^3). The converged results are also in good agreement with all-electron calculations (1.568 Å^3). Not surprisingly, our results computed with the PBE functional overestimated the experimental value of 1.47 Å^3 [101]. This is mainly due to the delocalization error introduced by semilocal functionals such as PBE [66]. Hybrid functionals, e.g. PBE0, reduce this error and PBE0 yielded a value in better agreement with experiments, 1.46 Å^3, when using the refined MLWF methods.

Our results for the quadrupole moments and quadrupole polarizabilities are shown in Fig. 3.3; it is seen that the calculated traceless quadrupole and the components of the quadrupole polarizabilities converge for cell size of 30 a.u.. The quadrupole polarizability

component $A_{Z,ZZ}$ converges slightly slower than the other components probably because of the spurious Coloumb interaction between the Z-component of the dipole associated with water molecules in neighboring cells. In Table 3.1, we show that the traceless quadrupole computed with the refined MLWF method does not depend on the molecular orientation. The MLWF method, on the other hand, showed poor convergence and yielded inconsistent results for different molecular orientations. As shown in Table 3.1, our results are also consistent with all electron calculations, as well as experimental data reported in Ref. [97].

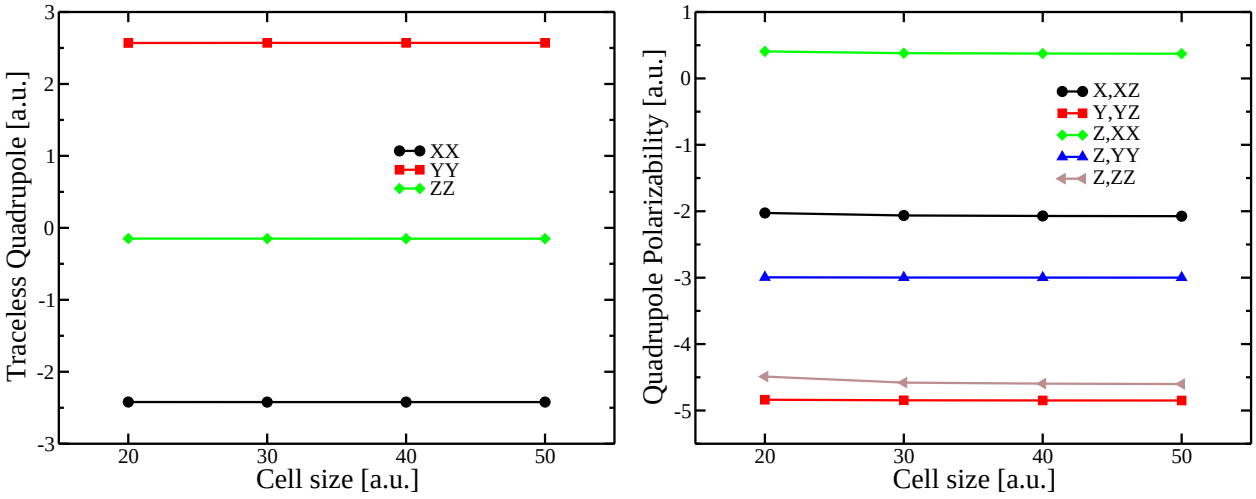


Figure 3.3: Diagonal elements of the traceless quadrupole moment (a) and the quadrupole polarizability (b) of a water molecule computed using refined MLWF method (see Section 3.2.1) at different cell sizes. Here X, Y and Z denote the axis in the molecular plane, the axis perpendicular to the molecular plane and the dipole axis of the water molecule, respectively.

Now we turn to condensed systems and present the results obtained using the refined MLWF method for one configuration extracted from FPMD simulations of liquid water [2]. In particular we show the computed high-frequency dielectric constant, and molecular dipole and quadrupole moments. In Table 3.2, we show ϵ_∞ computed using Eq. 3.47 and several different methods for the polarization $\mathbf{P}$. We found that for calculations using the PBE functional, the refined MLWF and DFPT approaches give consistent results while the Berry phase approach predicts a slightly smaller ϵ_∞. This is not surprising as we already observed that the Berry phase method underestimates the dipole polarizability of an isolated water

Table 3.1: Diagonal elements of the traceless quadrupole moment, in Buckingham, of an isolated water molecule

	Config. 1[a]			Config. 2[b]		
	X	Y	Z	X	Y	Z
Refined MLWF[c]						
PBE cell 30	2.571	-2.421	-0.150	2.571	-2.421	-0.149
PBE cell 50	2.571	-2.421	-0.150	2.571	-2.421	-0.150
PBE0 cell 30	2.572	-2.428	-0.145	2.572	-2.427	-0.145
MLWF[d]						
PBE cell 30	2.614	-2.375	-0.238	2.650	-2.341	-0.309
PBE cell 50	2.587	-2.404	-0.183	2.584	-2.406	-0.177
All electron[e]						
PBE	2.555	-2.408	-0.147			
PBE0	2.551	-2.407	-0.144			
Other work						
MLWF PBE [f]	2.58	-2.45	-0.13			
Expt.[g]	2.63 ± 0.02	-2.50 ± 0.02	-0.13 ± 0.03			

[a] Calculations carried out for a water molecule with its molecular axis (X, Y and Z, which denote the axis in the molecular plane, the axis perpendicular to the molecular plane and the dipole axis of the water molecule, respectively) parallel to the supercell edges. [b] Calculations carried out for a water molecule with the same geometry as Config. 1, but at a different, random orientation. [c] Results obtained using the refined MLWF method (Eqs. 3.41 and 3.44), with PBE [59, 60] or PBE0 [68] functionals; the specified cell dimensions are in a.u.. [d] Results obtained using the MLWF method (Eqs. 3.33 and 3.44) adn the PBE functional. [e] All electron calculations carried out using the NWChem code [96] (see text). [f] Results obtained using the MLWF method and the PBE functional [91]. [g] Experimental results from Ref. [97].

Table 3.2: High-frequency dielectric constant, ϵ_∞, of a 64-molecule liquid water sample extracted from FPMD trajectories obtained in the NVE ensemble and the PBE functional, with the Qbox code. Calculations of ϵ_∞ were carried out with both GGA (PBE) and hybrid (PBE0) functionals. The Berry phase and refined MLWF methods are described in Section 3.2.1, and the DFPT approach is described in Section 3.1.

Method	ϵ_∞
Refined MLWF PBE	1.873
DFPT PBE	1.879
Berry phase PBE	1.796
Refined MLWF PBE0	1.782

molecule at this cell size (23.464 Å) (see Fig. 3.2c). Note that the value predicted by the Berry phase method and the PBE functional is in good agreement with experimental value of about 1.77, mainly due to error cancellations. Similar results have been found in previous FPMD simulations using the same approach [102]. As shown in Table 3.2, by using the PBE0 functional and the refined MLWF method, one observes converged results in good agreement with the experimental value. In addition to ambient conditions, we also used the refined MLWF method with the PBE0 functional to accurately compute the high-frequency dielectric constant of liquid water under high temperature and pressure conditions [39].

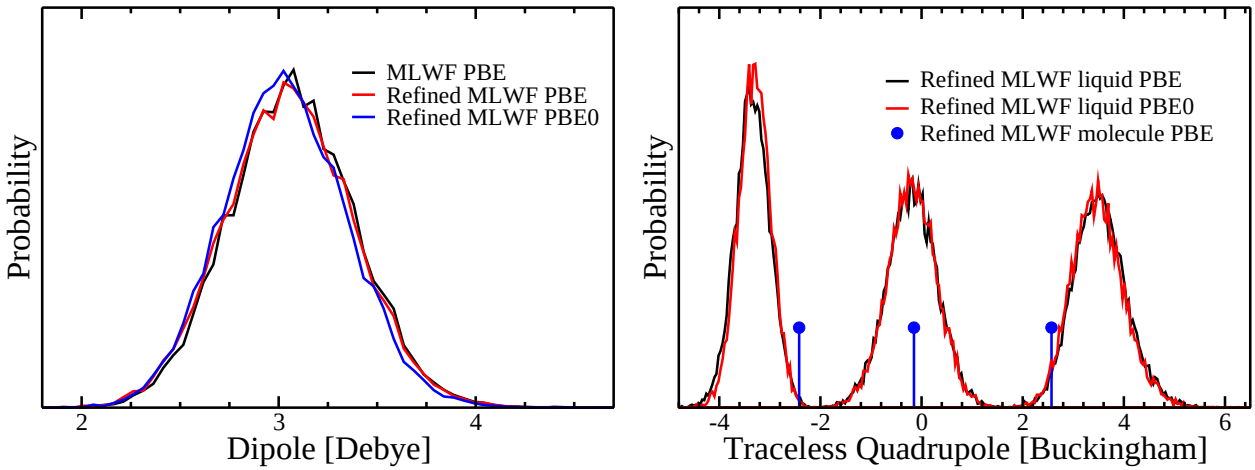


Figure 3.4: (a) Distribution of dipole moments of water molecules computed using the MLWF method and PBE functional (black), the refined MLWF method and PBE functional (red) and the refined MLWF and PBE0 functional (blue) for 200 snapshots extracted from an FPMD simulation using PBE functional and 64 molecule liquid water sample [2]. (b) Distribution of traceless quadrupole moments of water molecules computed using the refined MLWF method in combination with PBE (black) and PBE0 functionals (red) for the same snapshots. The blue dots and vertical lines indicate values computed for an isolated water molecule using the refined MLWF method and PBE functional.

In Fig. 3.4, we show the distribution of the computed dipole and traceless quadrupole moments of the water molecules taken from 200 configurations extracted from a 50-ps FPMD simulation [2]. We found that the dipole moment contribution computed from MLWF and refined MLWF methods are nearly identical: 3.06 and 3.07 Debye, respectively, with the PBE functional. The dipole moment computed with the PBE0 functional using the same

Table 3.3: Average diagonal elements of the traceless quadrupole moment tensor of water molecules in liquid computed using the refined MLWF method and the PBE or PBE0 functionals and 200 snapshots extracted from FPMD simulations using PBE [2] and PBE0 [14] functionals.

Refined MLWF Method	X	Y	Z
PBE@PBE	3.53	-3.34	-0.19
PBE0@PBE	3.49	-3.30	-0.19
PBE@PBE0	3.32	-3.13	-0.19
PBE0@PBE0	3.28	-3.10	-0.18

configurations is slightly smaller, 3.03 Debye. This value is much larger than the dipole moment of a single water molecule, 1.82 Debye, in agreement with results obtained in previous FPMD simulations [62, 91]. In Fig. 3.4b, we show the traceless quadrupole moment computed using the refined MLWF method and the PBE functional. The average of its three components are shown in Table 3.3 and are much larger than the values obtained for a single water molecule, shown in Table 3.1. We note that predictions of enhanced quadrupole moments of water molecules in the liquid, compared to the gas phase, have been reported by several FPMD studies [91, 103] and quantum chemistry calculations [104–106]. Large dipole and quadrupole moments of water molecules in the liquid originates from strong polarization due to the strong electric field in the liquid, which also affects the molecular polarizabilities as shown in previous studies, [2, 107, 108].

3.2.4 Conclusions

In summary, we proposed an accurate *ab initio* method to carry out electronic structure calculations in the presence of finite electric fields, by combining the refinement scheme [49] based on MLWF with an electric enthalpy functional [45]. We also extended the refinement scheme to the computation of multipole moments, e.g. quadrupole moments. Our method works equally well with semilocal and hybrid functionals, at variance with DFPT which is not straightforward to implement when using hybrid functionals. We showed that the proposed method converges much faster with respect to cell size than conventional Γ-only

calculations, based on Berry phase and MLWF approaches, and it can be implemented with low computational costs. Our predictions of electrostatic properties using hybrid functionals, e.g. dipole and quadrupole moments and their polarizabilities in the condensed phase, are in very good agreement with experiments.

The predicted multipole moments and polarizabilities may be useful for the parametrization of empirical force fields and will be used in the next chapters for the calculation of vibrational spectra. Another potential application of our method is in FPMD simulations in the presence of electric fields, which may be useful to study electrochemical systems under working conditions.

CHAPTER 4

VIBRATIONAL SIGNATURES OF CHARGE FLUCTUATIONS IN THE HYDROGEN BOND NETWORK OF LIQUID WATER

In the previous chapter, we described first-principles linear response and finite field methods for first-principles calculations in the presence of electric fields and their implementation in the Qbox Code. Starting from this chapter, we will describe the application of these methods for the calculation of vibrational spectra of aqueous systems. In particular, we will present results for Raman spectra of liquid water and sum-frequency generation (SFG) spectra of ice surfaces in this and the next chapters, respectively. We used FPMD and the linear response DFPT approach for Raman spectra simulations. For the SFG spectra calculation, we used DFPT for the evaluation of dipole polarizabilities and refined MLWF methods for computing quadrupole moments and quadrupole polarizabilities.

4.1 Introduction to Vibrational Spectroscopy

In this section we describe the basic theoretical formulations of vibrational spectroscopies, with the most well studied ones being infrared and Raman spectroscopies. In general vibrational spectra can be computed using two approaches, by diagonalizing the dynamical matrix to get normal modes or by computing time correlation functions (TCF). The dynamical matrix is computed from the second derivatives of the total energy (first derivative of the forces), obtained by finite differences. Within the TCF approach, spectra are obtained as Fourier transforms of the TCF of dipole moments and polarizabilities for infrared and Raman spectra, respectively. The normal mode approach is usually computationally less expensive than TCF approaches because it requires a small number of calculations at different atomic geometries. But this approach can only be applied to solids, and the anharmonic effects are usually neglected. TCF approaches, on the other hand, can be used for disordered systems, e.g. liquids, and they include anharmonic effects. However obtaining converged

TCF requires long molecular dynamics simulations, and it may be computationally quite demanding.

4.1.1 Normal Mode and TCF Approaches

Within the normal mode approach, the dynamical Matrix $\mathcal{D}$ is obtained for a geometrical configuration corresponding to a local minimum of the potential energy surface. The size of the matrix is determined by the number of the degrees of freedom ($3N_{atom}$), and its elements are given by:

$$\mathcal{D}_{pq} = \frac{1}{\sqrt{M_p M_q}} \frac{\partial^2 E}{\partial R_p \partial R_q} = -\frac{1}{\sqrt{M_p M_q}} \frac{\partial F_p}{\partial R_q}, \tag{4.1}$$

where E is the total energy, R_q, M_q and F_q are the position, mass and force of the qth degree of freedom. The qth eigenvalue and eigenvector of $\mathcal{D}$ are the squared frequency ω_q^2 and normal displacement S_q of the qth vibrational normal mode, respectively, within the harmonic approximation. A generic spectral intensity $I(\omega)$, e.g. of infrared or Raman spectra, can be computed using the intensity A_q of the qth normal mode:

$$I(\omega) = \sum_q A_q \delta(\omega - \omega_q), \tag{4.2}$$

where $\delta(\omega)$ is the Dirac function, which can be replaced by a Gaussian distribution function to obtain smooth spectra. We shall show explicit expressions of A_q for infrared and Raman spectra in Sections 4.1.2 and 4.1.3, respectively.

Making use of the Fermi's golden rule, the infrared and Raman intensities may be computed as Fourier transform of the TCF of the system's dipole moments and polarizabilities, respectively [31]. Within the TCF approach, a generic spectral intensity $I(\omega)$ is given by:

$$I(\omega) \propto \int \left\langle B(0)B(t) \right\rangle \exp(-i\omega t) dt, \tag{4.3}$$

where B denotes total dipole moment or polarizability of the system for infrared or Raman

spectra calculations, respectively. In Eq. 4.3, $\langle \cdots \rangle$ denotes ensemble average and $\langle B(0)B(t) \rangle$ is called the TCF of the quantity B. Note that here we used a classical TCF to approximate its quantum mechanical counterpart. In practice, a quantum correction factor [109] is often used to correct classical TCFs for better reproduction of the behavior of the corresponding quantum mechanical TCF.

One useful technique to analyze the spectra for molecular systems is to decompose B into molecular quantities b_i:

$$B = \sum_i b_i. \tag{4.4}$$

Then the total spectra can be decomposed into intra- and intermolecular contributions:

$$I_B^{intra}(\omega) \propto \sum_i \int \left\langle b_i(0) b_i(t) \right\rangle \exp(-i\omega t) dt, \tag{4.5}$$

$$I_B^{inter}(\omega) \propto \sum_{i \neq j} \int \left\langle b_i(0) b_j(t) \right\rangle \exp(-i\omega t) dt. \tag{4.6}$$

Analysis on intra- and intermolecular contributions turned out to be very useful for studying subtle intermolecular interactions, e.g. hydrogen bond interactions in aqueous systems [23, 35].

4.1.2 Infrared Spectroscopy

Infrared spectroscopy measures the absorption of infrared radiation at frequencies in resonance with the system's vibrational modes. A vibrational mode is infrared active if it has a net change in dipole moment $\mathbf{M}$ upon adsorption of radiation. Within the normal mode approach, the infrared intensity for the qth normal mode is given by [31]:

$$A_q^{IR} = \frac{\pi}{3cV} \left| \frac{\partial \mathbf{M}}{\partial S_q} \right|^2. \tag{4.7}$$

Using TCF instead, the infrared intensity is

$$I^{IR}(\omega) \propto \int \left\langle \mathbf{M}(0) \cdot \mathbf{M}(t) \right\rangle \exp(-i\omega t) dt. \tag{4.8}$$

Infrared spectra measurements of liquid systems usually report absorption coefficients per unit length [23, 35, 110]:

$$a(\omega) = \frac{2\pi\omega^2\beta}{3cVn(\omega)} \int \sum_{ij} \left\langle \mathbf{M}_i(0) \cdot \mathbf{M}_j(t) \right\rangle \exp(-i\omega t) dt. \tag{4.9}$$

where $\beta = 1/kT$ is the Boltzmann factor, c is the speed of light, $n(\omega)$ is the frequency-dependent refractive index and the summation is performed over all molecular dipole moment within a volume V. Here we use a harmonic approximation quantum correction factor [109].

4.1.3 Raman Spectroscopy

A Raman scattering effect occurs when the frequency of the scattered light is different from that of the incident radiation. This change in frequency, called the Raman shift, arises from the interaction between the incident radiation with vibrational modes of the system. Raman spectra measures the intensity of the scattered light as a function of the Raman shift. Raman intensities may be computed from polarizability ($\boldsymbol{\alpha}$) calculations [111]. First it is desirable to decompose the polarizability tensor $\boldsymbol{\alpha}$ into isotropic and anisotropic parts:

$$\boldsymbol{\alpha} = \bar{\alpha}\mathbf{I} + \boldsymbol{\beta}, \tag{4.10}$$

where $\mathbf{I}$ is the identity tensor and $\bar{\alpha}\mathbf{I}$ and $\boldsymbol{\beta}$ are the spherical and anisotropic parts of the $\boldsymbol{\alpha}$ tensor.

$$I_q^{Raman} = 45\tilde{a}_q^2 + 7\tilde{b}_q^2, \tag{4.11}$$

where

$$\tilde{a}_q = \frac{\partial \bar{\alpha}}{\partial S_q}, \tag{4.12}$$

$$\tilde{b}_q^2 = \frac{3}{2}\mathrm{Tr}\left(\frac{\partial \boldsymbol{\beta}}{\partial S_q}\frac{\partial \boldsymbol{\beta}}{\partial S_q}\right) \tag{4.13}$$

$$= \frac{1}{2}\Bigg(\big(\frac{\partial \alpha_{xx}}{\partial S_q} - \frac{\partial \alpha_{yy}}{\partial S_q}\big)^2 + \big(\frac{\partial \alpha_{yy}}{\partial S_q} - \frac{\partial \alpha_{zz}}{\partial S_q}\big)^2 + \big(\frac{\partial \alpha_{zz}}{\partial S_q} - \frac{\partial \alpha_{xx}}{\partial S_q}\big)^2 \tag{4.14}$$

$$+ 6\big(\frac{\partial \alpha_{xy}}{\partial S_q}\big)^2 + 6\big(\frac{\partial \alpha_{yz}}{\partial S_q}\big)^2 + 6\big(\frac{\partial \alpha_{zx}}{\partial S_q}\big)^2\Bigg). \tag{4.15}$$

Here $\tilde{a}^2$ and $\tilde{b}_q^2$ correspond to the isotropic and anisotropic Raman intensities. Tr denotes the trace operator. It can be clearly seen from the above expressions that the isotropic intensity describes the Raman intensity arising from isotropic or spatially averaged polarizabilities, obtained as the trace of the polarizability tensor, while the anisotropic one describes the anisotropy of the tensor. We can also calculate the depolarization ratio:

$$\eta = \frac{3\tilde{b}_q^2}{45\tilde{a}_q^2 + 4\tilde{b}_q^2}, \tag{4.16}$$

which describes the ratio between the intensities of the scattered light with polarizations perpendicular (VH) or parallel (VV) to the that of the incident light.

In the TCF formulation, Raman intensities can be expressed in terms of the Fourier transform of the TCF of the polarizability, similar to the infrared spectra in Eq. 4.8 [31]:

$$I_{iso}^{Raman}(\omega) \propto \int dt e^{-i\omega t}\left\langle \bar{\alpha}(0)\bar{\alpha}(t)\right\rangle \tag{4.17}$$

$$I_{aniso}^{Raman}(\omega) \propto \int dt e^{-i\omega t}\left\langle \frac{2}{15}\mathrm{Tr}\boldsymbol{\beta}(0)\boldsymbol{\beta}(t)\right\rangle. \tag{4.18}$$

These two intensities correspond to the contributions of spatially averaged and anisotropic polarizabilities, respectively. Alternatively we can derive the expression for the experimental

measurable VV and VH spectra:

$$I_{VV}^{Raman}(\omega) \propto \int dt e^{-i\omega t} \left\langle \bar{\alpha}(0)\bar{\alpha}(t) + \frac{2}{15}\mathrm{Tr}\boldsymbol{\beta}(0)\boldsymbol{\beta}(t) \right\rangle \tag{4.19}$$

$$I_{VH}^{Raman}(\omega) \propto \int dt e^{-i\omega t} \left\langle \frac{1}{10}\mathrm{Tr}\boldsymbol{\beta}(0)\boldsymbol{\beta}(t) \right\rangle. \tag{4.20}$$

4.2 Raman Spectroscopy for Liquid Water

Raman spectroscopy was extensively employed to investigate the O-H stretching band [3, 4, 112–115] and low-frequency translational and librational bands [5, 112, 116–123] of liquid water at ambient conditions and under pressure [124–127], as well as the vibrational properties of solvated ions [128–131] and biological molecules [132, 133] in solutions, and of nano-confined water [134, 135].

Most of the theoretical studies carried out to interpret the measured Raman spectra of water are currently based on classical simulations [132, 136–144], with Raman intensities computed from the time-correlation function (TCF) of the polarizability tensor [31]. The latter is obtained from molecular polarizabilities either derived empirically or calculated from small water clusters using *ab initio* electronic structure methods. In principle quantum TCFs are required for an accurate description of the vibrational excitations. However, in most simulations [132, 136–141] quantum TCFs are approximated with classical ones, and then multiplied by a so-called quantum correction factor [109]. The authors of Ref. [141–144] proposed a method to compute quantum TCFs within a semi-classical approximation, encompassing the generation of trajectories using classical molecular dynamics (MD) and the calculation of Raman intensities of clusters extracted from MD trajectories. These calculations yielded very good agreement with experiments, but they were restricted to the analysis of the O-H stretching band. Most results for the low-frequency part of the Raman

spectra were not as satisfactory [132, 136–138, 140], and a number of open questions remain in the interpretation of the Raman spectra of liquid water at ambient condition.

In particular, the origin of the low-frequency bands at 60 cm^{-1} and 200 cm^{-1} are not fully understood [145–147], and their interpretation calls for the use of *ab initio* calculations, where electrons are treated explicitly. As pointed out in the case of infrared spectra [23, 35, 63], it is important to properly account for dynamical changes in the electronic structure of water molecules and of the HB network to accurately describe the vibrational properties of the liquid.

In this chapter we report the first *ab initio* simulation of the Raman spectra of liquid water, obtained by using density functional perturbation theory (DFPT) [43]. *Ab initio* calculations of Raman spectra based on classical TCF and polarizabilities derived from the modern theory of polarization were so far only reported for crystalline systems [40, 41], but not for liquids. We also developed a systematic strategy to interpret the computed Raman spectra, which is of general applicability to solid and liquid phases of molecular systems. Our analysis is based on maximally localized Wannier functions (MLWF) [48], previously used to compute the infrared spectrum of liquid water [23, 35, 62, 63, 148]. Our interpretation of the high-frequency O-H stretching band is overall consistent with previous experimental and theoretical studies. In the low-frequency region, our analysis of the isotropic spectrum revealed intermolecular charge fluctuations accompanying HB stretching vibrations, despite the absence of any Raman intensity. Such fluctuations, identified at 200 cm^{-1} in the infrared spectra [22–24], exhibit signals on a wider frequency range in our calculations, up to 270 cm^{-1}. We further defined molecular polarizabilities [107, 108, 149, 150], which may be useful in the parameterization of polarizable force fields for classical MD simulations of water.

The rest of this chapter is organized as follows: in the next section we describe our computational and analysis methods used in Raman spectra simulations of liquid water. In the Results and Discussion section, first we compare the calculated Raman spectra with experimental results; we then focus on the interpretation of the low frequency part of the

spectrum, and we show evidence of intermolecular charge fluctuations using several complementary analysis tools. Finally, we present our conclusions.

4.3 Theoretical Methods

4.3.1 Simulation Details

We carried out Born-Oppenheimer *ab initio* MD simulations of liquid water using the Qbox code [46] and the semilocal exchange-correlation functional PBE [59, 60]. Based on previous studies with the PBE functional, we used a 64 heavy water sample at the experimental density and an elevated temperature around 400K, to better reproduce the liquid radial distribution functions and diffusion coefficient [10, 35]. We used a plane wave basis set with a kinetic energy cutoff of 85 Ry, norm conserving pseudopotentials [71] of the HSCV type [72] and only the Γ point to sample the Brillouin zone. Our MD simulations were carried out with a time step of 10 a.u. (0.24 fs) in the NVE ensemble for 50 ps after a 5 ps equilibration in the NVT ensemble, using the thermostat proposed in Ref. [151]. We adopted the diagonalization algorithm of Ref. [93] to compute the MLWFs [48].

We computed the Raman spectra as the Fourier transform of the TCF of the system's polarizabilities [31]. In this paper, we report the Bose-Einstein (BE) reduced [152] isotropic and anisotropic Raman spectra:

$$R_{iso}(\omega) \propto \frac{\hbar\omega}{kT} \int dt e^{-i\omega t} \Big\langle \bar{\alpha}(0)\bar{\alpha}(t) \Big\rangle, \tag{4.21}$$

$$R_{aniso}(\omega) \propto \frac{\hbar\omega}{kT} \int dt e^{-i\omega t} \Big\langle \frac{2}{15} \mathrm{Tr}\beta(0)\beta(t) \Big\rangle. \tag{4.22}$$

In Eqs. 4.21 and 4.22, ω is the frequency, Tr denotes a trace operator, T is the temperature and k is the Boltzmann constant. $\bar{\alpha}$ and β are defined in Eq. 4.10. The prefactor $\frac{\hbar\omega}{kT}$ in Eqs. 4.21 and 4.22 is the product of the BE factor [152], $1 - e^{-\hbar\omega/kT}$, and the harmonic approximation quantum correction factor [109], $\frac{\hbar\omega/kT}{(1-e^{-\hbar\omega/kT})}$. The BE factor is useful

in analyzing the low-frequency part of the Raman spectra, making features obscured by the high-intensity Rayleigh scattering peak more easily identifiable [152]. The use of the BE factor also ensures the correct temperature dependence of the Raman intensity in the low-frequency region [112]. The quantum correction factor has been previously used in *ab initio* simulations of infrared spectra [23, 35, 110] to approximate quantum time correlation functions (TCF) by their classical counterparts.

4.3.2 Effective molecular polarizabilities

We computed the polarizability using density functional perturbation theory (DFPT) [43], as implemented in the Qbox code [46] (see details in Section 3.1). In order to understand the contribution of individual water molecules to the total polarizability, we defined the effective molecular polarizability $\boldsymbol{\alpha}_i^{\text{eff}}$ of the ith water molecule by projecting the total polarizability $\boldsymbol{\alpha}$ onto MLWFs. By applying the MLWF transformation (see Section 3.2) to $\Delta^{\mathbf{E}}\psi$ and $\bar{\psi}^{\mu}$, we obtain the response wavefunctions $\Delta^{\mathbf{E}}w_n$ and $\bar{w}^m u_n$, corresponding to each MLWFs. By replacing $\Delta^{\mathbf{E}}\psi_n$ and $\bar{\psi}_n$ in Eq. 3.9 with $\Delta^{\mathbf{E}}w_n$ and $\bar{w}_n$, we obtain the projection of the total polarization onto the nth MLWF:

$$P_{w_n}^{\mu} = -\frac{4e}{V}\langle \Delta^{\mathbf{E}} w_n | \bar{w}_n^{\mu} \rangle. \tag{4.23}$$

We then define the effective polarizability associated to each MLWF as:

$$V\mathbf{P}_{w_n} = \Delta\mathbf{M}_{w_n} = \boldsymbol{\alpha}_{w_n}^{\text{eff}}\mathbf{E}. \tag{4.24}$$

In the case of water molecules, four valence MLWFs are associated to the four doubly occupied valence eigenstates: two lone pairs (LPs) and two bond pairs (BPs). Thus the effective molecular polarizability of each molecule is defined as:

$$\boldsymbol{\alpha}_i^{\text{eff}} = \boldsymbol{\alpha}_{i,\text{LP1}}^{\text{eff}} + \boldsymbol{\alpha}_{i,\text{LP2}}^{\text{eff}} + \boldsymbol{\alpha}_{i,\text{BP1}}^{\text{eff}} + \boldsymbol{\alpha}_{i,\text{BP2}}^{\text{eff}}. \tag{4.25}$$

The induced dipole $\Delta\mathbf{M}_i$ corresponding to the ith molecule is therefore:

$$\Delta\mathbf{M}_i = \boldsymbol{\alpha}_i^{\text{eff}}\mathbf{E}, \tag{4.26}$$

where $\boldsymbol{\alpha} = \sum_i \boldsymbol{\alpha}_i^{\text{eff}}$, and $\boldsymbol{\alpha}_i^{\text{eff}} = \sum_c \boldsymbol{\alpha}_{i,c}^{\text{eff}}$; $\boldsymbol{\alpha}_{i,c}^{\text{eff}}$ is the effective polarizability associated to each of the four MLWFs (c) belonging to a water molecule (2 bond pairs and 2 lone pairs MLWFs, constructed from a unitary transformation of valence eigenstates of the Hamiltonian).

The BE reduced Raman intensities (Eqs. 4.21 and 4.22) can be therefore be represented using effective molecular polarizabilities:

$$R_{iso}(\omega) \propto \frac{\hbar\omega}{kT}\int dt e^{-i\omega t}\left\langle \sum_{i,j}\bar{\alpha}_i^{\text{eff}}(0)\bar{\alpha}_j^{\text{eff}}(t)\right\rangle, \tag{4.27}$$

$$R_{aniso}(\omega) \propto \frac{\hbar\omega}{kT}\int dt e^{-i\omega t}\left\langle \frac{2}{15}\text{Tr}\sum_{i,j}\beta_i^{\text{eff}}(0)\beta_j^{\text{eff}}(t)\right\rangle, \tag{4.28}$$

where $\bar{\alpha}_i^{\text{eff}} = \frac{1}{3}\text{Tr}\boldsymbol{\alpha}_i^{\text{eff}}$ and $\boldsymbol{\alpha}_i^{\text{eff}} = \bar{\alpha}_i^{\text{eff}}\mathbf{I} + \boldsymbol{\beta}_i^{\text{eff}}$. Here Tr denotes the trace operator. By separating the $i = j$ and $i \neq j$ terms in the summation in Eqs. 4.21 and 4.22, we define the intra- and intermolecular contributions to the total Raman spectra, similar to previous infrared spectra studies [23, 35, 63]. The $i = j$ terms, $R_{iso}^{intra}(\omega)$ and $R_{aniso}^{intra}(\omega)$, are contributions from localized vibrations on each molecule or intramolecular terms, while the $i \neq j$ terms, $R_{iso}^{inter}(\omega)$ and $R_{aniso}^{inter}(\omega)$, are contributions from coupled vibrations on different molecules or intermolecular terms.

4.3.3 Molecular Polarizabilities

Following Refs. [107, 108, 149, 150] we defined the molecular polarizabilities, $\boldsymbol{\alpha}_i$, of the ith water molecule in the liquid as:

$$\Delta\mathbf{M}_i = \boldsymbol{\alpha}_i\mathbf{E}_i^{\text{loc}}, \tag{4.29}$$

where $\Delta\mathbf{M}_i$ is the induced dipole of the ith molecule as defined in Eq. 4.26 and $\mathbf{E}_i^{\mathrm{loc}}$ is the local electric field acting on the ith molecule; note that $\mathbf{E}_i^{\mathrm{loc}}$ contains the contribution of both the applied field $\mathbf{E}$ and the field induced by the induced dipole $\Delta\mathbf{M}_j$ ($j \neq i$) of all other molecules in the system via dipole-induced dipole (DID) interactions [132, 137, 138, 140, 141, 153]. This method is widely used in classical calculations of Raman spectra to obtain the total polarizability of the extended system from the molecular polarizabilities. In principle, higher multipoles also contributes to E_i^{loc}, but these effects are usually small compared to DID interactions [154]; hence $\mathbf{E}_i^{\mathrm{loc}}$ can be approximated by

$$\mathbf{E}_i^{\mathrm{loc}} = \mathbf{E} + \sum_{i \neq j} \mathbf{T}_{ij} \Delta\mathbf{M}_j, \tag{4.30}$$

where $\mathbf{T}_{ij}$ is the DID interaction tensor:

$$T_{ij}^{\mu\nu} = \frac{r_{ij}^2 \delta_{\mu\nu} - 3 r_{ij}^{\mu} r_{ij}^{\nu}}{r_{ij}^5}, \tag{4.31}$$

and μ and ν denote Cartesian axis; r_{ij}^{μ} and r_{ij}^{ν} are the μth and the νth Cartesian components of the vector between molecular positions $\mathbf{r}_{ij}$ and $\delta_{\mu\nu}$ is the Kronecker delta. The numerical evaluation of the second term on the right hand side of 4.30 was carried out using Ewald summation techniques [155–158].

We note that $\boldsymbol{\alpha}_i^{\mathrm{eff}}$ (Eq. 4.26) and $\boldsymbol{\alpha}_i$ (Eq. 4.29) are fundamentally different quantities: $\boldsymbol{\alpha}_i$ is an intrinsic molecular property depending solely on the electronic structure of the molecule in the system, that is if we carved a molecule out of the extended system, keeping its wavefunctions (and hence MLWFs) and coordinates unchanged and we placed it in vacuum, its polarizability would be exactly $\boldsymbol{\alpha}_i$. Instead $\boldsymbol{\alpha}_i^{\mathrm{eff}}$ is deduced from the projection of the total polarization onto each molecule. Thus it contains also induced polarization terms arising from the environment:

$$\boldsymbol{\alpha}_i^{\mathrm{eff}} = \boldsymbol{\alpha}_i + \boldsymbol{\alpha}_i^{\mathrm{DID}}, \tag{4.32}$$

where $\boldsymbol{\alpha}_i^{\mathrm{DID}}$ denotes the DID contributions from the environment.

4.4 Results and Discussion

4.4.1 Calculated Raman Spectra: Comparison with Experiments

Our calculated isotropic and anisotropic Raman spectra are presented in Fig. 4.1a, together with several experimental results. In the case of the low-frequency anisotropic spectrum, we compared our results with measurements for hydrogenated water since it has been shown experimentally that the anisotropic spectra of H_2O and D_2O are essentially the same in this region [122].

As expected, the position of the calculated high-frequency O-D stretching band is red-shifted by about 200 cm^{-1} compared with the experimental band, and it is broader. This error is mainly due to the use of the semilocal functional PBE as discussed in Refs. [10, 20, 35] in the case of infrare spectra, where it was shown that the use of the PBE0 functional can greatly improve the description of the position of the water stretching band, as well as its width. The use of some van der Waals functionals (see Ref. [20]) may also lead to an improvement of the position of the band.

There are three main features in the measured isotropic Raman spectrum: two intense peaks at 2400 cm^{-1} and 2500 cm^{-1} and a shoulder at 2700 cm^{-1}. The intensities of the two peaks vary as a function of temperature from 0 to 100 °C [3, 4]. As shown in the upper panel of Fig. 4.1b, the peak at 2200 cm^{-1} arises mainly from intermolecular contributions, while the one at 2400 cm^{-1} has an intramolecular nature. Although the peak positions are red shifted in our simulations and their separation is overestimated, due to the use of the PBE functional, our results are consistent with experiments and with those of previous studies [16, 114, 143, 144]. In our calculations, the position of the shoulder is in better agreement with experiments than that of the two main peaks; this shoulder is predominantly determined by the vibrations of non-hydrogen-bonded (NHB) species, with both inter- and intramolecular

contributions. In the isotropic spectrum, in agreement with experiments [113], we did not detect any significant signal at frequencies lower than those of the stretching band.

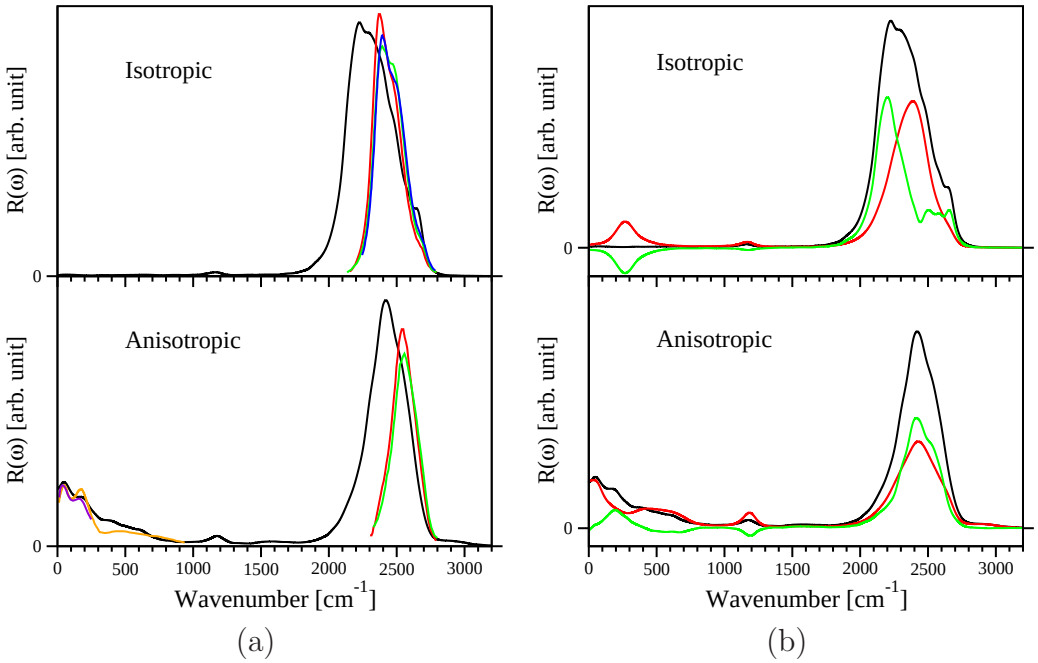


Figure 4.1: (a) Calculated isotropic (upper panel) and anisotropic (lower panel) Raman spectra (black lines) compared with several experiments. Red, blue and green curves are experimental Raman spectra of heavy water measured at 283K [3], 293K [4] and 303K [3] respectively. Orange and purple curves show experimental Raman spectra of hydrogenated water at 278K and 308K [5]. (b) Calculated isotropic (upper panel) and anisotropic (lower panel) Raman spectra (black lines) decomposed into intra- (red lines) and inter- (green lines) molecular contributions.

4.4.2 Low-Frequency Bands

In the region below 300 cm^{-1}, the isotropic Raman spectrum (upper panel of Fig. 4.1a) shows no features; the anisotropic spectrum (lower panel) exhibits instead two distinct peaks at 60 cm^{-1} and 200 cm^{-1}, in remarkable agreement with experiments [5] and at variance with previous theoretical studies [132, 136–138, 140] based on classical models. The slight deviation between theory and experiments (broad band over 300 cm^{-1}) in the librational region is likely due to the different mass of the hydrogen atom in experiments and calculations. The weak peak around 1200 cm^{-1} arises from the water bending mode. The origin of the peaks

at 60 cm^{-1} and 200 cm^{-1} has been greatly debated [145–147]. Some studies [112, 123, 146] assigned the peaks to transverse and longitudinal acoustic modes respectively, others to HB bending and stretching modes [112, 127]. Ref. [123] related the peak at 60 cm^{-1} to bifurcated HB. Recent *ab initio* simulations of the infrared spectrum of liquid water, where two peaks were detected at the same frequencies (although with different intensities), suggested that the 200 cm^{-1} peak arises from intermolecular charge fluctuations [23, 24] and from the presence of the tetrahedral HB network [23]. It was also suggested that the peak at 60 cm^{-1} corresponds to localized vibrations, while the peak at 200 cm^{-1} corresponds to coupling between HB molecules [63, 153].

We analyzed the 60 and 200 cm^{-1} peaks in terms of intra- and intermolecular contributions (Fig. 4.1b) and found that the former comes mainly from intramolecular vibrations, while the latter shows a significant intermolecular contribution, in agreement with previous studies [63, 153]. In addition, we found that at 270 cm^{-1}, positive intra- and negative intermolecular contributions cancel each other (see upper panel of Fig. 4.1b), with the total spectra showing no significant intensity. This indicates the presence of anti-correlated vibrations on different molecules, that are in anti-phase but have the same amplitude. As shown below such anticorrelation signal the presence of intermolecular charge fluctuations between HB water molecules, consistent with the correlation between the polarizability of the water molecules and the bond order of the HB O-D bond recently reported in Ref. [108].

In order to identify charge fluctuations, we used the sum of the spreads of the MLWFs centered on a given molecule to describe the charge transfer from that molecule to neighboring ones. The spread S_{w_n} of MLWFs is defined as [48]:

$$S_{w_n} = \langle w_n | r^2 | w_n \rangle - \langle w_n | \mathbf{r} | w_n \rangle^2, \tag{4.33}$$

and describes the spatial extension of each MLWFs. To analyze the spectral properties of the charge density, we computed the power spectrum of the spread using Eq. 4.21, where

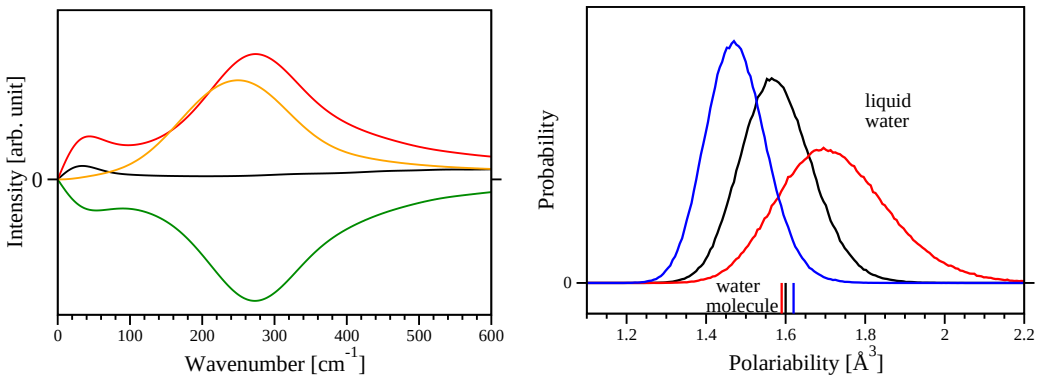


Figure 4.2: (a) The spectrum (black line) calculated from the spread of the MLWFs (see text) is decomposed into intra- (red line) and inter- (green line) molecular contributions. The vibrational density of states of the relative speed between oxygen atoms of HB water molecules is also shown (orange line). (b) Distribution of different components of the molecular polarizabilities of water molecules obtained in the simulation, compared with the polarizability of an isolated water molecule in different directions (vertical bars): black, red and blue lines represent polarizabilities along the dipole axis, the axis perpendicular to and the axis within the molecular plane, respectively.

we replaced $\bar{\alpha}$ by the sum of the spreads of the MLWFs. We further decomposed the spread power spectrum into intra- and intermolecular contributions, and we found (Fig. 4.2a) the same anticorrelation behavior at 270 cm^{-1}as observed in the isotropic Raman spectrum.

To understand the origin of the charge transfer feature at 270 cm^{-1}, we analyzed the vibrational density of states (VDOS) obtained from the relative speed between the oxygen atoms of HB molecular pairs:

$$I = \int dt e^{-i\omega t} \langle |\mathbf{v}_{\text{O-O}}(0)||\mathbf{v}_{\text{O-O}}(t)| \rangle. \tag{4.34}$$

Eq. 4.34 includes only the HB stretching vibrational motion, and rotations of water pairs are ignored. As shown in Fig. 4.2a, the resulting spectra exhibits a peak at 250 cm^{-1}, very close to 270 cm^{-1}, indicating a clear relation between HB stretching and intermolecular charge fluctuations.

In summary, our analysis showed that in liquid water at ambient conditions, intermolec-

ular HB stretching mode, centered at 270 cm^{-1}, are accompanied by intermolecular charge fluctuations. While there is no peak at this frequency in the isotropic Raman spectrum because of the anticorrelation nature of intermolecular charge-fluctuations, a peak is present in the anisotropic Raman spectrum and also in the infrared spectrum [23, 24], albeit at a slightly lower frequency of 200 cm^{-1}. This difference in frequency is probably caused by different selection rules determining isotropic Raman, anisotropic Raman and infrared spectra. We note that so far the study of charge transfer between water molecules had been restricted to gas-phase molecules and small clusters using either molecular-beam scattering experiments [159] or quantum chemistry calculations [160–163]. Although various studies including classical [24, 164] and *ab initio* [23, 108] simulations suggested the presence of intermolecular charge fluctuations in liquid water, most classical water models, such as SPCE [165, 166] and TIP4P [167], completely neglected this effect. Only very recently, water models that incorporated intermolecular charge transfer have been reported [168–170].

4.4.3 Molecular Polarizabilities

We now turn to the analysis of the molecular polarizabilities of water molecules in the liquid [107, 149]. We first compare them with the polarizability of water in the gas phase and we then estimate the contribution of molecular polarizabilities to Raman spectra. We then separate and analyze contributions from individual molecules and the environment .

The ensemble average of the isotropic molecular polarizability of water molecules computed from our simulations is 1.60, the same as for an isolated molecule. The experimental value of the molecular polarizability in the liquid is 1.47, again the same as in the gas phase; it was estimated from the measured refraction index n using the Lorentz-Lorenz equation: $\frac{n^2-1}{n^2+2} = \frac{4\pi}{3} N \bar{\alpha}_M$, where N is the number of molecule per unit volume and $\bar{\alpha}_M$ is the isotropic molecular polarizability. Although the calculated values overestimate experiments, due to the use of the PBE functional [99, 111], our results show that the molecular polarizability of water is the same in the liquid and for an isolated molecule, in agreement with experiments

and consistent with previous calculations [107, 108].

The distribution of the different components of the $\boldsymbol{\alpha}_i$ tensor computed for liquid water is shown in Fig. 4.2b. The polarizability projected on the molecular dipole axis is the same as in the isolated molecules, while that perpendicular to the molecular plane is significantly enhanced and the one on the in-plane axis decreased. The out-of-plane axis represents the direction least sensitive to the screening of the nuclei and thus the polarizability component on this axis is the most affected by the environment [107], and its distribution is broader. The in-plane component is, instead, smaller than in the gas phase and displays a narrower distribution. This indicates that, not surprisingly, the molecular polarizability of water is more anisotropic in the liquid than in an isolated molecule, consistent with previous classical simulations [132, 139, 140]. We note that while in qualitative agreement with results reported in the literature [107, 108], the distribution of $\boldsymbol{\alpha}_i$ shown in Fig. 4.2b is broader and the anisotropy is larger compared to previous studies. This is probably due to the use of different functionals (PBE versus BLYP).

We also note that many polarizable force fields adopted in the literature to simulate water (see a description of selected water models, e.g. in Ref. [171]) use a value of the polarizability smaller than the experimental one [101] in order to reproduce structural and diffusive properties of the liquid; in addition most often it is assumed that the polarizability tensor is isotropic [172], although models with different polarizability components [173] have been suggested. Hence no parameterization of existing force fields is consistent with the *ab initio* results reported here, and in previous first principles studies [107, 108]. Our first principles results may be useful for tuning the parameters of polarizable water models, although absolute values of polarizabilities should probably be obtained using hybrid functionals [99].

We now turn to analyzing the contribution of molecular polarizabilities to Raman intensities. We compared our results with the total Raman spectra, as shown in Fig. 4.3a. The difference between the two spectra represents the environmental part of the polarizability $\boldsymbol{\alpha}_i^{\mathrm{DID}}$. The $\boldsymbol{\alpha}_i$ contribution is approximately proportional to the total Raman spectra in the

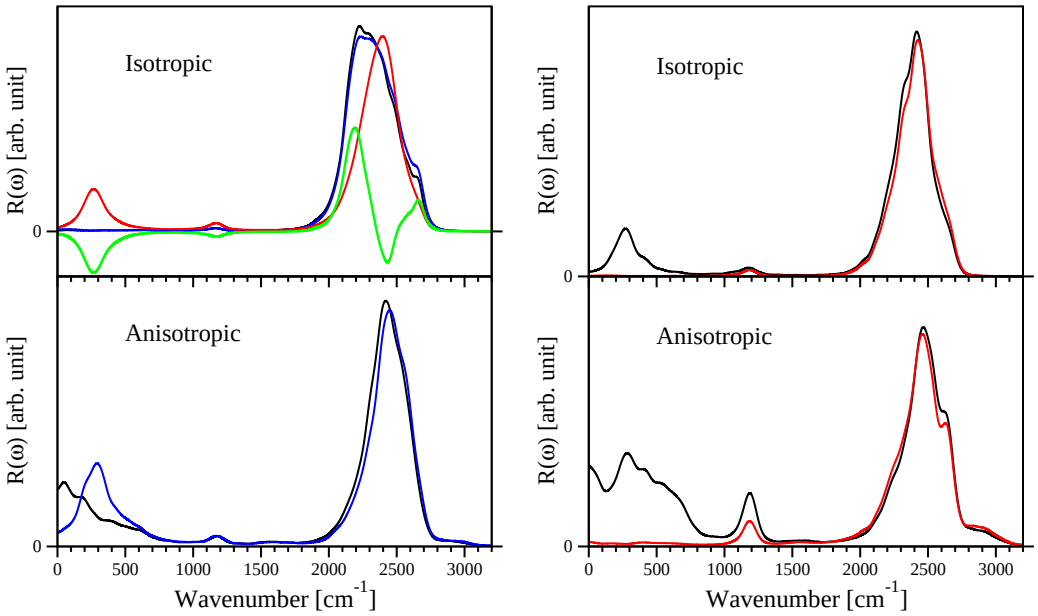


Figure 4.3: (a) Isotropic (upper panel) and anisotropic (lower panel) Raman spectra (scaled; black lines) compared with the spectra computed from molecular polarizabilities $\boldsymbol{\alpha}_i$ (blue lines) and their intra- (red lines) and inter- (green lines) molecular contributions (see text). (b) Raman spectra arising from the $\boldsymbol{\alpha}_i$ (black lines) and $\boldsymbol{\alpha}'_i$ (polarizability of a single, isolated water molecule at the geometry of the molecule in the liquid (see text); red lines)

bending and stretching regions over 1000 cm^{-1}, however it differs substantially from the total one below this frequency. This can be understood by observing (see Eqs. 4.26, 4.29, 4.30 and 4.31) that the difference between $\boldsymbol{\alpha}_i^{\text{eff}}$ and $\boldsymbol{\alpha}_i$ is a function of intermolecular distances, r_{ij} (Eq. 4.31). At high frequencies, r_{ij} can be assumed constant since water molecules move much slower than the O-D stretching motions, and $\boldsymbol{\alpha}_i$ is therefore proportional to $\boldsymbol{\alpha}_i^{\text{eff}}$. At low frequencies, however, r_{ij} may not be assumed constant, and the DID interactions induced by the collective molecular motions contribute to the difference between the spectra computed from $\boldsymbol{\alpha}_i^{\text{eff}}$ or $\boldsymbol{\alpha}_i$. This analysis suggests that the practice of assuming constant molecular polarizabilities of molecules, as in previous studies [132, 136, 137, 139, 140] of Raman spectra, may yield good predictions in the high-frequency stretching region, but rather inaccurate results for the low-frequency bands.

The low-frequency part of the anisotropic $\boldsymbol{\alpha}_i$ spectra (Fig. 4.3a) is particularly interest-

ing. It shows a peak at around ~ 300 cm^{-1}, i.e. in the same region where we identified intermolecular charge fluctuations. In fact, charge fluctuations are a major contribution to this peak, as charge transfer between molecules is responsible for a change of $\boldsymbol{\alpha}_i$, a mere molecular property. As shown in Fig. 4.3a, the intra- and intermolecular contributions of the isotropic $\boldsymbol{\alpha}_i$ spectrum also show an anticorrelation behavior, proving that the feature identified in the Raman spectra at 270 cm^{-1} does not arise solely from environmental effects (DID interactions), but from the fluctuation of molecular polarizabilities induced by intermolecular charge fluctuations. The peak at 200 cm^{-1} in the anisotropic Raman spectrum is instead a combination of intermolecular charge fluctuations, which contributes to $\boldsymbol{\alpha}_i$, and of environmental effects. Since no feature is present in the $\boldsymbol{\alpha}_i$ spectra at 60 cm^{-1}, we conclude that the peak at this frequency mostly originates from DID interactions, which correspond to collective motions of the system.

In order to understand how molecular polarizabilities are modified by the environment in the liquid, we computed the polarizabilities of a single isolated water molecule at the geometry of the molecule in the liquid. We call such polarizability $\boldsymbol{\alpha}'_i$. Therefore, the difference between $\boldsymbol{\alpha}'_i$ and $\boldsymbol{\alpha}_i$ accounts for the difference in electronic structure induced by the environment. Because of the high computational cost, we calculated $\boldsymbol{\alpha}'_i$ for several but not all water molecules in our simulation, and we computed the $\boldsymbol{\alpha}'_i$ spectra by replacing $\boldsymbol{\alpha}$ by $\boldsymbol{\alpha}'_i$ in Eqs. 4.21 and 4.22. In Fig. 4.3b, we show the comparison of the $\boldsymbol{\alpha}_i$ and $\boldsymbol{\alpha}'_i$ spectra for only one of the water molecules in our sample. While the two spectra agree well in the high-frequency O-D stretching region, we observed clear differences in the low frequency region. The difference between the two spectra is therefore caused by the change of the electronic structure of the water molecules by the environment in the liquid, namely by intermolecular charge fluctuations.

4.5 Summary and Conclusions

In summary, we computed the Raman spectra of liquid heavy water from *ab initio* MD simulations using DFPT. These calculations represent the first *ab initio* study of the Raman spectrum of a molecular liquid. Our results are in good agreement with experiments, especially in the low-frequency region of the anisotropic spectrum. We analyzed the Raman spectra by decomposing the intensities into intra- and intermolecular contributions using MLWFs. In the case of the high-frequency O-D stretching band of the isotropic spectrum, our findings indicate that the intense feature found experimentally at 2500 cm^{-1} arises from intramolecular vibrations, while the one at 2400 cm^{-1} stems from intermolecular coupling; the higher frequency shoulder at 2700 cm^{-1}originates from NHB species, with both inter- and intramolecular contributions. Overall our results for the high frequency band are consistent with previous studies [143, 144].

In the low-frequency part of the Raman spectra, we identified the presence of intermolecular charge fluctuations at 270 cm^{-1}, accompanying intermolecular HB stretching modes. To support such an identification, we combined a decomposition analysis in terms of intra- and intermolecular contributions, with that of the spread of MLWFs and of molecular polarizabilities. In the anisotropic Raman spectrum, we found that the peak at 200 cm^{-1} has an intermolecular nature and it originates partly from the charge fluctuations and partly from environmental effects. The peak at 60 cm^{-1} is instead due to intramolecular modes determined by DID interactions.

Overall our simulations showed the importance of *ab initio* methods to accurately describe and interpret the Raman spectra of liquid water, especially in the low-frequency region where intermolecular charge transfer is involved.

Following previous studies [107, 108, 149], we defined the molecular polarizabilities of water molecules in the liquid, and we found that its average value is the same in the liquid and in an isolated molecule, while the polarizability anisotropy of water molecules is enhanced in the liquid. These results may be helpful for the parameterization of polarizable force

fields of liquid water. We also investigated the contributions of the molecular polarizability $\boldsymbol{\alpha}_i$ to the total Raman spectra and compared it with the spectra computed using $\boldsymbol{\alpha}'_i$, i.e. the polarizability of water molecules extracted from the liquid with their geometry kept unchanged. The comparison showed that using constant polarizabilities (or $\boldsymbol{\alpha}'_i$ to replace $\boldsymbol{\alpha}_i$) in computing Raman spectra, as in many classical simulations [132, 136, 137, 139–141], may be a good approximation for predicting the stretching band of the Raman spectra, but not for the low-frequency region.

Our simulation elucidated the role of intermolecular charge fluctuations in the low-frequency Raman spectra and provided a comprehensive interpretations of the spectra in this region. We note that our *ab initio* methods to calculate and interpret the Raman spectra can be readily used to study aqueous solutions and in general other non-metallic systems. In addition, only simple modifications are needed to extend our methods to study other vibrational probes such as the surface specific sum frequency vibrational spectroscopy [27, 174].

CHAPTER 5

A FIRST-PRINCIPLES FRAMEWORK TO COMPUTE SUM-FREQUENCY GENERATION VIBRATIONAL SPECTRA OF SEMICONDUCTORS AND INSULATORS

5.1 Introduction

The determination of the atomistic structure of surfaces and interfaces is a central problem in materials physics and in nanoscience [175, 176]. Among surface sensitive probes, non-linear optical spectroscopies such as sum-frequency generation (SFG) have been widely used for decades to study the structure and dynamics of surfaces and interfaces [26, 27]. However the interpretation of experimental data is often difficult and, in many instances, controversial [29, 177, 178]; in addition most theoretical models adopted to interpret experiments have been based on multiple approximations and on the use of empirical force fields [174, 179].

In an SFG experiment (see Fig. 5.1), an interface is illuminated by two beams, of frequency ω_1 in the visible and ω_2 in the infrared range, and the emitted sum-frequency (SF) light with frequency $\omega_s = \omega_1 + \omega_2$ is detected. The total emitted SFG intensity is proportional to the square of the complex effective second-order nonlinear (or SFG) susceptibility $\chi^{(2)}_{\text{eff}}$ of the medium. Indeed, within the dipole approximation, the SFG signal of centrosymmetric bulk regions vanishes, and the intensity of the emitted SF light originates only from atoms or molecules present at the interface, where centrosymmetry is broken. We note that it is the absence of quadrupole and higher order contributions that ultimately determines the surface specificity of SFG spectroscopy. Experimentally, it is difficult to validate the dipole approximation, and to establish whether quadrupole contributions are negligible in SFG signals; only part of quadrupole contributions may be identified experimentally, by com-

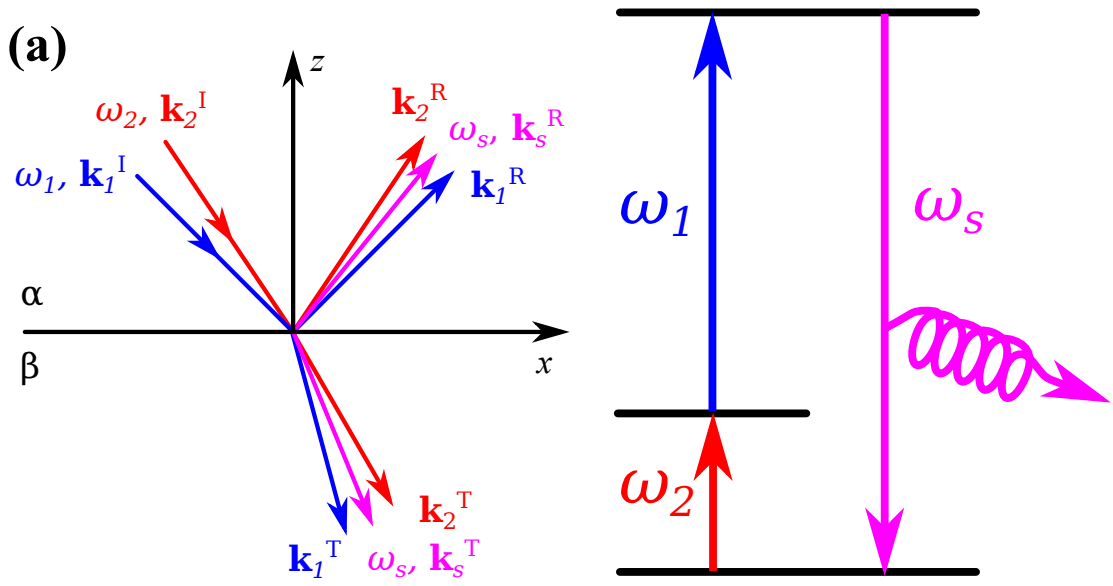


Figure 5.1: Representative geometry of of a Sum Frequency Generation (SFG) experiment (left panel) where an interface between two media α and β is illuminated by two light beams of frequency ω_1 and ω_2 (the corresponding wavevectors $\mathbf{k}$ are also indicated); I, R and T indicate incident, reflected and transmitted light, respectively. The emitted SF light has frequency $\omega_s = \omega_1 + \omega_2$ (right panel).

paring the intensities detected in transmission and reflectance optical geometries [180–182]. Hence to fully understand and interpret SFG signals, a robust and predictive computational framework is required.

The theoretical foundation of SFG vibrational spectroscopy dates back more than two decades [181–184]; nevertheless calculations of SFG intensities (including quadrupole contributions) and comparisons between computed and measured spectra remain challenging tasks for several reasons: (i) The explicit inclusion of quadrupole contributions requires the calculation of both interfacial and bulk terms to obtain the total susceptibility (unlike the case of the dipole contribution, where only interfacial terms are to be evaluated). In atomistic simulations using slab geometries to model interfaces, the consistent and non-empirical evaluation of bulk and interfacial terms is an extremely demanding task, from a computational standpoint. (ii) The calculation of quadrupole contributions requires the evaluation of quadrupole moments, whose calculation requires special care in the case of dipolar molecules, in order

to avoid origin dependent results. (iii) The evaluation of SFG susceptibilities encompasses the calculation of local dielectric functions, necessary to properly evaluate the variation of the electric field at the interface; in many studies such variation was simply ignored.

Recently, Shiratori and Morita [98] proposed a computational framework for the evaluation of SFG spectra, using mixed quantum and molecular mechanics methods to compute dipole and quadrupole contributions from molecular non-linear polarizabilities; they considered local field correction factors to account for the variation of the electric field at the interface, hence effectively including a spatial dependent dielectric function in their calculations. These authors also suggested a way to solve the origin dependence problem encountered in the calculation of quadrupole moments of dipolar systems. However their method was applied only to a non-polar molecular liquid [185]. Byrnes *et al.* [186] proposed instead a technique relying on cancellation between interface and bulk contributions, where the variation of the electric field at the interface was ignored. As most of the theoretical investigations on SFG appeared in the literature[174], the study of Ref. [185] used empirical force fields. To our knowledge only one *ab initio* simulation of SFG spectra has been reported so far, on the surface of liquid water [42], which however contained numerous approximations. The authors did not include quadrupole contributions, nor the variation of the electric field at the surface; in addition, they made use of possibly severe approximations in evaluating the time correlation functions (TCF) [42].

In this chapter, we propose a first principles theoretical framework to predict SFG spectra, which is based on density functional theory and includes both dipole and quadrupole contributions arising from surface and bulk regions. We computed SFG susceptibilities using Maximally Localized Wannier functions (MLWFs) [48, 93], within a general formulation valid for semiconductors and insulators. Our framework included the evaluation of local dielectric functions to account for the variation of the electric field across the interface, and the calculation of quadrupole moments with a technique yielding origin independent results. We present an application of the method to the study of the SFG spectra of the ice I_h

basal surface and discuss the importance of considering the local dielectric constant and quadrupole contributions when interpreting experiments. [7, 8, 28, 187–189]. To our knowledge our study represents the first *ab initio* simulation of SFG spectra including quadrupole contributions and the electric field variations at the interface, as well as the first *ab initio* simulation of the SFG spectra of ice [187, 190].

5.2 Basic Theories of SFG Spectroscopy

Both infrared and Raman transitions are involved in an SFG process (see Fig. 5.1): a system in its vibrational ground state first absorbs an infrared photon of frequency ω_2 and it is then further excited to a virtual state by absorbing a visible photon (ω_1). When the system decays, it emits a photon of frequency $\omega_s = \omega_1 + \omega_2$. The SFG susceptibility tensor $\boldsymbol{\chi}^{(2)}_{\text{eff}}$ relates the effective induced polarization of the system $\mathbf{P}^{(2)}_{\text{eff}}(\omega_s)$ to the electric fields of the incident beams, $\mathbf{E}^0(\omega_1)$ and $\mathbf{E}^0(\omega_2)$:

$$P^{(2)}_{\text{eff},\nu}(\omega_s) = \chi^{(2)}_{\text{eff},\nu\mu\xi} E^0_\mu(\omega_1) E^0_\xi(\omega_2), \tag{5.1}$$

where $\nu\mu\xi\zeta$ denote Cartesian coordinates. For a specific combination of incident angles and polarization directions of the beams, $\chi^{(2)}_{\text{eff}} = \mathbf{e}_s \cdot \boldsymbol{\chi}^{(2)}_{\text{eff}} : \mathbf{e}_1\mathbf{e}_2$, where $\mathbf{e}_s$, $\mathbf{e}_1$ and $\mathbf{e}_2$ are the polarization directions of the SF, visible and infrared lights, respectively. Experiments usually report $\chi^{(2)}_{\text{eff}}(\omega_2)$, and ω_s and ω_1 are kept constant. Hereafter for simplicity we do not explicitly show the frequency dependency (ω_2) of the second-order nonlinear susceptibilities.

SFG spectra are surface-specific probes if $\chi^{(2)}_{\text{eff}}$ is determined only by dipolar transitions. If quadrupole transitions are included to describe one of the three processes shown in Fig. 5.1, then $\boldsymbol{\chi}^{(2)}_{\text{eff}} \simeq \boldsymbol{\chi}^{\text{D}}_{\text{eff}} + \boldsymbol{\chi}^{\text{Q}}_{\text{eff}}$ where in this expression non-resonant background, magnetic multipole terms and higher electric multipole terms are neglected. The dipole contribution $\boldsymbol{\chi}^{\text{D}}_{\text{eff}}$ describes the response to a uniform electric field perturbations while $\boldsymbol{\chi}^{\text{Q}}_{\text{eff}}$ includes the response to electric field gradients.

5.2.1 Interfacial and Bulk Contributions

Using an approach similar to that of Ref. [98], we further decompose $\boldsymbol{\chi}^{\mathrm{D}}_{\mathrm{eff}}$ and $\boldsymbol{\chi}^{\mathrm{Q}}_{\mathrm{eff}}$ into interfacial (I) and bulk (B) contributions: $\boldsymbol{\chi}^{\mathrm{D}}_{\mathrm{eff}} = \boldsymbol{\chi}^{\mathrm{ID}}_{\mathrm{eff}} + \boldsymbol{\chi}^{\mathrm{BD}}_{\mathrm{eff}}$, and $\boldsymbol{\chi}^{\mathrm{Q}}_{\mathrm{eff}} = \boldsymbol{\chi}^{\mathrm{IQ}}_{\mathrm{eff}} + \boldsymbol{\chi}^{\mathrm{BQ}}_{\mathrm{eff}}$. Note that $\boldsymbol{\chi}^{\mathrm{BD}}_{\mathrm{eff}} = 0$ for an isotropic bulk system, e.g. bulk ice I_h. This decomposition is conceptually not necessary but of key importance in any simulation with slab geometries to represent interfaces or surfaces: computing both bulk and interface terms in the same simulation cell with periodic boundary conditions only in the x and y direction (z is perpendicular to the slab) would require samples prohibitively large and simulation time prohibitively long for any *ab initio* simulation.

The ID term $\boldsymbol{\chi}^{\mathrm{ID}}_{\mathrm{eff}}$ is often considered the major contribution to $\boldsymbol{\chi}^{(2)}_{\mathrm{eff}}$ and has been the focus of most studies in the literature [174, 179]:

$$\chi^{\mathrm{ID}}_{\mathrm{eff},\nu\mu\xi} = \int_{z_b}^{\infty} dz L_{\nu\nu}(\omega_s, z)\chi^{\mathrm{D}}_{\nu\mu\xi}(z)L_{\mu\mu}(\omega_1, z)L_{\xi\xi}(\omega_2, z), \tag{5.2}$$

where we assumed that the system is homogeneous in the x and y directions. The position z_b is supposed to be well beneath the surface so that the integral in Eq. 5.2 includes the whole interfacial area. $\boldsymbol{\chi}^{\mathrm{D}}(z)$ is the second-order nonlinear susceptibility for dipole transitions at position z (see below). We also assumed that the wavelength of the light is much larger than the size of the interfacial region so that the phase associated to the electric field can be neglected. The Fresnel coefficients $\mathbf{L}(\omega, z)$ are the elements of a diagonal tensor which relates the electric field at position z, $\mathbf{E}(\omega, z)$, at the interface to $\mathbf{E}^0(\omega)$:

$$\mathbf{E}(\omega, z) = \mathbf{L}(\omega, z)\mathbf{E}^0(\omega). \tag{5.3}$$

The explicit expression of the Fresnel coefficients are:

$$
\begin{aligned}
L_{xx}(\omega) &= \frac{2\epsilon^{\alpha}(\omega)k_z^{\beta}(\omega)}{\epsilon^{\beta}(\omega)k_z^{\alpha}(\omega)+\epsilon^{\alpha}(\omega)k_z^{\beta}(\omega)} \\
L_{yy}(\omega) &= \frac{2k_z^{\alpha}(\omega)}{k_z^{\alpha}(\omega)+k_z^{\beta}(\omega)} \\
L_{zz}(\omega,z) &= \frac{2\epsilon^{\alpha}(\omega)\epsilon^{\beta}(\omega)k_z^{\alpha}(\omega)}{\epsilon^{\beta}(\omega)k_z^{\alpha}+\epsilon^{\alpha}(\omega)k_z^{\beta}(\omega)}\frac{1}{\epsilon_z(\omega,z)}.
\end{aligned}
\tag{5.4}
$$

Here $\epsilon^{\gamma}(\omega)$ and $\mathbf{k}^{\gamma}(\omega)$, $\gamma=\alpha,\beta$, are the dielectric constant and light wavevector in the bulk region of phase γ (the interface is between two systems, α and β, see Fig. 5.1). The coefficient $\mathbf{L}(\omega)$ also depends on the incident angles of the light beams and its value depends on the optical geometry chosen in experiments. Only $L_{zz}(\omega,z)$ is position dependent, through the function $1/\epsilon_z(\omega,z)$, while $L_{xx}(\omega)$ and $L_{yy}(\omega)$ are position independent [7]. Our definition of Fresnel coefficients is slightly different from the definition in Ref. [7], which is based on the local field correction factors. We note that Fresnel coefficients, and $\epsilon(\omega,\mathbf{r})$, are of key importance to account for the variation of the electric field across the interface.

The interfacial quadrupole (IQ) contribution includes response to the electric field as well as the electric field gradient. The latter, $\nabla\mathbf{E}(\omega,\mathbf{r})$, comes solely from the gradient of the Fresnel coefficient, $d\mathbf{L}(\omega,z)/dz$ (see Eq. 5.3). The expression of the interfacial quadrupole contributions is [98, 181]:

$$
\begin{aligned}
\chi^{\mathrm{IQ}}_{\mathrm{eff},\nu\mu\xi} = \int_{z_b}^{\infty} dz \Big\{ & L_{\nu\nu}(\omega_s,z)\chi^{\mathrm{Q_1}}_{\nu\mu\xi z}(z)\frac{d}{dz}L_{\mu\mu}(\omega_1,z)L_{\xi\xi}(\omega_2,z) \\
& + L_{\nu\nu}(\omega_s,z)\chi^{\mathrm{Q_2}}_{\nu\mu\xi z}(z)L_{\mu\mu}(\omega_1,z)\frac{d}{dz}L_{\xi\xi}(\omega_2,z) \\
& + \frac{d}{dz}L_{\nu\nu}(\omega_s,z)\chi^{\mathrm{Q_s}}_{\nu\mu\xi z}(z)L_{\mu\mu}(\omega_1,z)L_{\xi\xi}(\omega_2,z)\Big\} \; using \\
& + L_{\nu\nu}(\omega_s,z_b)\chi^{\mathrm{Q_s,B}}_{\nu\mu\xi z}(z_b)L_{\mu\mu}(\omega_1,z_b)L_{\xi\xi}(\omega_2,z_b),
\end{aligned}
\tag{5.5}
$$

where $\boldsymbol{\chi}^{\mathrm{Q_1}}(z)$, $\boldsymbol{\chi}^{\mathrm{Q_2}}(z)$ and $\boldsymbol{\chi}^{\mathrm{Q_s}}(z)$ are quadrupole second-order nonlinear susceptibilities, which are rank-4 tensors, and describe the response to both the electric field and the electric

field gradient at position z. We note that the first three terms on the rhs arise from the electric field (dielectric constant) gradient across the interface, and we call the sum of these terms $\boldsymbol{\chi}_{\text{eff}}^{\text{IQI}}$. The fourth terms on the rhs, which arises from integration by part, depends solely on the property of the bulk region, despite the fact that it is obtained as part of the interfacial contribution. We call this contribution $\boldsymbol{\chi}_{\text{eff}}^{\text{IQB}}$. Therefore we have $\boldsymbol{\chi}_{\text{eff}}^{\text{IQ}} = \boldsymbol{\chi}_{\text{eff}}^{\text{IQI}} + \boldsymbol{\chi}_{\text{eff}}^{\text{IQB}}$ where:

$$\begin{aligned}\chi_{\text{eff},\nu\mu\xi}^{\text{IQI}} \equiv \int_{int} dz \Big\{ & L_{\nu\nu}(\omega_s, z)\chi_{\nu\mu\xi z}^{\text{Q}_1}(z)\frac{d}{dz}L_{\mu\mu}(\omega_1, z)L_{\xi\xi}(\omega_2, z) \\ & + L_{\nu\nu}(\omega_s, z)\chi_{\nu\mu\xi z}^{\text{Q}_2}(z)L_{\mu\mu}(\omega_1, z)\frac{d}{dz}L_{\xi\xi}(\omega_2, z) \\ & + \frac{d}{dz}L_{\nu\nu}(\omega_s, z)\chi_{\nu\mu\xi z}^{\text{Q}_\text{s}}(z)L_{\mu\mu}(\omega_1, z)L_{\xi\xi}(\omega_2, z)\Big\},\end{aligned} \tag{5.6}$$

$$\chi_{\text{eff},\nu\mu\xi}^{\text{IQB}} \equiv L_{\nu\nu}(\omega_s, z_b)\chi_{\nu\mu\xi z}^{\text{Q}_\text{s}}(z_b)L_{\mu\mu}(\omega_1, z_b)L_{\xi\xi}(\omega_2, z_b). \tag{5.7}$$

When computing the bulk quadrupole contribution (BQ), we can simplify its expression by considering $\mathbf{L}(\omega, \mathbf{r})$ as a position independent quantity, since $\epsilon_z(\omega, z)$ can be approximated as the bulk dielectric constant in the isotropic bulk region (see Eq. 5.4). Because the light wavelength are smaller than the size of the bulk region, the electric field phase factor $\exp(i\mathbf{k}\cdot\mathbf{r})$ has to be explicitly considered in the evaluation of $\boldsymbol{\chi}_{\text{eff}}^{\text{BQ}}$. The expression of the bulk quadrupole contributions for both reflection (R) and transmission (T) geometries (G = R, T) is [98, 181]:

$$\begin{aligned}\chi_{\text{eff},\nu\mu\xi}^{\text{BQ}} = & \frac{1}{k_{1,z}^{\text{T}} + k_{2,z}^{\text{T}} - k_{s,z}^{\text{G}}} L_{\nu\nu}(\omega_s)L_{\mu\mu}(\omega_1)L_{\xi\xi}(\omega_2) \\ & \sum_{\zeta} \chi_{\nu\mu\xi\zeta}^{\text{Q}_1}(z_b)k_{1,\zeta}^{\text{T}} + \chi_{\nu\mu\xi\zeta}^{\text{Q}_2}(z_b)k_{2,\zeta}^{\text{T}} \\ & - \chi_{\nu\mu\xi\zeta}^{\text{Q}_\text{s}}(z_b)(k_{1,\zeta}^{\text{T}} + k_{2,\zeta}^{\text{T}}).\end{aligned} \tag{5.8}$$

Note that due to the different signs of $k_{s,z}^{\text{R}}$ and $k_{s,z}^{\text{T}}$ (see Fig. 5.1), $\boldsymbol{\chi}_{\text{eff}}^{\text{BQ}}$ is different for reflection and transmission optical geometries while $\boldsymbol{\chi}_{\text{eff}}^{\text{ID}}$ and $\boldsymbol{\chi}_{\text{eff}}^{\text{IQ}}$ are the same. Therefore $\boldsymbol{\chi}_{\text{eff}}^{\text{BQusing}}$ can be experimentally determined by comparing spectra measured in the two different optical geometries [182]. On the contrary the $\boldsymbol{\chi}_{\text{eff}}^{\text{IQB}}$ term *cannot* be determined

experimentally.

In sum in our calculation the SFG susceptibility is decomposed as:

$$\boldsymbol{\chi}^{(2)}_{\text{eff}} = \boldsymbol{\chi}^{\text{ID}}_{\text{eff}} + \boldsymbol{\chi}^{\text{IQI}}_{\text{eff}} + \boldsymbol{\chi}^{\text{IQB}}_{\text{eff}} + \boldsymbol{\chi}^{\text{BQ}}_{\text{eff}}, \tag{5.9}$$

where the first two terms on the rhs are obtained by carrying out calculations in slab geometries and the last two terms are instead evaluated by bulk calculations with 3D periodicity.

5.2.2 The Calculation of the linear and nonlinear susceptibilities

In order to compute each contributions to $\boldsymbol{\chi}^{(2)}_{\text{eff}}$ from first principles, we first computed the local frequency-dependent dielectric constant $\boldsymbol{\epsilon}(\omega, z)$ entering the definition of the Fresnel coefficients (Eq. 5.4) and hence the first-order linear susceptibility $\boldsymbol{\chi}^{\text{L}}(\omega)$:

$$\boldsymbol{\epsilon}(\omega, z) = \boldsymbol{\epsilon}_{\infty}(z) + 4\pi\boldsymbol{\chi}^{\text{L}}(\omega, z), \tag{5.10}$$

where $\boldsymbol{\epsilon}_{\infty}(z)$ is the local high-frequency dielectric constant used to approximate $\boldsymbol{\epsilon}(\omega_s, z)$ and $\boldsymbol{\epsilon}(\omega_1, z)$ since the dispersion of $\boldsymbol{\epsilon}(\omega, z)$ may be neglected in the visible frequency range. Ignoring anharmonic effects, all the linear and nonlinear susceptibilities share the same form:

$$\boldsymbol{\chi}^{\text{X}}(\omega, z) = \sum_q \frac{\mathbf{A}^{\text{X}}_q(z)}{\omega - \omega_q + i\Gamma_q}, \tag{5.11}$$

where $\text{X} = \text{L}, \text{D}, \text{Q}_1, \text{Q}_2, \text{Q}_\text{s}$ and ω_q, Γ_q and $\mathbf{A}^{\text{X}}_q(z)$ are the frequency, lifetime and amplitude of the qth vibrational normal mode. Hence the calculation of total SFG signals reduces to the calculation of $\mathbf{A}^{\text{X}}_q(z)$, which are computed using quantities defined within the entire system (slab or bulk models), not in terms of molecular quantities, such as molecular polarizabilities

[98]:

$$
\begin{aligned}
A^{\mathrm{L}}_{q,\nu\mu}(z) &= -\frac{1}{2\omega_q V}\frac{\partial M_\nu}{\partial S_q}\frac{\partial M_\mu}{\partial S_q} \\
A^{\mathrm{D}}_{q,\nu\mu\xi}(z) &= -\frac{1}{2\omega_q V}\frac{\partial M_\xi}{\partial S_q}\frac{\partial \Pi_{\nu\mu}}{\partial S_q} \\
A^{\mathrm{Q_1}}_{q,\nu\mu\xi\zeta}(z) &= -\frac{1}{2\omega_q V}\frac{\partial M_\xi}{\partial S_q}\frac{\partial(\partial Q_{\zeta\mu}/\partial E_\nu)}{\partial S_q} \\
A^{\mathrm{Q_2}}_{q,\nu\mu\xi\zeta}(z) &= -\frac{1}{2\omega_q V}\frac{\partial Q_{\xi\zeta}}{\partial S_q}\frac{\partial \Pi_{\nu\mu}}{\partial S_q} \\
A^{\mathrm{Q_s}}_{q,\nu\mu\xi\zeta}(z) &= -\frac{1}{2\omega_q V}\frac{\partial M_\xi}{\partial S_q}\frac{\partial(\partial Q_{\zeta\nu}/\partial E_\mu)}{\partial S_q},
\end{aligned}
\tag{5.12}
$$

where S_q is the qth normal mode, $\mathbf{E}$, $\mathbf{M}$, $\mathbf{Q}$ and $\mathbf{\Pi}$ are the total electric field, dipole, quadrupole and effective polarizability, respectively, within a volume V in the vicinity of the position z. We note that $\mathbf{\Pi}$ represents the electric dipolar response to the total electric field $\mathbf{E}$, $\delta M_\nu = \Pi_{\nu\mu}\delta E_\mu$, which is to be distinguished from the response to the local electric field, represented by the molecular polarizabilities (used, e.g., in Ref. [98]). The expressions of the $\mathbf{A}^{\mathrm{X}}_q(z)$ coefficients may be recast into TCFs and computed using molecular dynamics simulations [98]. It is important to choose the volume V big enough to obtain a smooth susceptibility profile but small compared to the length scale of the entire system so as to include the variation of $\mathbf{E}$. In a system with well defined molecular units, it is reasonable to choose V to include only one molecule, e.g. one water molecule in the ice system studied here. using Four MLWFs are associated to each water molecule, if only valence electrons ($1s$ for H and $2s^2$ and $2p^4$ for oxygen) are considered. In practice we used V to be the average volume of a molecule in the bulk material ($V = V_{cell}/N_{mol} = 32.7$ Å^3 for ice I$_h$), to approximate V for both the slab and the bulk system assuming that the volume of water molecules does not change significantly in bulk and surface environments. Then $\mathbf{M}$ reduces to a molecular dipole moment that can be computed using MLWFs [23, 48, 62, 93], and $\mathbf{\Pi}$ reduces to an effective molecular polarizability, as defined in Ref. [2]. The latter can be evaluated efficiently by projecting the polarizability obtained from DFPT onto MLWFs (see Chapter 4). The quadrupole moment $\mathbf{Q}$ can be computed using the formula reported in the

Appendix of Ref. [91]. However, instead of the position operator [87, 191] used in Ref. [91], we adopted a real-space scheme [49] to evaluate the position and second moment of the nth MLWF, $\langle r_\mu \rangle_n$ and $\langle r_\mu r_\nu \rangle_n$ (see Section 3.2). The latter quantitie converges much faster with respect to the size of the simulation cell, and their value is more accurate for the system sizes used in our study [49], than with the scheme used in Ref. [91]. The choice of the origin in the quadrupole moment calculation will be discussed below. In practice, we used the same origin to compute the quadrupole moment for all molecules in the system (see Section 5.4.6). We calculated the derivatives of quadrupole moments $\partial Q_{\zeta\nu}/\partial E_\mu$ using finite differences (FD) and an electric enthalpy functional [45].

Using MLWFs, we could obtain the value of $\mathbf{A}_q^{\mathrm{X}}(z)$ for each molecule by summing over all MLWFs in one molecule. The intermolecular contributions [2, 35] were properly included in our calculation. For example, the amplitude for the linear susceptibility, including intra- and intermolecular contributions of the ith molecule, is

$$A^{\mathrm{L}}_{q,i,\nu\mu} = -\frac{1}{2\omega_q V} \sum_j \frac{\partial M_{i,\nu}}{\partial S_q} \frac{\partial M_{j,\mu}}{\partial S_q},$$

where $\mathbf{M}_i$ is the dipole of the ith molecule, and the summation is performed over all molecules in the system. Similar expressions were used to compute the amplitudes in Eq. 5.12 for each molecule. The linear and nonlinear susceptibilities and the local dielectric constant can be calculated, for each molecule, from the corresponding amplitudes using Eqs. 5.10 and 5.11. The Fresnel coefficient tensor can be defined for each molecules as well, as it is directly related to the local dielectric constant (see Eq. 5.4). Thus all quantities needed to compute the ID, IQ and BQ contributions (see Eqs. 5.2, 5.6, 5.7 and 5.8) can be defined for each molecule. When computing these contributions, we performed summations (for ID and IQI) or averages (for IQB and BQ) over the corresponding molecular quantities instead of carrying out an explicit integration or an average over the position z, respectively.

5.2.3 The Origin Dependence Problem in the Calculation of Quadrupole Moments

The calculation of the quadrupole moment in $\mathbf{A}_q^{\mathrm{X}}(z)$ requires special care. For a system with zero dipole moment, the total quadrupole moment tensor $Q_{\nu\mu} = \frac{1}{2}\int \rho(\mathbf{r}) r_\nu r_\mu d\mathbf{r}$ ($\rho(\mathbf{r})$ is the total charge density of the system) is well defined and origin independent. Likewise, only if $\boldsymbol{\chi}^{\mathrm{D}} = 0$, second-order nonlinear susceptibilities for quadrupole transitions are well defined and origin independent. Hence to properly compute quadruple contributions from isotropic bulk regions, we used the total charge density of the supercell; we avoided summations over quadrupole contributions associated to single molecular units, at variance from several formulations presented in the literature [98, 186]. The essence of our method is that we consider the entire simulation cell as the smallest unit in the calculation, instead of individual water molecules. It is key to use properties of the entire system, such as the total charge density, to evaluate $\boldsymbol{\chi}^{\mathrm{Q1}}$, $\boldsymbol{\chi}^{\mathrm{Q2}}$ and $\boldsymbol{\chi}^{\mathrm{Qs}}$, thus insuring that, as long as the entire sample under consideration is centrosymmetric ($\boldsymbol{\chi}^{\mathrm{D}} = 0$), the results are origin independent.

In principle, the exact same formulation as adopted here for ice (see below) could be applied to disordered systems consisting of polar molecules, such as liquid water. Indeed in an isotropic liquid $\boldsymbol{\chi}^{\mathrm{D}} = 0$ and hence $\boldsymbol{\chi}^{\mathrm{Q1}}$, $\boldsymbol{\chi}^{\mathrm{Q2}}$ and $\boldsymbol{\chi}^{\mathrm{Qs}}$ are guaranteed to be origin independent when a fixed origin in the lab coordinates is used, instead of a moving one such as the center of mass of molecules. However in practice for a liquid, the second-order nonlinear susceptibilities need to be defined in terms of TCFs; obtaining $\chi^{\mathrm{D}}_{\nu\mu\xi} \propto \int \langle M_\xi(t)\Pi_{\nu\mu}(t+\Delta t)\rangle dt \exp(i\omega_2\Delta t) \simeq 0$ (and hence recovering the origin independence of $\boldsymbol{\chi}^{\mathrm{Q1}}$ $\boldsymbol{\chi}^{\mathrm{Q2}}$ and $\boldsymbol{\chi}^{\mathrm{Qs}}$) requires simulations of the order of nanoseconds, which are not yet affordable with *ab initio* calculations.

5.2.4 Polarization Combinations

In experimental measurements, specific polarization directions of the SF, visible and infrared light beams are usually reported. Commonly used polarization combinations include *ssp* and *ppp*, where *s*-polarization is in the y direction and *p*-polarization is in the xz plane (see Fig. 5.1). The explicit expressions for $\chi^{(2)}_{\text{eff}}$ in terms of different polarization combinations are [7]:

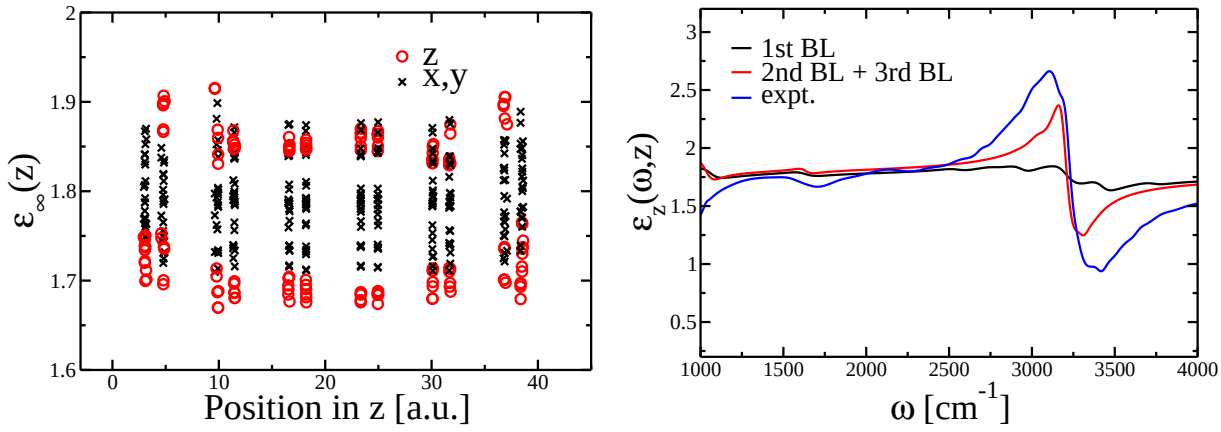
$$\begin{aligned}
\chi^{(2)}_{\text{eff},ssp} =& \sin(\theta_2)\chi^{(2)}_{\text{eff},yyz} \\
\chi^{(2)}_{\text{eff},ppp} =& \sin(\theta_s)\sin(\theta_1)\sin(\theta_2)\chi^{(2)}_{\text{eff},zzz} \\
&- \cos(\theta_s)\cos(\theta_1)\sin(\theta_2)\chi^{(2)}_{\text{eff},xxz} \\
&+ \sin(\theta_s)\cos(\theta_1)\cos(\theta_2)\chi^{(2)}_{\text{eff},zxx} \\
&- \cos(\theta_s)\sin(\theta_1)\cos(\theta_2)\chi^{(2)}_{\text{eff},xzx},
\end{aligned} \tag{5.13}$$

where θ_s, θ_1 and θ_2 are the incident angles of the SF, visible and infrared light, respectively. If we only consider the dipole contributions, the above expression becomes:

$$\chi^{(2)}_{\text{eff},ssp} = \int_{int} dz L_{yy}(\omega_s)L_{yy}(\omega_1)L_{zz}(\omega_2,z)\sin(\theta_2)\chi^{\text{D}}_{yyz}(z) + \cdots \tag{5.14}$$

$$\begin{aligned}
\chi^{(2)}_{\text{eff},ppp} = \int_{int} & dz L_{zz}(\omega_s,z)L_{zz}(\omega_1,z)L_{zz}(\omega_2,z)\sin(\theta_s)\sin(\theta_1)\sin(\theta_2)\chi^{\text{D}}_{zzz}(z) \\
& -L_{xx}(\omega_s)L_{xx}(\omega_1)L_{zz}(\omega_2,z)\cos(\theta_s)\cos(\theta_1)\sin(\theta_2)\chi^{\text{D}}_{xxz}(z) \\
& +L_{zz}(\omega_s,z)L_{xx}(\omega_1)L_{xx}(\omega_2)\sin(\theta_s)\cos(\theta_1)\cos(\theta_2)\chi^{\text{D}}_{zxx}(z) \\
& -L_{xx}(\omega_s)L_{zz}(\omega_1,z)L_{xx}(\omega_2)\cos(\theta_s)\sin(\theta_1)\cos(\theta_2)\chi^{\text{D}}_{xzx}(z) + \cdots .
\end{aligned} \tag{5.15}$$

For symmetry reasons, $\chi^{\text{D}}_{xzx}(z) = \chi^{\text{D}}_{zxx}(z)$ and in most experimental setup, θ_1, θ_2 and θ_s are usually close to 45° [7, 8]. Therefore the last two terms on the rhs of Eq. 5.15 nearly cancel each other. The local dielectric constant $\epsilon(\omega,z)$ enters the definition of the z component of the Fresnel factor in $L_{zz}(\omega,z)$ as shown in Eq. 5.4. Therefore, Eq. 5.14 and 5.15 may be



(a) (b)

Figure 5.2: Computed z component and x and y components of the high-frequency dielectric constant tensor ϵ_∞ as a function of the distance z in the direction perpendicular to the surface (a); local frequency-dependent dielectric constant, $\epsilon_z(\omega)$, averaged over the first (1st), second (2nd) and third (3rd) BLs (b) in one of the ice slab models studies in this work. The latter is compared with experimental measurement of $\epsilon(\omega)$ for bulk ice I$_h$ [6].

approximated as:

$$\begin{aligned}\chi^{(2)}_{\text{eff},ssp} &\sim \frac{C_1}{\epsilon_z(\omega_2)}\chi^{\text{D}}_{yyz} \\ \chi^{(2)}_{\text{eff},ssp} &\sim \frac{C_2}{\epsilon_z(\omega_2)\epsilon^2_\infty}\chi^{\text{D}}_{zzz} - \frac{C_3}{\epsilon_z(\omega_2)}\chi^{\text{D}}_{xxz},\end{aligned} \tag{5.16}$$

where C_1, C_2, C_3 are constants. Note that, changing the incident angles θ_1, θ_2 and θ_s may lead to a change in the relative intensities of different features in *ppp* spectra, but not in *ssp* spectra (see Eqs. 5.14 and 5.15). No approximate expressions have been used in this work.

5.3 Computational Details

We simulated ice I$_h$ basal surfaces using slab models and periodic boundary conditions. Since bulk ice I$_h$ is proton disordered, we created 8 slabs, four with proton ordered surfaces and four with proton disordered surfaces, to accumulate statistics. Each slab contained 6 bilayers (BLs) and a total of 144 water molecules. The proton disordered slabs were generated using the algorithm described in Refs. [190, 192]. All of the 6 BLs were proton disordered in the four slabs with proton disordered surfaces. The four slabs with proton ordered surfaces (in

the Fletcher's striped phase, which was predicted to be one of the most stable phases of ice I_h basal surfaces [190]) were built by manually ordering water molecules in the surface BLs, starting from each of the four slabs with proton disordered surfaces. Hence the slabs with proton ordered surfaces turned out to have four proton disordered BLs within the slab. In total we sampled 8 proton disordered and 8 proton ordered surfaces (2 surfaces per slab). We used a rectangular supercell of size 13.55 Å×15.65 Å×47.63 Å. We carefully checked that our results for the SFG spectra were converged as a function of the cell and slab size to ensure the convergence of the computed spectra. We computed the interfacial contributions $\boldsymbol{\chi}_{\text{eff}}^{\text{ID}}$ and $\boldsymbol{\chi}_{\text{eff}}^{\text{IQI}}$ using the slab models just described. We computed the bulk contributions $\boldsymbol{\chi}_{\text{eff}}^{\text{BQ}}$ and $\boldsymbol{\chi}_{\text{eff}}^{\text{IQB}}$ using four bulk proton disordered ice I_h models, each with 96 water molecules in a rectangular supercell of size 13.55 Å×15.65 Å×14.76 Å.

To compute the amplitudes $\mathbf{A}_q^{\text{X}}(z)$ for linear and nonlinear susceptibilities (Eq. 5.12), we obtained the dynamical matrix by computing derivatives of forces by FD with a displacement of atomic position of 0.01 a.u.. Normal modes were obtained by diagonalizing the dynamical matrix. We compared our results with those computed from *ab initio* molecular dynamics simulations and found small differences in peak positions and intensities (see Section 5.4.5). This indicates that anharmonic effects are small in the vibrational spectra of ice I_h basal surfaces and justifies our use of FD techniques to compute SFG spectra. The details of the computation of multipole moments and polarizabilities are reported in Section 5.2.2. We carried out all these calculations with the Qbox Code [47] using the PBE functional [59, 60], HSCV pseudopotentials [71, 72] and a planewave basis set with a kinetic energy cutoff of 85 Ry. The optical geometries, i.e. incident angles, for computing the Fresnel coefficients were adapted from Ref. [7].

Note that the calculations of $\Pi_{\nu\mu}$ and $\partial Q_{\nu\mu}/\partial E_\xi$ with slab models includes the local dielectric constant since within DFPT the applied perturbation in the z direction is that of the electric displacement field $\mathbf{D}$ rather than the total electric field $\mathbf{E}$ [193]. Therefore the ratio between $\Pi_{\nu z}(z)$ and the polarizability computed from DFPT, $\Pi_{\nu z}^{\text{DFPT}}(z)$, is given by

the inverse of the local high-frequency dielectric constant $\epsilon_{\infty,z}(z)$:

$$\Pi_{\nu z}(z) = \Pi_{\nu z}^{\mathrm{DFPT}}(z)/\epsilon_{\infty,z}(z). \tag{5.17}$$

For bulk models $\Pi_{\nu\mu}(z) = \Pi_{\nu\mu}^{\mathrm{DFPT}}(z)$ due to the absence of any surface polarization.

5.4 Tests and Validation of the Method

5.4.1 Local Dielectric Constant Profile

In Fig. 5.2 we show the computed $\epsilon_z(\omega)$ profile of a representative slab sample, showing good agreement with experiment. It was computed taking into account corrections for the spurious electrostatic interactions between slab images [194]. The expression we used to compute $\epsilon_{\infty,z}$ is:

$$\epsilon_{\infty,z}(z) = \frac{D_z}{E_z(z)} = \frac{D_z}{D_z - 4\pi P_z(z)} = \frac{1}{1 - 4\pi\Pi_{zz}^{\mathrm{DFPT}}(z)/V}. \tag{5.18}$$

Our computed $\epsilon_\infty \sim 1.8$ was also in good agreement with the experimental value (~ 1.7). We note however that the local dielectric constant $\epsilon_{\infty,z}(z)$ of the dangling OH bonds of surface water molecules was estimated to be about 1.31 [7] from experiments, a value smaller than the average presented in Fig. 5.2a. This apparent discrepancy is likely due to the fact that ϵ reported Ref. [7] is defined in terms of the local field correction factors, different from our definitions. Nevertheless we note that we observe a modest decrease of $\epsilon_{\infty,z}(z)$ (of about 0.1) in approaching the surface from the bulk region.

We found that the x, y and most of the z components of the local high-frequency dielectric constant $\boldsymbol{\epsilon}_\infty(z)$ show very little variations throughout the slab, although the z component decreases significantly at the top half of the surface BL (see Fig. 5.2a). The behavior of the local frequency-dependent dielectric constant $\epsilon_z(\omega)$ of the surface BL is significantly different from that of subsurface ones as well (see Fig. 5.2b). The variation of $\epsilon_z(\omega, z)$ as a function

of the z coordinate may affect weights of surface and subsurface contributions to $\boldsymbol{\chi}_{\text{eff}}^{\text{ID}}$ since $\epsilon_z(\omega, z)$ enters the denominator in the definition of $\mathbf{L}_{zz}(\omega, z)$ in Eq. 5.4.

Because of the small variation in $\epsilon_z(\omega, z)$ in the subsurface BLs, the gradient of the Fresnel coefficient $dL_{zz}(\omega, z)/dz$ in this region can be neglected (see Eq. 5.3). Therefore in the calculation of $\boldsymbol{\chi}_{\text{eff}}^{\text{IQI}}$, which is proportional to $dL_{zz}(\omega, z)/dz$ (see Eq. 5.6), we only need to include the surface BL. Given $d\epsilon_{\infty,z}(z)/dz$ estimated from Fig. 5.2a, we estimate that $dL_{zz}(\omega, z)/dz \simeq 0.05 L_{zz}(\omega, z)$ at the surface BL. Although this is a very rough estimate, it is sufficient to predict the order of magnitude of $\boldsymbol{\chi}_{\text{eff}}^{\text{IQI}}$. Since $\boldsymbol{\chi}_{\text{eff}}^{\text{IQI}}$ is a minor contribution to the SFG susceptibility, we expect that the error associated with our estimate will not significantly affect the total computed $\chi_{\text{eff}}^{(2)}$. More accurate predictions of $d\epsilon_{\infty,z}(z)/dz$ may be obtained by computing the change in the electrostatic potential under an applied electric field. We also note that the IQI contribution is partly included in the ID one [98]. For slab samples, Π as obtained within DFPT already includes the response to both the electric field and the electric field gradient, since the total field perturbation applied in DFPT in not constant (see Eq. 5.17).

5.4.2 Electrostatic Correction for Slab Models

The ID contribution to the spectra reported in Figs. 5.7 and 5.6 were computed without the corrections for the spurious electrostatic interactions between periodic images. Although without correction the dielectric constant turns out to be $\epsilon_{\infty,z} \approx 2.2$ is much larger than the corrected value of about 1.8, the spectra computed with and without corrections are very similar (see Fig. 5.3), due to the fact that in our models the vacuum region between slabs is sufficiently large so as to properly converge the electronic structure of the system. Inaccuracies in $\epsilon_{\infty,z}$ have little effect on the computed *ssp* spectra, since the dependences of $L_{zz}(\omega_2, z)$ (see Eq. 5.4) and $\boldsymbol{\chi}^{\text{D}}$ (see Eqs. 5.10, 5.12 and 5.17) on $\epsilon_{\infty,z}$ approximately cancel in the expression of *ssp* spectra given in Eq. 5.14. Instead in the expression of *ppp* spectra, Eq. 5.15, there is an explicit dependence on $\epsilon_{\infty,z}$ through the term $\chi_{zzz}^{\text{D}}(z)$ (see

also Eq. 5.16). However the signal contributing to $\chi^{\mathrm{D}}_{zzz}(z)$ arises mainly from the free OH stretching vibrations in the frequency region at about 3700 cm^{-1}. Therefore increasing $\epsilon_{\infty,z}$ only reduces the intensity of the free OH stretching peak, but does not significantly change the shape of the main peaks of the *ppp* spectra in the range of 3100 – 3500 cm^{-1}. We note that setting $\epsilon_z(\omega, z) = 1$ in our calculations (violet curves in Fig. 5.7) enhances the weight of $\chi^{\mathrm{D}}_{zzz}(z)$ in the expression of $\chi^{\mathrm{D}}_{\mathrm{eff},ppp}(z)$ in Eq. 5.15 and therefore significantly changes the shape of $\chi^{\mathrm{D}}_{\mathrm{eff},ppp}(z)$.

5.4.3 Effect of Incident Angles Used in Different Experiments

The incident angles of the input and output light beams in the two experimental measurements we compared our calculations with (see Fig. 4 in manuscript) [7, 8] are slightly different. Wei *et al.* [7] used θ_1 and θ_2 values of 45° and 57°, respectively, while Bisson and Shultz [8] used 50° and 60°, respectively. These parameters enter the final expression for the spectra in Eq. 5.14 and 5.15. The angles θ_s can be inferred from θ_1 and θ_2 using the equality $k_{1,x} + k_{2,x} = k_{s,x}$ and the phase matching condition $|\mathbf{k}_1| + |\mathbf{k}_2| = |\mathbf{k}_s|$. All of our spectra were computed using the optical geometry of Ref. [7]. As shown in Fig. 5.3 the difference between spectra computed with the incident angles of Ref. [7] and Ref. [8] are negligible.

5.4.4 Convergence of SFG spectra with respect to the slab thickness

In our calculation of $\chi^{\mathrm{ID}}_{\mathrm{eff}}$, we used ice slabs consisting of six BLs, three BLs for each surface. We show the ID contribution to the SFG spectra computed by including one, two or three surface BLs in Fig. 5.7. We found that $\chi^{\mathrm{ID}}_{\mathrm{eff}}$ computed from two or three surface BLs are very similar. This indicates that at least two surface BLs (about 7 Å) are needed in surface calculations in order to converge the computed spectra.

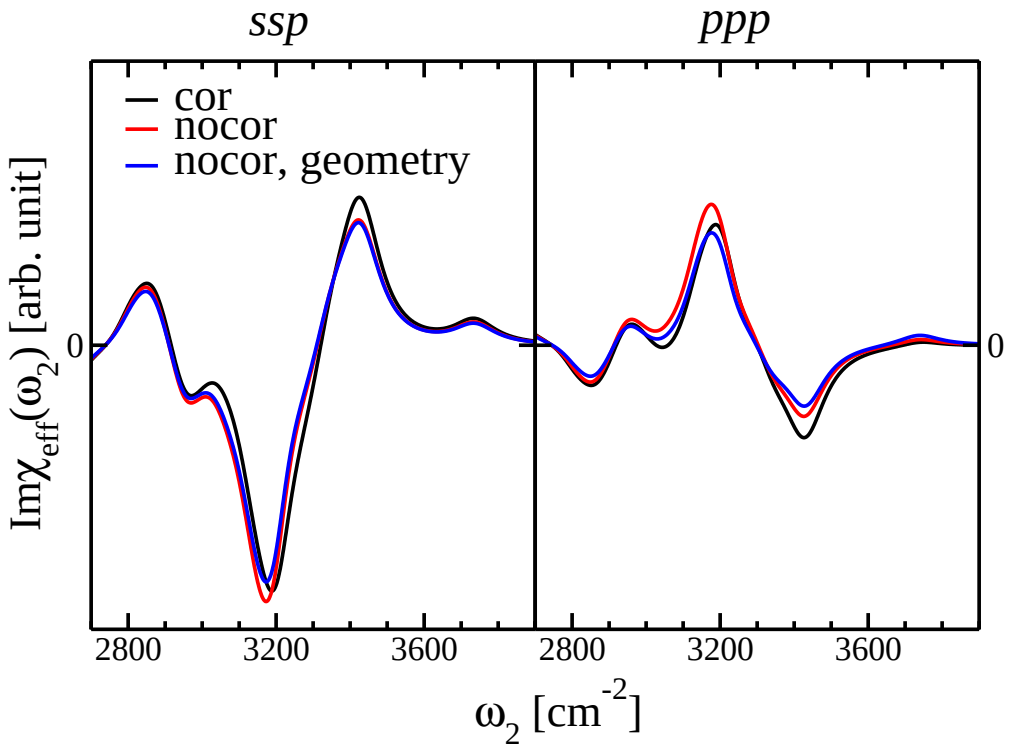


Figure 5.3: Imaginary part of the $\chi_{\text{eff}}^{\text{ID}}$ contribution to SFG spectra [the *ssp* (left) and *ppp* (right) polarization combinations] computed for one proton-disordered surface with (black) and without (red and blue) correction for the long-range interaction between slabs. We used the optical geometry (incident angles) of Ref. [7] for black and red curves and the optical geometry of Ref. [8] for the blue curve.

5.4.5 SFG spectra from FPMD simulations

In order to assess the importance of anharmonic effects on the computed SFG spectra of ice I_h basal surfaces, we carried out Born-Oppenheimer AIMD simulations on two slab samples (one with proton ordered and one with proton disordered surfaces). We used the same atomic geometries and cell parameters as used in FD calculations, with the PBE functional [59] and a time step of 5 a.u. (0.12 fs). After equilibrating the two samples for 20 ps in the NVT ensemble, we performed 70 ps production runs in the NVE ensemble. Then we computed $\boldsymbol{\chi}^{\mathrm{D}}$ from $\chi^{\mathrm{D}}_{\nu\mu\xi} \propto \int \langle M_{\xi}(t)\Pi_{\nu\mu}(t+\Delta t)\rangle dt \exp(i\omega_2 \Delta t)$. The χ^{ID} spectra computed from AIMD simulations and FD calculations are compared in Fig. 5.4. Note that to simplify the comparison, all spectra reported in Fig. 5.4 were computed by setting the Fresnel coefficients to the identity tensor. We observed small differences in the peak positions and intensities of AIMD and FD results, indicating that the contribution of anharmonic effect to the SFG spectra is small.

5.4.6 The Origin Dependency of Quadrupole Contributions

In this section, we discuss the accuracy of our method in solving the origin dependence problem in the calculation of the quadrupole contributions. As discussed in Section 5.2.2, we used the total charge density to compute the quadrupole moment and nonlinear susceptibilities $\boldsymbol{\chi}^{\mathrm{Q_1}}$, $\boldsymbol{\chi}^{\mathrm{Q_2}}$ and $\boldsymbol{\chi}^{\mathrm{Q_s}}$, from which we computed the IQB and BQ contributions using Eqs. 5.7 and 5.8. In Fig. 5.5, we show how the computed $|\chi^{\mathrm{IQB}}_{\mathrm{eff}}|^2$ varies as the origin is changed when computing quadrupole moments. We observed only small changes that do not affect the total computed spectra. The small origin dependency found here comes from a small deviation from centrosymmetry in our model bulk ice samples, which is expected due to our representation of proton disordered ice with a relatively small finite sample (96 molecules). This error can be reduced by using a larger sample so as to better recover centrosymmetry.

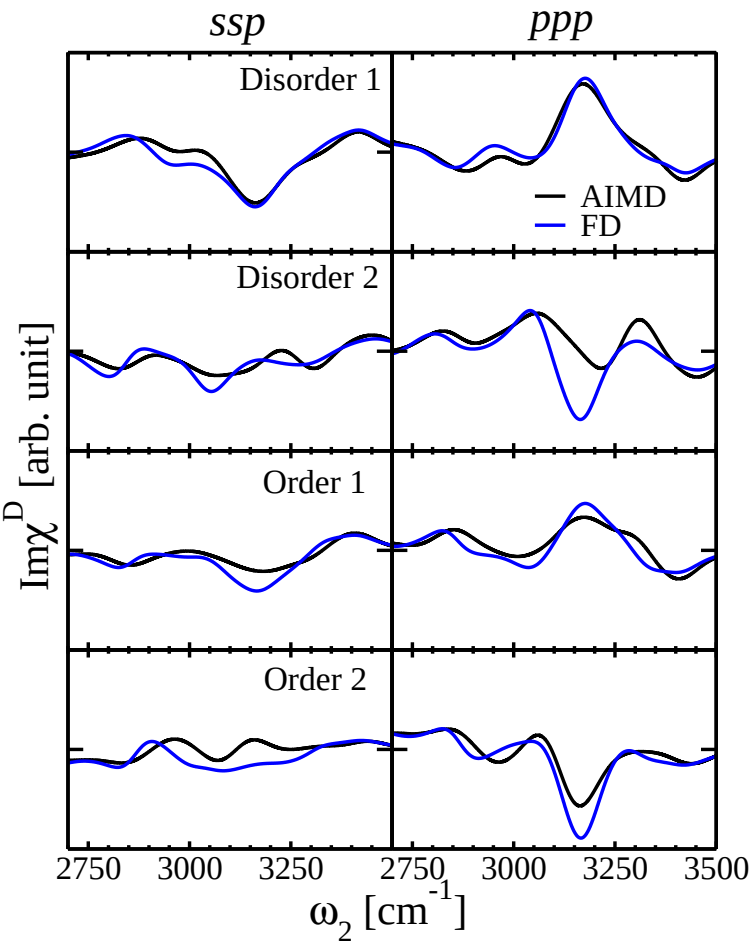


Figure 5.4: Imaginary part of the computed $\boldsymbol{\chi}_{\text{eff}}^{\text{ID}}$ contribution to the SFG spectra of *ssp* and *ppp* polarization combinations, as obtained from AIMD (black) and finite difference (FD, blue) calculations. The upper two panels show spectra computed for two proton disordered surfaces; the lower two panels for two proton ordered surfaces. All AIMD and FD spectra were obtained by including three surface BLs in the calculation and setting the Fresnel coefficients to the identity tensor.

5.5 Results and Discussion

The various contributions to the imaginary part of $\boldsymbol{\chi}_{\text{eff}}^{(2)}$ in the OH stretching region are shown in Fig. 5.6, for two polarization combinations *ssp* and *ppp* (see Section 5.4.1). The gradient of the Fresnel coefficients used in the calculation of $\chi_{\text{eff}}^{\text{IQI}}$ is estimated from the dielectric constant profile in Fig. 5.2 (see Section 5.4.1). Because $\chi_{\text{eff}}^{\text{IQI}}$ is much smaller than $\chi_{\text{eff}}^{\text{ID}}$ (see below), the errors arising from its approximate evaluation are expected to be negligible. In

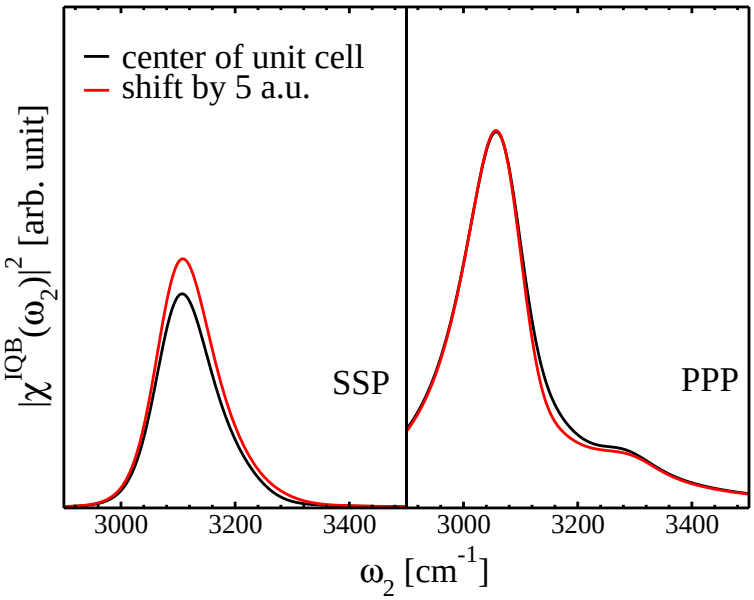


Figure 5.5: The computed IQB contribution $|\chi_{\rm eff}^{\rm IQB}(\omega_2)|^2$ of one bulk ice sample. The black and red curves were computed using (0,0,0) and (5,5,5), in a.u., as the origin (see text).

the upper and middle panels of Fig. 5.6, we show the real and imaginary part of the $\chi_{\rm eff}^{\rm ID}$ and $\chi_{\rm eff}^{\rm ID} + \chi_{\rm eff}^{\rm IQB}$ contributions to the SFG spectra. The results computed from the proton ordered and proton disordered ice I_h basal surface samples show similar features. It may be difficult to distinguish between the two surfaces based on these spectra. Interestingly, we found that in the *ppp* spectra, $\chi_{\rm eff}^{\rm IQB}$ is the only significant quadrupole contribution, mostly influencing the peak at about $\sim$ 3150 cm^{-1}. This finding is in agreement with previous speculations [188] that such peak includes significant quadrupole contributions. This high frequency peak has been previously assigned [188, 195] to a "bilayer-stitching mode", i.e. a stretching mode of the OH groups hydrogen bonded to water molecules in neighboring BLs, , which are vibrating along the z direction. These vibrations should indeed result in a large zzz component of $\chi_{\rm eff}^{\rm IQB}$, which is an important contribution to *ppp* spectra. Instead the contribution of $\chi_{\rm eff}^{\rm IQB}$ to *ssp* spectra is less intense since it has only yyz components.

In Fig. 5.7, we present the computed SFG intensities $|\chi_{\rm eff}^{(2)}|^2$ averaged over several surface models and compared with different experiments. The difference in the measurements [7, 8] likely originates from different temperatures and surface cleanness (slightly different incident

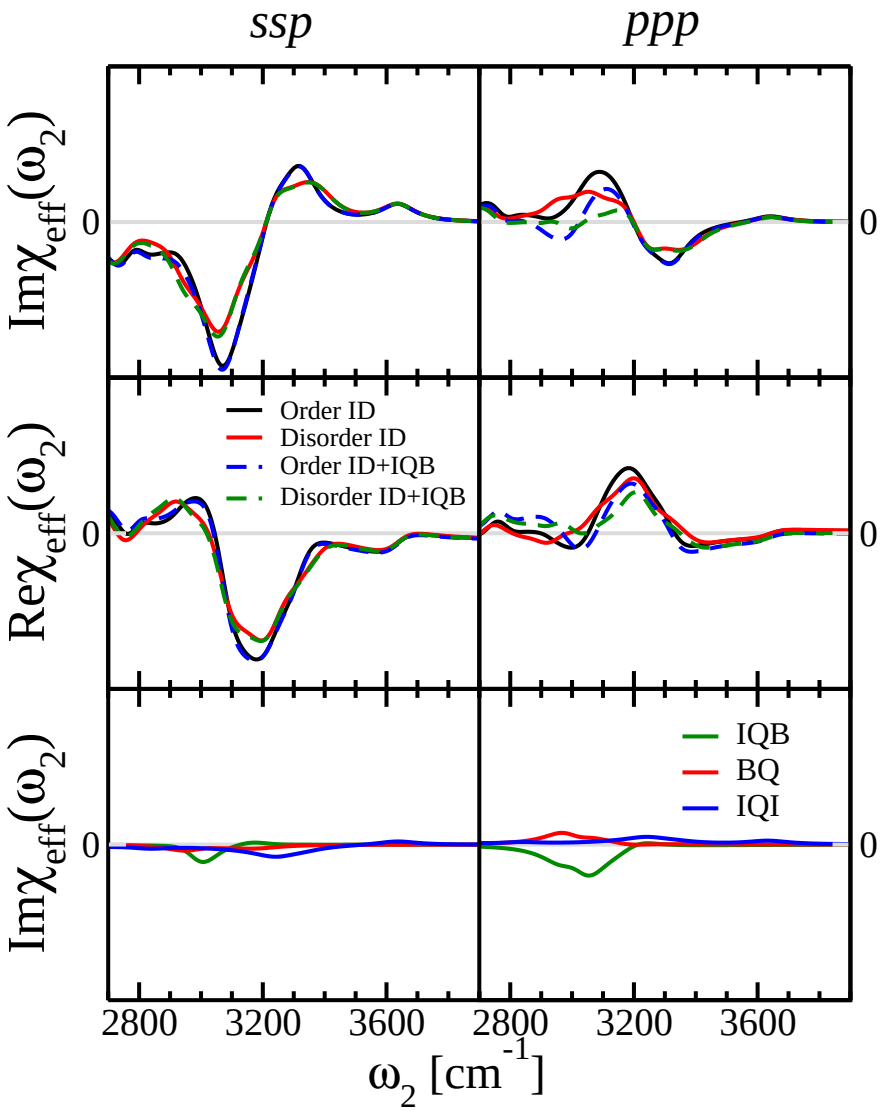


Figure 5.6: Real (top panels) and imaginary (middle panels) part of $\chi_{\text{eff}}^{\text{ID}}$ spectra computed for the proton ordered (black line) and proton disordered (red line) ice I_h basal surfaces. The corresponding $\chi_{\text{eff}}^{\text{ID}} + \chi_{\text{eff}}^{\text{IQB}}$ spectra for the two types surfaces (blue and green dashed lines) are shown as well. Imaginary part of the $\chi_{\text{eff}}^{\text{IQB}}$, $\chi_{\text{eff}}^{\text{BQ}}$ and $\chi_{\text{eff}}^{\text{IQI}}$ are shown in the bottom panels. The spectra for *ssp* and *ppp* polarization combinations are shown in left and right panels, respectively. The computed spectra were red-shifted by 100 cm^{-1} to align them with experiments at 3100 cm^{-1}. The discrepancy in peak positions is ascribed to a combined effect of the neglect of quantum effects [9] and the use of the PBE functional [10]. Anharmonic effects were not found to be significant for this system (Fig. 5.4).

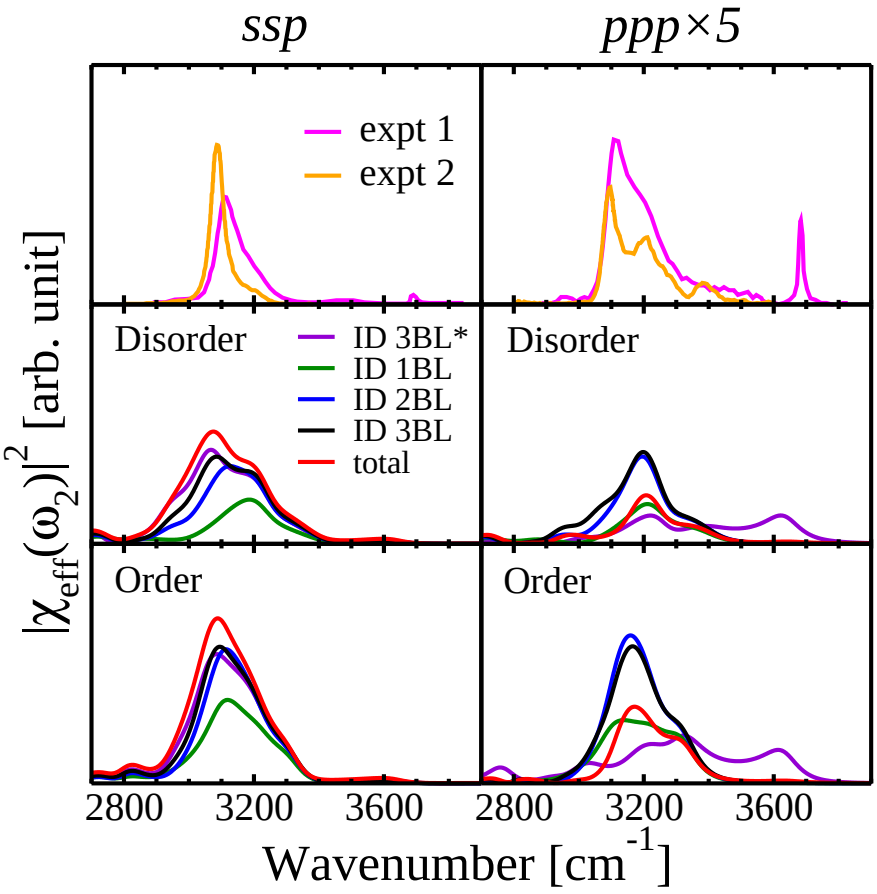


Figure 5.7: SFG intensities for *ssp* (left panels) and *ppp* (right panels) polarization combinations for ice I_h basal surfaces. We report measured (at 173 K, Ref. [7], and at 100 K, Ref. [8]; top panels) and computed results (as averages for proton disordered surface models in the middle panels, and proton ordered surface models in the bottom panels). The spectra were computed by including the ID contribution originating from the top one (green), two (blue) or three (black) surface BLs (see text). Violet curves show only the ID contribution from three surface BLs obtained by setting $\epsilon(\omega, z) = 1$ in the calculation. The total spectra $|\chi^{(2)}_{\text{eff}}|^2$, including ID contribution from three surface BLs as well as IQI, IQB and BQ contributions, are shown by the red curve. The computed spectra were red-shifted by 100 cm^{-1} to align them with experiments at 3100 cm^{-1}.

angles used in the two experiments are not expected to significantly change the spectra as explained in Section 5.2.4). The dipole contribution to computed SFG intensities was considered converged when including three BLs, with the first two BLs ($\simeq$ 7 Å) yielding most of the total contribution (see Section 5.4.4). Similar to our results for $\mathrm{Im}\chi^{(2)}_{\mathrm{eff}}$ spectra, we found that quadrupole contributions are only significant in the *ppp* spectra. We also found that it is important to properly include the frequency dependent dielectric constant (shown in the right panel of Fig. 5.2) in computing the *ppp* spectra, as it is seen by comparing calculations with $\epsilon(\omega, z)$ set to 1 (violet curves in Fig. 5.7) with those including the proper $\epsilon(\omega, z)$. As shown in Eq. 5.16, $\chi^{(2)}_{\mathrm{eff},ssp}$ and $\chi^{(2)}_{\mathrm{eff},ppp}$ have different dependencies on $\epsilon(\omega)$ and the former is largely unaffected when setting $\epsilon(\omega) = 1$.

Our results show that the proton ordered and disordered surfaces do exhibit different contributions to the SFG spectra arising from the first BL: in the *ssp* spectra (Fig. 5.7), the proton disordered surfaces show a feature about 100 cm^{-1} higher in frequency than proton ordered ones. However, when including three BLs and quadrupole contributions, the total spectra of proton ordered and disordered surfaces cannot be unambiguously distinguished. Our results are consistent with results reported by Ishiyama *et al.* [195], who found significant contributions originating from subsurface BLs. As shown in the top panels of Fig. 5.7, both experimental *ssp* spectra show a major peak at $\sim 3100\ \mathrm{cm}^{-1}$ and a shoulder at $\sim 3200\ \mathrm{cm}^{-1}$; however the spectrum measured at a lower temperature (100 K) and better surface conditions [8] show a weaker shoulder than the one taken at higher temperature (173 K) [7]. Overall our results appear to point to a proton ordered surface at low temperature, in agreement with the prediction of Ref. [190].

5.6 Conclusion

In summary, we developed a theoretical and computational method to compute SFG spectra from first principles. We consistently took into account the varying electric field across the interface and we included dipole and quadrupole contributions from both surface and bulk

regions. We also provided a simple computational framework to avoid origin dependent terms in the calculation of quadrupole contributions. We used the method to compute the SFG spectra of the ice I_h basal surface and obtained good agreement with experiments. Our results highlight the importance of including electric field gradients and quadrupole contributions to properly describe the SFG spectra of the *ppp* polarization combinations. The first principles methodology developed here is general and applicable to ordered and disordered semiconductors and insulators.

CHAPTER 6

SOLVATION PROPERTIES OF MICROHYDRATED SULFATE ANION CLUSTERS: INSIGHTS FROM *AB INITIO* CALCULATIONS

In addition to water and ice, we used electronic structure calculations and FPMD to study the properties of ions in water. In particular, in this chapter we present a study of the solvation (surface vs internal) of sulfate anions in small water clusters, which may affect the nucleation process of sulfate-containing aerosols. We show the results for our FPMD simulation of sulfuric acid aqueous solutions in the next chapter, where we discuss the energy level alignment between ion and water states and the implication of our results for the mechanism of oxygen evolution reactions in water splitting experiments.

6.1 Introduction

Sulfate-water clusters play an important role in industrial and environmental processes, such as nucleation sites in clouds [196–199], and the formation of aerosols and thus of acid rain. In addition, sulfate-water clusters are prototype systems for the investigation of the complex structural and dynamical properties of anion solvation [200, 201]. For these reasons, small hydrated clusters of the sulfate dianion, $SO_4^{2-} \cdot (H_2O)_n$, have been extensively investigated both experimentally [11, 202–208] and theoretically [209–214]. Isolated sulfate dianion does not exist in the gas phase, due to the strong repulsion between its two negative charges and at least three water molecules are necessary to stabilize the excess charge [209, 215]. Blackbody Infrared Radiative Dissociation (BIRD) experiments [207] have shown that $SO_4^{2-} \cdot (H_2O)_6$ and $SO_4^{2-} \cdot (H_2O)_{12}$ clusters are more stable, compared to neighboring ones, and their stability has been interpreted as indicating the presence either of a complete solvation shell or of

particularly strong hydrogen bonds. However a recent theoretical study [214] found several new structures for n=6-7, that are iso-energetic with the internally solvated anion. Infrared [11] and photodetachment [202, 203] experiments on $SO_4^{2-} \cdot (H_2O)_{12}$ have been interpreted as signatures of a complete, symmetric solvation shell around the anion, although neither experiment could lead to unequivocal structural assignments.

Sulfate ions are believed to be absent from air-aqueous solution interfaces [216]. Classical molecular dynamics (MD) simulations of sulfate-water cluster consisting of 60 water molecules suggests a preference for internal solvation of the sulfate dianion [211].In addition, a joint experimental and theoretical study based on sum frequency generation (SFG) spectroscopy and empirical MD simulations of ammonium and sodium sulfate aqueous solutions [217], and of sulfuric acid [218] in water indicated that at room temperature, the sulfate ion is absent within at least 6Å from the interface with air, in all solutions. However, for small size clusters, such a clear preference for internal solvation is yet unclear.

In this chapter we investigate the stability, vibrational and electronic properties of $SO_4^{2-} \cdot (H_2O)_n$, with $n = 12$ and 13, using *ab initio* molecular dynamics [32, 46] and electronic structure calculations [219] based on Density Functional Theory (DFT), with semilocal [59, 60] and hybrid functionals [68]. We considered 12 water molecules, as $SO_4^{2-} \cdot (H_2O)_{12}$ is the most stable among the small sulfate clusters and it was expected to be the first in the series with a complete solvation shell [11, 207]. Both internally and surface solvated structures were investigated. We then added an additional water molecule to investigate in detail the competition between water-water and ion-water interactions. For both 12- and 13-water anion clusters, we studied the stability of different geometrical configurations by computing total energies and vibrational free energies from harmonic frequencies. We then computed infrared (IR) spectra using both MD simulations and finite difference calculations and compared our results with available infrared multi-photon dissociation (IRMPD) experiments [11]. We found a more complex potential energy surface for the13 water cluster than previously recognized and obtained similar vibrational signatures for clusters with different

structures, making it difficult to use IR spectra to unequivocally assign cluster geometries. Finally we computed the electronic properties of the internally and surface solvated 12-water sulfate dianion clusters and found very similar electronic density of states and vertical ionization potentials. Our findings are consistent with IR and photodetachment experiments and show that a surface solvated anion configuration is energetically favored, compared to an internally solvated one. We suggest that mixtures of surface and internally solvated configurations may be present in experimental samples, with metastable configurations kinetically trapped in the fast cooling process at low temperature in the case of IRMPD experiments. While structural and stability properties obtained with PBE and PBE0 functionals are similar, the use of a hybrid functional is necessary to obtain quantitative agreement with measured vibrational frequencies and to compute ionization potentials.

The rest of the paper is organized as follow: in the next section we describe our computational methods and then we present our results, discussing IR spectra of the 12 and 13 water hydrated sulfate anion first, followed by an analysis of their electronic properties. Finally we present our conclusions in the last section.

6.2 Theoretical Methods

We performed *ab initio* Born-Oppenheimer (BO) MD simulations of $SO_4^{2-}(H_2O)_n$ and $SO_4^{2-}(D_2O)_n$ clusters with n=12 and 13, with the generalized-gradient exchange-correlation functional PBE [59, 60] and the hybrid functional PBE0 [68], using the Qbox code [46]. We simulated isolated clusters in a cubic super cell of length 30 atomic units (a.u.). The negative charge of the anion was neutralized by a positive uniform background [44]. We used plane waves basis sets with a cutoff of 80 Ry and norm-conserving pseudo-potentials [71, 72]. Our simulations were carried out at 100 K in the NVE ensemble, after equilibrating the system in the constant volume and temperature ensemble (NVT), where the temperature was controlled with a stochastic velocity rescaling thermostat [151]. We used a time step of 5 and 10 a.u. (10 a.u. = 0.24 fs) for light and heavy water, respectively.

Starting from selected configurations from MD trajectories, we optimized the cluster geometries. In the calculation of total energies we included zero-point vibrational energies, obtained using vibrational frequencies determined from harmonic finite difference calculations. We also computed the vertical ionization potential (IP) of selected configurations, defined as $IP = E_{SO_4^{2-}\cdot(H_2O)_n} - E_{SO_4^{-}\cdot(H_2O)_n}$, with both energies computed at the optimized dianion cluster geometry, and using the Makov-Payne correction [220] implemented in the Quantum Espresso code [219].

Vibrational Infrared (IR) spectra were obtained both from finite difference calculations and from dipole dipole correlation functions computed over MD trajectories. The IR intensity (I_q) of a normal mode (S_q) is proportional to the square of the derivative of the cluster's dipole moment with respect to the normal mode coordinate [221]:

$$I_q \propto |\frac{\partial \mathbf{M}}{\partial S_q}|^2. \tag{6.1}$$

The normal modes can be obtained by finite difference calculations of the atomic positions.

IR spectra can also be derived from the Fourier transform of the dipole auto-correlation function computed over an MD trajectory:

$$I(\omega) \propto Q_{QC} \cdot \int_{-\infty}^{\infty} dt e^{-i\omega t} \langle \mathbf{M}(t) \cdot \mathbf{M}(0) \rangle, \tag{6.2}$$

where the so called quantum correction term Q_{QC} corrects for the use of the classical expression of the dipole correlation function instead of the quantum mechanical one. Here we adopted the harmonic approximation with $Q_{QC} = \beta\hbar\omega/[1 - \exp(-\beta\hbar\omega)]$ [109, 222, 223], which has been extensively used in calculations of IR spectra of bulk water [23, 35]. Our results are presented in the next section.

6.3 Results and Discussion

6.3.1 Structural Properties of 12-Water Sulfate Clusters

We focused on the solvation of the dianion; we did not include solvation studies of HSO_4^- since both IRMPD [11] and photodetachment [202, 203] experiments did not detect any signal of OH^-, thus ruling out the occurrence of reactions of SO_4^{2-} with water to form HSO_4^- and OH^-. We considered the two different geometries of $SO_4^{2-} \cdot (H_2O)_{12}$ proposed in Ref. [11, 207, 210, 212] and shown in Fig. 6.1. In the 12A structure [11, 207, 210] the sulfate ion is at the center of the cluster, i.e. it is internally solvated, whereas in the 12B configuration [212] all water molecules are on one side of the anion which is thus at the surface of the cluster. The 12A geometry belongs to the T symmetry group. Each water molecule is hydrogen bonded to one oxygen atom of the sulfate ion and to two neighboring water molecules, forming four membered rings, each composed of three molecules. All water molecules reside in a single coordination shell, maximizing the number of hydrogen bonds. On the contrary the 12B structure has only mirror symmetry, with 8 molecules directly hydrogen-bonded to the anion, forming two four-membered rings (no three-membered rings are present). In these four-membered rings, each molecule has three hydrogen bonds with oxygens of the sulfate anion, and one with another water molecule. We note that both three- and four-membered ring structures have been reported in the stable geometries of 6-water sulfate cluster based on IR exprements by Bush *et al.*[205]; they have also been recently found by Lambrecht *et al.*[214] in some of the many low-energy geometries of 6- and 7-water sulfate clusters in an exhaustive search for cluster geometries.

In our *ab initio* MD simulations, we found that both 12A and 12B geometries are stable at 100K (with light and heavy water), for over 20ps, consistent with the observed stability of 12-water sulfate clusters in BIRD experiments [207]. An analysis of hydrogen bond length and bond angles showed that the 12B structure exhibits shorter hydrogen bonds and larger bond angles than the 12 A, indicating the presence of stronger hydrogen bonds than in 12A

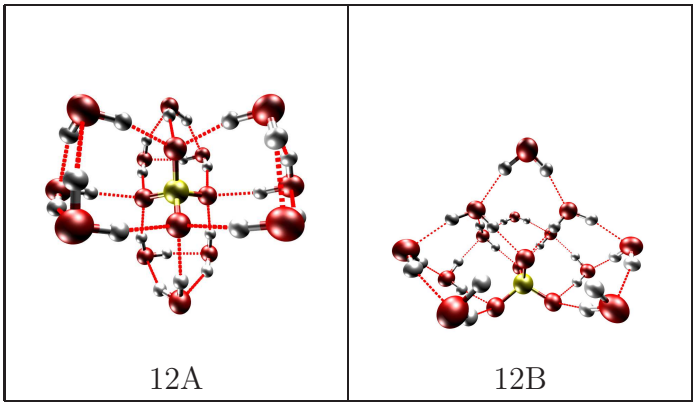


Figure 6.1: Internally (12A) and surface (12B) solvated configurations of the SO_4^{2-} $(H_2O)_{12}$ cluster. Oxygen, hydrogen and sulfur are represented in red, white and yellow, respectively. Red dotted lines indicate hydrogen bonds.

(see Table 6.2).

In Ref. [11] based on the comparison of measured and computed IR spectra, the geometry of the 12A cluster was considered as the most probable one. Photodetachment experiments [202, 203] were also interpreted as suggesting the stability of internally solvated configurations for the $SO_4^{2-}(H_2O)_{12}$ cluster, although these measurements do not permit to make unequivocal structural assignments. However DFT calculations [212] using the B3LYP exchange and correlation functional found that the 12B geometry is energetically favored over that of 12A by about 0.2eV [212]. Our total energy calculations with the PBE and PBE0 functionals, inclusive of zero point energies, confirm these findings, as shown in Table 6.1.

In order to investigate the apparent disagreement between theory and experiments on the cluster stability, we computed IR spectra for the 12A and 12B configurations and compared our results with available IRMPD experiments carried out at 17K [11] (see Fig. 6.2). We note that IRMPD data are representative of linear IR absorption spectra only under the assumption of fast internal vibrational redistribution [200]. The accuracy of this hypothesis is difficult to assess and thus the comparison with experiments presented here is not expected to provide a one-to-one correspondence between computed and measured peaks and it is considered to be qualitative. IR spectra were obtained both from finite difference calculations

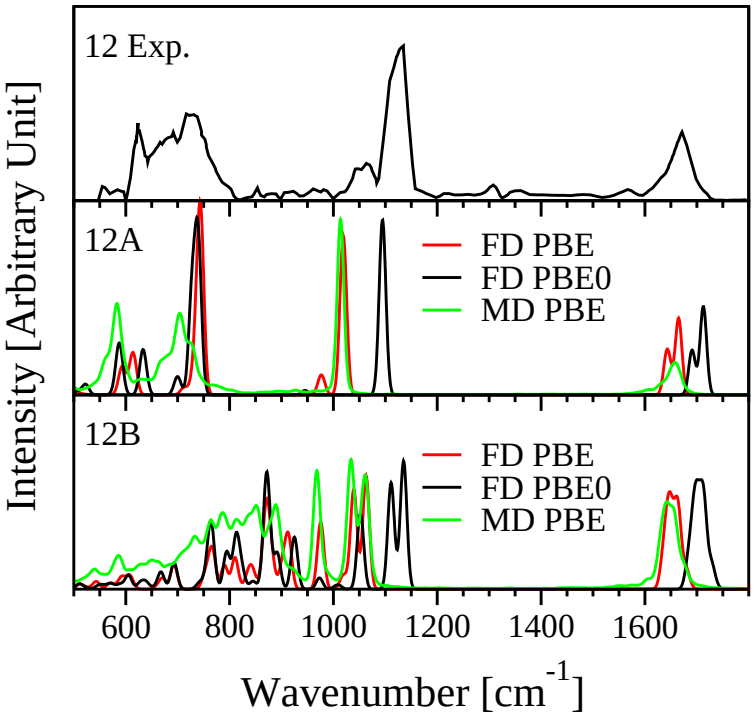


Figure 6.2: Infrared spectra of 12-water sulfate clusters $SO_4^{2-}(H_2O)_{12}$. The uppermost panel shows the experimental IRMPD spectra[11]. The middle and lower panels show computed spectra for the 12A and 12B clusters, respectively (see Fig. 6.1), obtained using finite differences (FD) and molecular dynamics (MD) simulations, with either the PBE or PBE0 functional.

and from the Fourier transform of the time correlation function of the cluster dipole moment over a ∼20 ps long MD trajectory with light water. In the region between 600-800 cm^{-1}, calculated IR spectra for the 12A cluster show a satisfactory agreement with the experimental data. The broad peak within $600 - 800$ cm^{-1} arises from the superposition of the sulfate bending mode at ~ 600 cm^{-1} and water librational modes at ~ 800 cm^{-1}. In this region of the spectrum we did not find significant differences between results obtained with the PBE and PBE0 functional. IR spectra from MD simulations using PBE show broader features, with librational modes shifted at slightly lower frequencies, compared to those computed from finite differences with the same functional; these differences stem from the inclusion of anharmonic effects in the MD simulations, resulting in a better agreement with experiment. Below 800 cm^{-1}, the agreement between experiments and the computed spectrum of the

12B structure is only fair, and inferior to that found for the 12A cluster. However in the higher frequency region, 1000-1200 cm^{-1}, we found that our results for 12B compare better to experiments than those for 12A. While the 12B structure exhibits peaks both at ∼1100 and 1050 cm^{-1}, similar to experiment, only one sharp peak and a very weak shoulder (close to the main peak) are observed for the 12A cluster. The intense peak at $\sim$ 1100 cm^{-1} corresponds to a sulfate stretching mode, which is split in the non symmetric 12B cluster; its position is underestimated in PBE calculations by about 80 cm^{-1}, while it is well reproduced using PBE0.

Table 6.1: Energy difference (eV) between cluster geometries represented in Fig. 6.1 and 6.3 for the 12 and 13 water hydrated sulfate dianion clusters. Calculations were carried out using semilocal (PBE) and hybrid (PBE0) functionals, and adding the harmonic zero point energy contribution (ZPE) computed from finite difference calculations.

Geometry	PBE	PBE+ZPE	PBE0	PBE0+ZPE
12A	0	0	0	0
12B	-0.23	-0.20	-0.19	-0.17
13A	0	0	0	0
13B	-0.28	-0.23	-0.18	-0.14
13C	-0.06	-0.04	-0.05	-0.04

These results suggest that a mixture of symmetric 12A and non symmetric 12B sulfate-water clusters are likely to exist in the experimental samples, with the energetically less favorable structure 12A being kinetically trapped under rapid cooling at very low temperature [205]. This conclusion is also supported by our electronic structure analysis, reported in the electronic structure section. Before discussing the electronic properties, we report below our results for the stability and IR spectra of $SO_4^{2-} \cdot (H_2O)_{13}$.

6.3.2 Structural Properties of 13-Water Sulfate Clusters

We first considered the structure of the 13 water hydrated sulfate cluster denoted as 13A in Fig. 6.3, proposed on the basis of IR experiments [11]; this configuration comprises a water molecule residing in a separate, second solvation shell, while the other 12 molecules remain

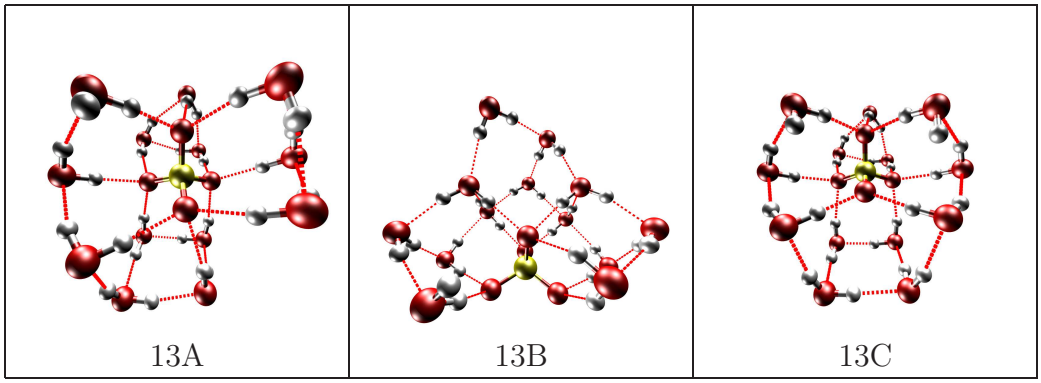


Figure 6.3: Internally (13A and 13C) and surface (13B) solvated configurations of the $SO_4^{2-}(H_2O)_{13}$ cluster. Oxygen, hydrogen and sulfur are represented in red, white and yellow, respectively. Red dotted lines indicate hydrogen bonds.

in the first shell. The calculated IR spectra for the 13A geometry reported in Ref. [11] shows a relatively good agreement with the experimental data, suggesting that 13A may be the stable structure for a 13-water sulfate cluster. The authors of Ref. [11] therefore concluded that the addition of an extra water molecule to a 12-water sulfate cluster does not affect the structure of the first solvation shell.

Table 6.2: Average bond lengths (Å) and average bond angles (°) for water-ion and water-water hydrogen bonds, computed over 20 ps *ab initio* MD trajectories. No results are shown for 13A,which was found to be unstable in our simulations at finite temperature.

		12A	12B	13B	13C
Bond length	Ion-Water	1.97	1.84	1.84	1.94
	Water-Water	1.96	1.94	1.92	1.92
Bond angle	Ion-Water	165.5	168.4	167.7	166.4
	Water-Water	157.0	163.3	164.8	162.9

We performed MD simulations of a 13-water cluster at 100 K starting from the 13A configuration of Fig. 6.3 and we found that the 13A geometry is not stable. In Fig. 6.4 we report the distance of two water molecules in the first and second coordination shell, respectively. After a short simulation time, one of the water molecules leaves the first coordination shell. A similar result was found for both light and heavy water clusters. The new structure found in our MD simulations, denoted as 13C in Fig. 6.3, has a lower energy

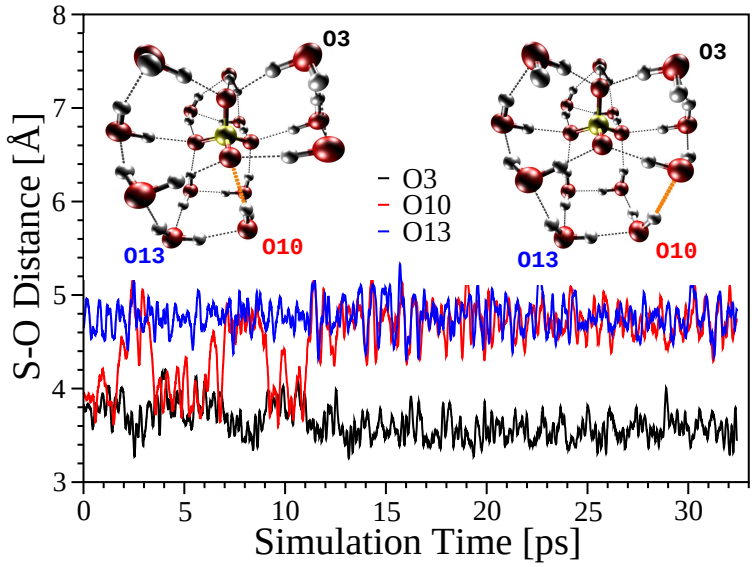


Figure 6.4: Distance between the sulfur atom and three oxygen atoms (denoted as O3, O10 and O13) belonging to solvating water molecules, during a ∼30ps *ab initio* molecular dynamics simulation at 100K. The starting (13A) and final (13C) configurations are shown as insets. The bond lengths between S and oxygen atoms 3, 10 and 13 are denoted by black, red and blue lines respectively. A jump of O10 from the first to the second solvation shell can be observed from the change in sulfur-oxygen distance: the affected hydrogen bonds are highlighted in orange in both the starting and final configurations.

than the starting one, 13A (see Table 6.1), after geometry optimization and inclusion of zero point energy. In Fig. 6.5 we report the simulated IR spectra for 13A and 13C, which turn out to be very similar. Zhou *et al.*[11] assumed the existence of 13A under experimental conditions based on the peak at 965 cm^{-1}, attributed to a vibrational mode localized on the water molecule outside the first solvation shell. Our results show that rather different geometries of the solvation shells exhibit almost identical IR spectra, in particular the same peak at 965 cm^{-1}. These findings indicate that the geometry of the solvated dianion cannot be unequivocally determined based solely on IR experiments.

To further investigate the stability of the first solvation shell, we simulated a 13-light-water cluster starting from the 12B geometry with one extra water molecule (see Fig. 6.6). We started our simulation by placing the extra water molecule close to the bare side of the anion. After a short simulation time (∼ 2 ps), the extra water molecule (circled in green

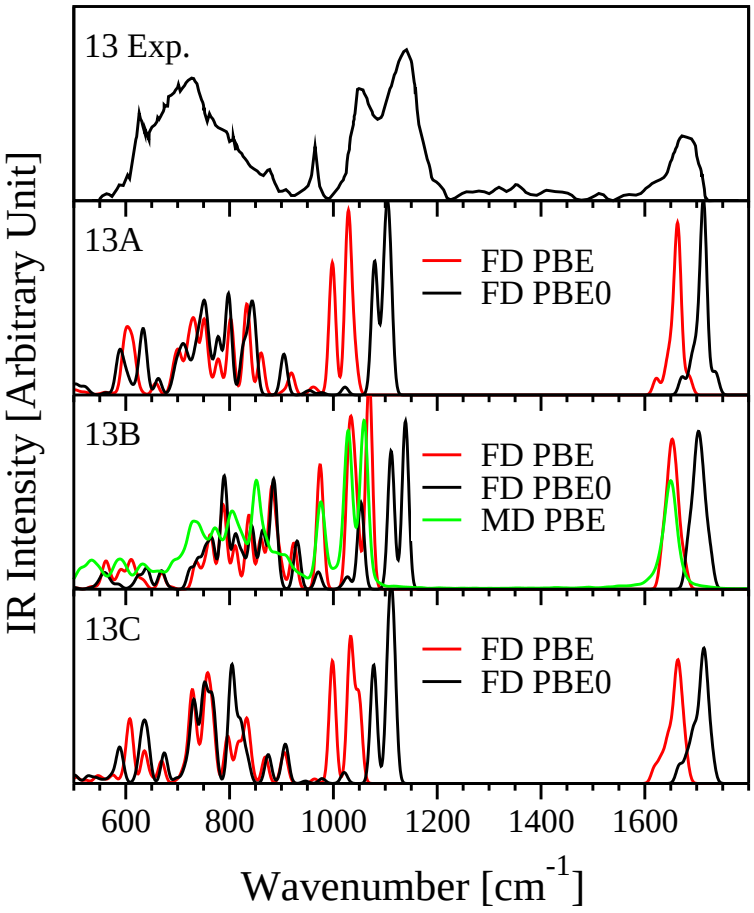


Figure 6.5: Infrared spectra of 13-water sulfate anion clusters $SO_4^{2-}(H_2O)_{13}$. The uppermost panel shows the experimental IRMPD spectra[11]. The remaining panels show computed spectra for the 13A, 13B and 13C clusters, respectively (see Fig. 6.3), obtained using finite differences (FD) and molecular dynamics (MD) simulations, with either the PBE or PBE0 functional.

in Fig. 6.6) migrates to the opposite side of the cluster, leaving the anion at the surface. The new surface-solvated configuration is labeled 13B in Fig. 6.3. The extra water molecule does not bind to the anion, but rather it prefers to form hydrogen bonds with the other water molecules (see also Table 6.2). This preference for water molecules to bind with water rather than the sulfate anion was also predicted by calculations [212, 214] of smaller sulfate clusters. The configuration 13B is energetically favored with respect to 13A and 13C (see Table 6.1) by about 0.15 eV, an energy difference of the same order as that found between 12A and 12B.

Simulated IR spectra for 13B are shown in Fig. 6.5. Similarities are evident in the IR features of 12B and 13B, especially in the splitting of the sulfate streching mode in the 1000-1100cm^{-1} region. All the three structures investigated here, 13A, 13B and 13C, show IR spectra with features that match experimental measurements. Overall our findings indicate that for $SO_4^{2-} \cdot (H_2O)_n$ clusters with n=12 and 13, configurations where the anion is at the surface are expected to be present in experimental samples, likely in a mixture of internally and surface solvated geometries. We now turn to the discussion of the cluster electronic properties and show that our results on surface solvation are also consistent with photodetachment data, although photodetachment experiments have been interpreted as showing internal solvation of the anion in $SO_4^{2-} \cdot (H_2O)_{12}$.

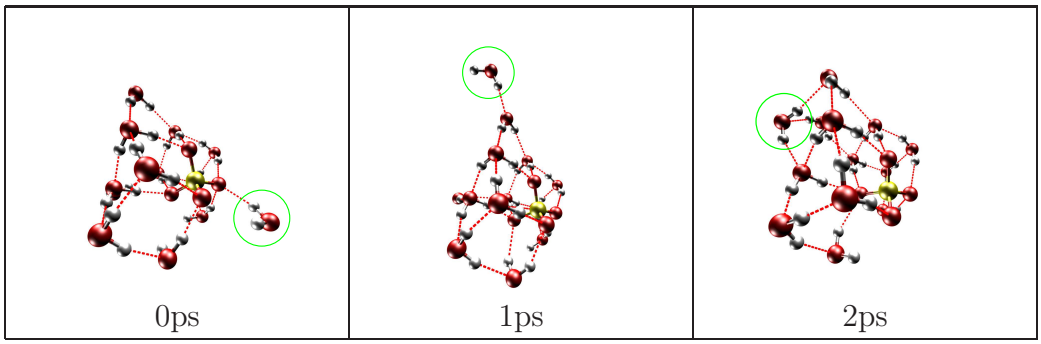


Figure 6.6: Structural changes observed during the first 2 ps of an *ab initio* MD simulation of a 13-water hydrated sulfate anion cluster (see text).

6.3.3 Electronic Properties

The description of the electronic properties of solvated anions is in general more complex than that of simple cations. For example, it has been shown [66, 67] that local and semilocal exchange and correlation functionals incorrectly describe the charge distribution around negative ions in water, e.g. Cl^-, yielding a delocalized charge distributed over the first shell of solvating water molecules. The use of hybrid functionals, that include a portion of Hartree-Fock exact exchange, may alleviate this problem and lead to a better localization of the charge around the anion.

We computed:

$$\Delta\rho = \rho_{SO_4^{2-}\cdot(H_2O)_n} - \rho_{SO_4^{-}\cdot(H_2O)_n}, \tag{6.3}$$

where $\rho_{SO_4^{2-}\cdot(H_2O)_n}$ is the charge density of the original cluster and $\rho_{SO_4^{-}\cdot(H_2O)_n}$ the charge density of a cluster with one eletron removed. Our results are reported in Fig. 6.7. $\Delta\rho$ may be regarded as the probability distribution of the quasi-particle HOMO of the system and describes the most probable distribution of the electron that is removed from the doubly charged cluster in a vertical ionization process. This charge distribution is in general different from that of the Kohn-Sham orbital corresponding to the highest eigenvalue in a DFT calculation. Similar to results reported for Cl^- [66], when using PBE $\Delta\rho$ is partially delocalized over the water molecules around the anion, for both the 12A and 12B configurations (see left panel of Fig. 6.7). The use of PBE0 leads instead to a more localized charge distribution around the sulfate (see right panel of Fig. 6.7).

One expects a higher bond order for the S-O bond and an higher vibrational frequency for its stretching mode when the electronic charge is more localized on the anion. Indeed our calculated IR spectra (Fig. 6.2 and 6.5) obtained with the PBE0 functional show a sulfate stretching mode at a higher frequency than PBE calculations, and in much better agreement with experiments. We note that the position of the sulfate stretching mode at 1100 cm^{-1} is a good indicator of the accuracy of the theory, as this peak appears to be decoupled from other water and sulfate vibrational modes. In addition, nuclear quantum effects may be regarded as negligible in this case, as only S and O atoms contribute to the mode.

Table 6.3: Calculated vertical ionization potential (eV) of 12A and 12B sulfate dianion clusters (see Fig. 6.1), obtained with semilocal (PBE) and hybrid (PBE0) functionals.

Functional	12A	12B
PBE	2.9	2.6
PBE0	3.9	3.6

In Table 6.3 we report the calculated vertical ionization potential (IP) of the 12 water cluster, defined as $IP = E_{SO_4^{2-}\cdot(H_2O)_n} - E_{SO_4^{-}\cdot(H_2O)_n}$. As expected, the IP obtained within

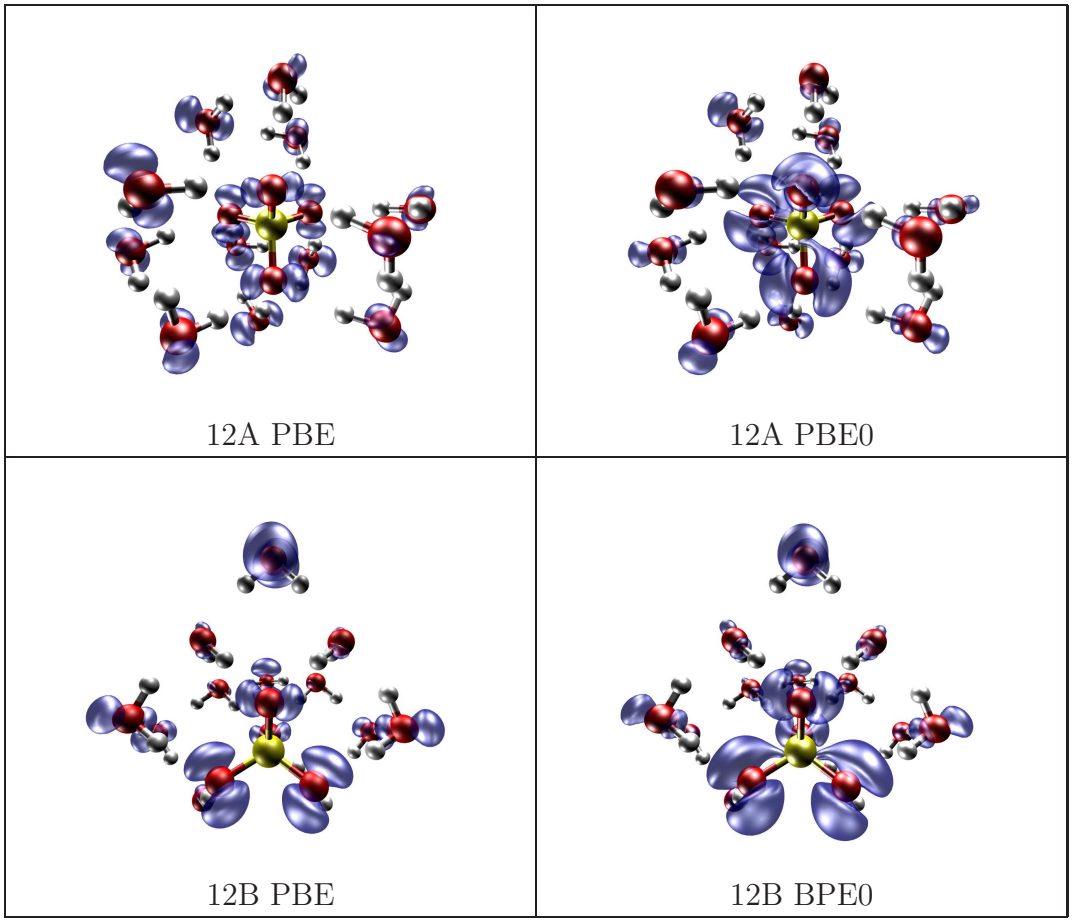


Figure 6.7: Charge density differences between the 12 water hydrated SO_4^{2-} and SO_4^{-} clusters in the 12 A (two upper panels) and 12B (two lower panels) configurations (see Fig. 6.1), obtained using the PBE (left panels) and PBE0 (right panels) functionals. Note the difference in charge localization obtained with semilocal and hybrid functionals. The same value of charge density is plotted in all cases.

PBE0 is higher than the value obtained at the PBE level of theory and in better agreement with the experimental value of ∼3.7 eV estimated from the position of the outermost peak of the photodetachment spectrum [203].

The computed value for 12B appears to be in better agreement with experiments, suggesting that this configuration may be the most probable one found in photodetachment experiments. However it is unclear whether the accuracy of the theory at the PBE0 level is sufficient for a definite assignment of the geometry based on computed IP compared to experiments. In addition, the computed vertical IPs of 12A and 12B differ by only 0.3

eV and the experimental peaks are rather broad (roughly 1eV), since the experiments were performed at room temperature. Most likely the experimental samples contain an ensemble of various configurations.

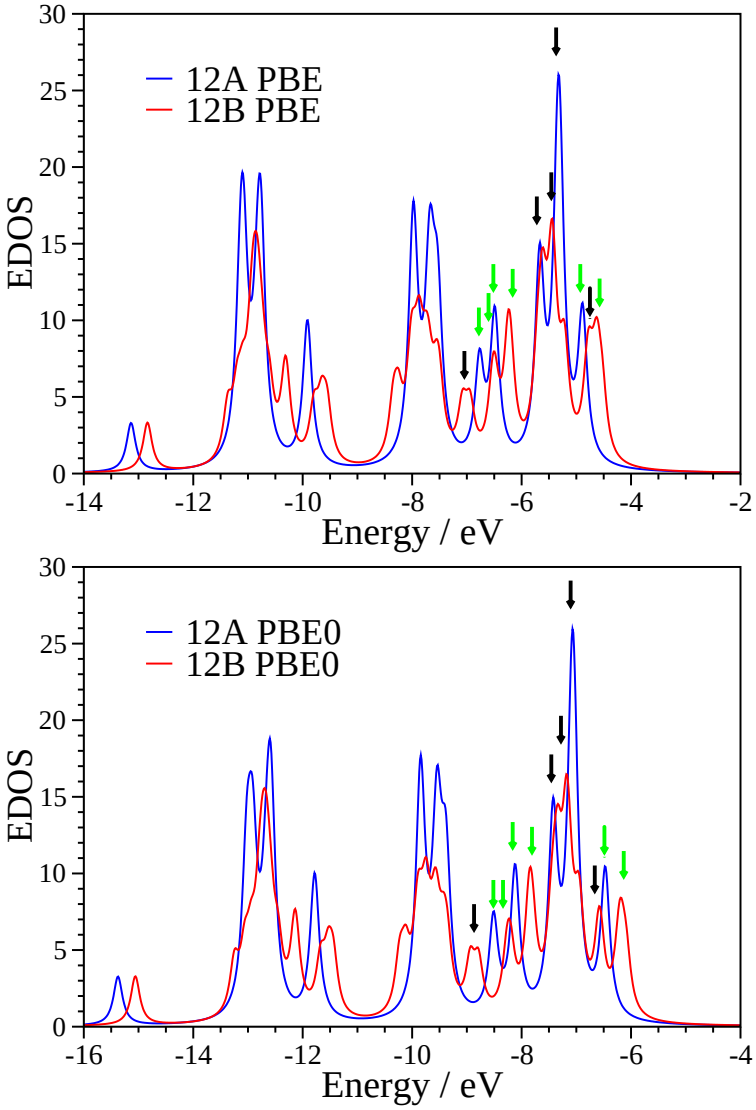


Figure 6.8: Electronic density of states (EDOS) of the 12A and 12B (Fig. 6.1) computed using the PBE (upper panel) and PBE0 (lower panel) exchange-correlation functionals. Green and black arrows indicate the first three orbitals belonging to sulfate ion, water and sulfate ion respectively (see text).

We also computed the electronic density of states (EDOS) of the 12A and 12B geometries: these are shown in Fig. 6.8, where we drew green and black arrows in correspondence of peaks

predominantly arising from anion states and from water states, respectively. PBE and PBE0 calculations yield similar EDOS distributions; however the energy separation between the uppermost anion and water states is substantially larger at the PBE0 level, indicating, as expected, a different level alignment between anion and water states at the two levels of theory. Interestingly the EDOS of the 12A and 12B geometries are also rather similar: in both cases the uppermost peak belongs to the anion, followed by water states, and then by an anion states about 1.8 eV lower in energy than the outermost one. Although a direct comparison between EDOS and photodetachment spectra is not possible (as no information about escape barriers is contained in the computed EDOS), our results strongly suggest that it is not possible to discern between internally solvated and surface solvated geometries of the 12 water hydrated dianion on the basis of photodetachment experiments at room temperature.

Finally we computed the vertical ionization potential using the PBE0 functional and the PBE single particle orbitals and vice versa. The results showed insensitivity to the orbitals used and a strong dependence on the functional, consistent with the findings of Ref. [67]. Therefore $\Delta\rho$ appears to be a reliable indicator of the accuracy of the theory for electronic properties calculations. We note that despite inaccurate results for ionization potentials and the fair predictions of vibrational frequencies, PBE can still provide qualitatively correct results on the structural properties and the overall shape of IR spectra.

6.4 Conclusions

We revisited the interpretation of spectroscopic (IR) and photodetachment experiments on hydrated sulfate anions with 12 and 13 water molecules. $SO_4^{2-} \cdot (H_2O)_{12}$ is a particularly stable cluster and was believed to be the first one of the series to present a full hydration shell. We also considered an additional water molecule (that is $SO_4^{2-} \cdot (H_2O)_{13}$) to investigate in detail the competing water-water and water-anion interactions. Using *ab initio* MD and electronic structure calculations we found that at zero temperature, surface hydrated config-

urations are more stable than fully solvated ones for both the 12 and 13 water clusters, by about 0.15 eV. Computed IR spectra suggest that at low temperature, the samples produced in IRMPD experiments are likely to contain an ensemble of surface and internally solvated geometries, with the less stable ones being kinetically trapped. Computed electronic density of states and ionization potentials for surface solvated and internally hydrated clusters are both fully consistent with photodetachment experiments conducted at room temperature. Therefore these experiments may not be used to discern between the different cluster geometries for the small cluster sizes considered here. Likewise, IR spectra may not be used to make definitive structural assignments as we found that different geometries may have similar vibrational signatures, with striking similarities especially in the case of 13 water clusters. It was previously reported that adding a water molecule to $SO_4^{2-} \cdot (H_2O)_{12}$ has a negligible influence on its first solvation shell. However our MD simulations showed that the presence of an extra molecule does affect the cluster. Our calculations indicate that it is only in clusters larger than previously believed that one observes a clear preference for internal solvation of the sulfate dianion, and that surface solvated geometries are likely to be present, possibly as the majority population, in several experiments with small clusters. Since internally and surface solvated clusters behave differently as nucleation sites in water droplets, our results may have implications in understanding aerosol processes that involve sulfate dianion. Finally we note that PBE0 calculations are superior to PBE ones in describing vibrational spectra and ionization potentials. However we found that PBE does provide qualitatively correct results on the structural properties and overall shape of IR spectra.

CHAPTER 7

ELECTRONIC STRUCTURE OF AQUEOUS SULFURIC ACID FROM FIRST PRINCIPLES SIMULATIONS WITH HYBRID FUNCTIONALS

7.1 Introduction

Aqueous solutions of sulfuric acid H_2SO_4 and hydrated sulfate SO_4^{2-} and bisulfate HSO_4^- anions are of interest to several fields of chemistry [50, 51, 224–226]. For example, small clusters of hydrated HSO_4^- and SO_4^{2-} play an important role in nucleation processes in the atmosphere [224]. The sulfonic acid group -SO_3H is the effective group of nafion ionomers suitable for proton-exchange membranes in fuel cells [225, 226]. In addition, sulfuric acid solutions have been widely used in photoelectrochemical water splitting reactions [50]. Recent studies [50, 51] on the effect of electrolytes on the efficiency of oxygen evolution reactions (OER) highlighted the importance of understanding the electronic structure of these solutions.

Several experimental investigations, including X-ray and neutron diffraction [227–230], appeared in the literature in the last two decades, addressing the structural and electronic properties of solvated sulfuric acid in water. Most studies focused on concentrated solutions, e.g. above 5 mol/L, and found a water oxygen-oxygen distance $\sim$ 0.2 Å shorter than in pure water [230] and a high proton mobility, i.e. one order of magnitude larger than that of any other common ions in aqueous solutions [227, 228]. Photoelectron emission experiments were employed to determine the ionization potential of SO_4^{2-} in aqueous solutions [231, 232], and the degree of dissociation ($\alpha^{HSO_4^-}$) of HSO_4^- in sulfuric acid solutions [233], which was also investigated by means of Raman spectroscopy [234, 235]. It was found that the value of $\alpha^{HSO_4^-}$ is very sensitive to the temperature (e.g. $\alpha^{HSO_4^-}$ varies from 0.68 to 0.37 in the

range 240 - 290 K at a concentration of 1.13 mol/L [235]). The values reported by different experiments [233, 235] agree well at concentrations above 2 mol/L. However, photoelectron measurements [233] overestimated Raman spectroscopy results [235] by about 10% at ~ 1 mol/L.

Theoretical studies, in particular *ab initio* molecular dynamics (AIMD) simulations [12, 236] with semilocal BLYP [236], HCTH/120 [236] and PBE [12] functionals, have so far been focused on the structural properties of sulfuric acid aqueous solutions at 0.9 mol/L. The values of $\alpha^{\mathrm{HSO}_4^-}$ reported in these simulations varied from 0 in a PBE simulation at 302 K [12], 0.1 in a BLYP simulation at 326 K [236] to 0.2 or 0.5 in two HCTH/120 simulations at 320 K [236], to be compared with $0.3-0.35$ measured with Raman spectroscopy at ~ 300 K [235]. The discrepancies with experiments may stem partly from the use of semilocal functionals, which at the chosen simulation temperatures of $300-320$ K yield an overstructured hydrogen bond network [20, 64] and sluggish water dynamics [236].

Despite recent progress in understanding the properties of sulfuric acid in water, a qualitatively correct *ab initio* description of the dissociation of sulfuric acid is not yet available, and a thorough knowledge of the electronic structure of aqueous solutions of sulfuric acid is still lacking, with no *ab initio* calculations of its electronic properties reported so far. One difficulty in performing accurate first-principles simulations of solvated anions, i.e. HSO_4^- and SO_4^{2-}, is posed by the so called delocalization error [66] from which many semilocal density functional approximations suffer. This error is especially pronounced for negatively charged systems such as solvated anions [65, 237], leading to the prediction of an incorrect localization of the charge density. One way to remedy the delocalization error is to use hybrid density functionals which include a combination of semilocal exchange-correlation potentials and Hartree-Fock non-local exchange (also known as exact exchange) [66]. The calculation of the latter in condensed phases is computationally very demanding, and several techniques [75, 238, 239] were proposed to reduce the computational cost. Among these, the recursive subspace bisection algorithm [75, 76] was shown to be both efficient for water [76] and NaCl

solutions [69] and is our method of choice.

In this chapter, we investigate the structural and electronic properties of a 0.87 mol/L sulfuric acid aqueous solution, a concentration similar to that ($\sim$ 1.0 mol/L) used in recent OER experiments [51]. We present the first AIMD simulation of sulfuric acid in water carried out with hybrid functionals (PBE0 [68]), and we compare our results with those obtained with semilocal ones (PBE [59, 60]). We found that only PBE0 yields a qualitatively correct degree of dissociation of the HSO_4^- ion in water, while PBE largely overestimates the concentration of SO_4^{2-} in the solution. We also reported the first study of the electronic properties of the solution, so far investigated only experimentally: based on the relative alignment between the highest occupied molecular orbital (HOMO) of the anions and the valence band maximum (VBM) of water, we discuss the possible mechanisms involved in OER experiments.

7.2 Computational Methods

We considered a cubic supercell of linear demension of 12.42 Å, containing one deuterated sulfuric acid molecule and 62 heavy water molecules at the experimental density of 1.1 g/cm^3. (For simplicity, deuterium atoms are referred to with the element symbol H instead of D throughout the paper.) We carried out Born-Oppenheimer AIMD simulations with the Qbox code [46] and a time step of 10 a.u. (0.24 fs). We used a planewave basis set with a kinetic energy cutoff of 85 Ry, norm-conserving pseudopotentials [71, 72] and we sampled the Brillouin zone with only the Γ point. In both PBE and PBE0 simulations we considered that the total energy was converged in the electronic structure optimization when the total energy difference between two consecutive steps was less than 1×10^{-7} atomic unit. When using the PBE functional, we collected statistics over a 38 ps NVE simulation, after a 15 ps equilibration in the NVT ensemble with a target temperature of 380 K maintained by a Bussi-Donadio-Parrinello thermostat [151]. We chose an elevated temperature to reproduce the measured structural properties of water at room temperature [10]. Because of the high computational cost, when using the PBE0 functional, we utilized

the recursive subspace bisection algorithm [75, 76] with a threshold of 0.02 to speed up the calculations. We collected statistics on a 15 ps NVT trajectory, after 15 ps equilibration using the same target temperature and thermostat as with PBE. We note that statistical averages over trajectories obtained in the NVE and NVT ensembles yield the same results, provided the simulations are carried out at the same average temperature and density. The PBE0 simulations were carried out in the NVT ensemble for reasons of efficiency, after comparing the results obtained in the NVE and NVT ensemble for another solution containing anions (see Fig. 3 and 4 of Ref. [69]. NVE and NVT simulations yielded indistinguishable results for the structural and electronic properties examined in Ref. [69]). The averaged temperatures over the production runs using PBE and PBE0 functionals were chosen to be the same and they turned out to be 376 ± 18 K and 376 ± 21 K respectively.

7.3 Results and Discussion

7.3.1 Structural Properties

We first discuss the structural properties of the solution. The computed water oxygen-oxygen and oxygen-hydrogen radial distribution functions (RDF) are compared in Fig. 7.1a and 7.1b, in the presence and absence of the dissolved acid. Our results indicate that at the concentration considered here, the presence of the sulfuric acid does not significantly alter the average liquid water structure. We also found that the PBE0 simulation yields a less structured water and ion-water RDF than PBE (see Fig. 7.1), consistent with the results reported for solvated chloride ions [65] and recent simulations of water with PBE0+vdW functionals [240]. Note that the RDFs obtained in our work with the PBE functional are in better agreement with experiments than those of Ref. [12], where simulations were carried out at a much lower temperature (~ 300 K). We found a solvation shell of ~ 8 water molecules around the anion, with PBE0, compared to ~ 10 obtained from PBE simulations (Fig. 7.1c). A comparison with experiment is not straightforward, as the coordination number of SO_4^{2-}

reported by X-ray diffraction experiments [229] varies between 7 and 12, due to difficulties in defining a clear minimum in the correlation function of the anion and water molecules in the liquid [241]. The oxygen-oxygen RDF obtained with PBE0 yields a better agreement with that of recent X-ray diffraction experiments on pure liquid water [13] than the PBE one. However, the optimal temperature required for PBE0 simulations to reproduce experimental RDFs at ambient conditions is not known [10]. Considering that the computed PBE0 oxygen-oxygen RDF is slightly understructured compared to experiment [13], such temperature is estimated to be about 350 - 360 K, slightly lower than the one used here, i.e. 376 K, which was chosen to be the same as that of our PBE simulations.

7.3.2 Dissociation of Sulfuric Acid

One common property of polyprotic acids is their ability to release multiple hydronium ions in aqueous solutions. Sulfuric acid can release two hydronium cations in the following two consecutive reactions:

$$H_2SO_4(aq) + H_2O \rightarrow HSO_4^-(aq) + H_3O^+(aq), \tag{7.1}$$

$$HSO_4^-(aq) + H_2O \rightarrow SO_4^{2-}(aq) + H_3O^+(aq). \tag{7.2}$$

At ambient conditions and at the simulated concentration, the reaction 7.1 is considered complete, i.e. the neutral species H_2SO_4 is not present in solution, and only the HSO_4^- and SO_4^{2-} anions are. The degree of dissociation of HSO_4^- is defined as $\alpha^{HSO_4^-} = [SO_4^{2-}]/\big([HSO_4^-] + [SO_4^{2-}]\big)$. Raman [234, 235] experiments indicate that in sulfuric acid aqueous solutions at 300K and a concentration equal to the simulated one, $\alpha^{HSO_4^-}$ is about $0.3 - 0.35$. The values reported by photoelectron emission experiments were not used for comparison due to errors arising from the low signal-to-noise ratio at this concentration [233].

We computed the value of $\alpha^{HSO_4^-}$ for both of our simulations. In order to distinguish between HSO_4^- and SO_4^{2-} we used a simple geometrical criterion, where we assigned each

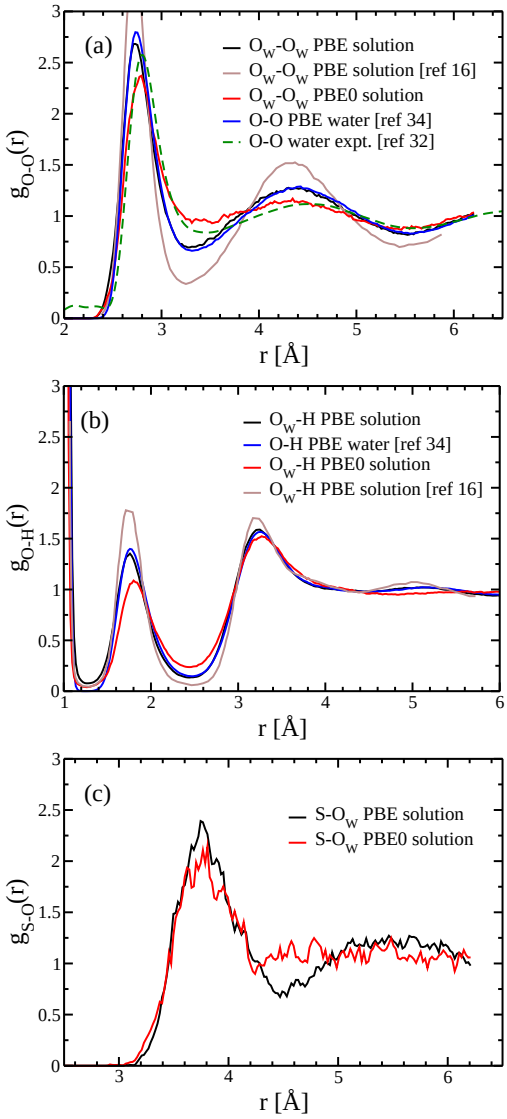


Figure 7.1: Oxygen-oxygen (a) , oxygen-hydrogen (b) and sulfur-oxygen (c) radial distribution functions (RDF), $g_{\mathrm{O-O}}(r)$, $g_{\mathrm{O-H}}(r)$ and $g_{\mathrm{S-O}}(r)$; $\mathrm{O_W}$ denotes oxygen atoms belonging to the water molecules in the liquid. Red and black solid curves denote our results for a 0.87 mol/L sulfuric acid solution obtained with simulations at 376 K using the PBE0 and PBE functionals, respectively. We compare our RDFs with those reported in Ref. [12] (brown curve). We also report the $g_{O-O}(r)$ of pure heavy water from a PBE simulation at 378 K [2] (blue curve) and the experimental water correlation function obtained from recent X-ray diffraction experiments at 295 K [13] (dashed green curve).

H atoms to its nearest neighboring oxygen atom, either a water oxygen or a sulfuric acid oxygen. In agreement with experiments [235], we did not find any neutral H_2SO_4 species. The computed values of $\alpha^{\mathrm{HSO_4^-}}$ are 0.73 and 0.25, at the PBE and PBE0 levels respectively

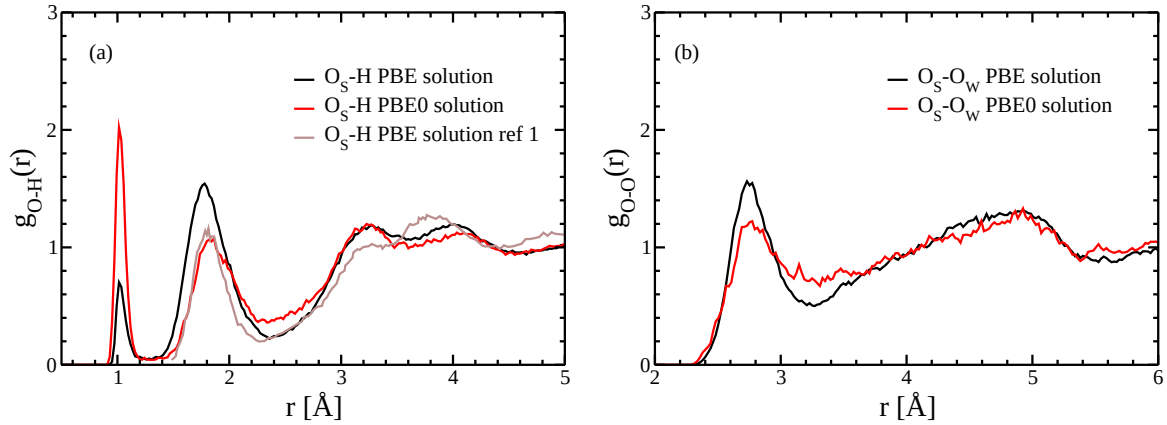


Figure 7.2: Computed oxygen-hydrogen (O_S-H) (a) and oxygen-water oxygen (O_S-O_W) radial distribution functions (RDFs), where O_S and O_W denote oxygen atoms belonging to sulfuric acid and to water molecules, respectively. Red and black solid curves denote our results for a 0.87 mol/L solution obtained with the PBE0 and PBE functionals, respectively. We compare our O_S-H RDF with that reported in Ref. [12] (brown curve).

(see top panels of Fig. 7.3). There is a sizable error associated to these computed values, estimated to be of the order of ± 0.1; this it due to the size of our simulation cell and the length of our simulations. But the comparison between the PBE and PBE0 results is meaningful. Clearly, our PBE0 simulation gives a more reasonable estimate of $\alpha^{HSO_4^-}$, in qualitative (and possibly quantitative) agreement with experiments. The qualitatively incorrect result obtained with the PBE functional is not surprising, and it is likely due to the delocalization error [66] that causes an overstabilization of the species with the highest charged state, e.g. SO_4^{2-}, and thus results in an overestimated $\alpha^{HSO_4^-}$. Hybrid functionals partly correct the delocalization error and give a qualitatively correct descriptions of the equilibrium between HSO_4^- and SO_4^{2-}. We note that Choe *et al.* [12] reported $\alpha^{HSO_4^-} = 0$ using AIMD and the PBE functional at the same sulfuric acid concentration, most likely because they simulated at a lower T (300 K), at which the PBE functional yields a severely overstructured liquid (see Fig. 7.1 (a) and (b)) with slow diffusion. AIMD simulations [236] with BLYP and HCTH/120 functionals at $\sim$ 320 K may suffer from the same error as well. These simulations [12, 236] yielded water self diffusion coefficients varying from 0.05×10^{-5} to 0.5×10^{-5} cm^2/s, much lower than the experimental value [242] of pure heavy water 1.9×10^{-5} cm^2/s; instead the

values computed in our simulations are $\sim 1.5\times10^{-5}$ and $\sim 2.7\times10^{-5}$ cm^2/s at the PBE and PBE0 levels of theory, respectively, in qualitative agreement with experiments. However we note that definitive conclusions on the values of the diffusion coefficients may most likely not be attained by using only a single trajectory of several tens of ps. Therefore only qualitative comparisons are meaningful between the numbers reported here.

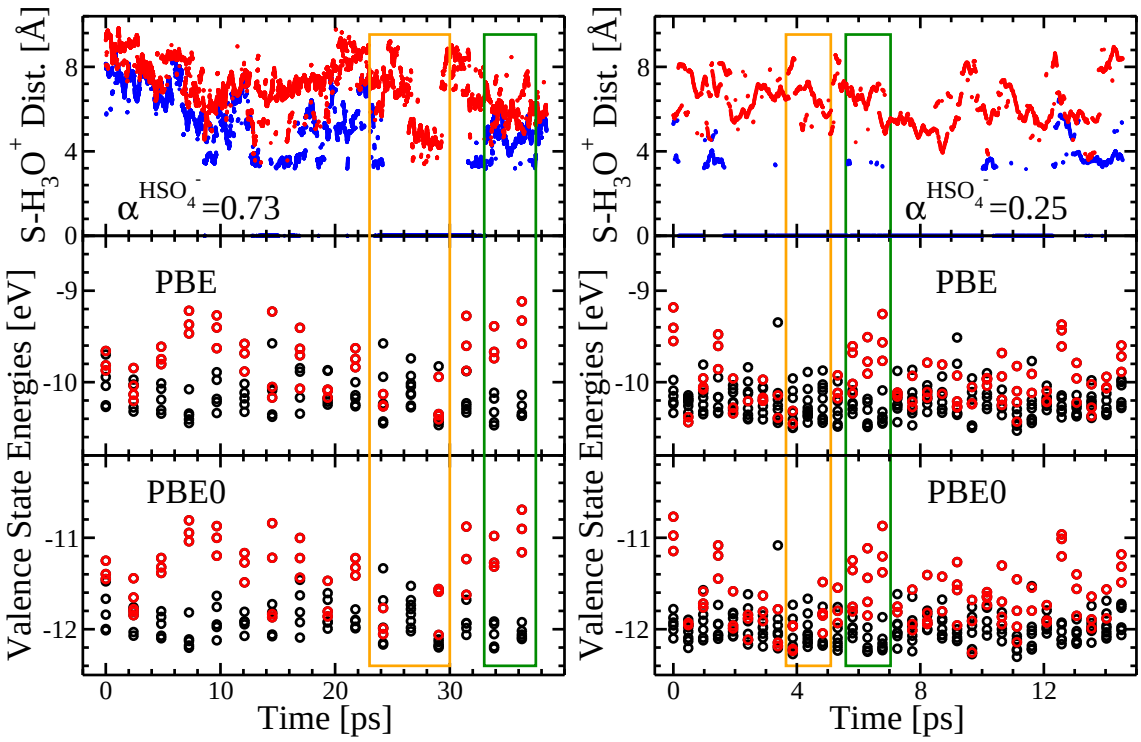


Figure 7.3: Distances (top panel) and single particle energies (middle and bottom panels) of a sulfuric acid aqueous solution (0.87 mol/L) computed from trajectories obtained at the PBE (left panels) and PBE0 (right panels) level of theory. In the upper panel the red and blue curves show the distance between the sulfur atom and the oxygen atom belonging to its first and second hydronium ion neighbor, respectively (the blue curve is zero when only one hydronium ion is present in the simulation). The middle and bottom panels show Kohn-Sham eigenvalues of the highest eight occupied orbitals (black circles), with states localized on the anion (HSO_4^- or SO_4^{2-}) shown in red. The electronic structure was computed using the PBE (middle panels) and PBE0 (bottom panels) functionals on samples extracted from PBE (left) and PBE0 (right) trajectories. The orange and green boxes highlight several typical portions of our simulations where the HOMO of the anion is significantly below or above the water VBM, respectively.

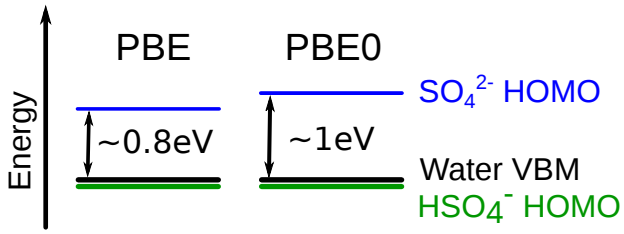


Figure 7.4: Relative positions of the highest occupied molecular orbital (HOMO) of the HSO_4^- (green) and SO_4^{2-} (blue) anions, with respect to the valence band maximum (VBM) of water (black), as obtained in our simulations with the PBE and PBE0 functional.

7.3.3 Electronic Properties

We now turn to the discussion of the electronic properties of the sulfuric acid solutions. We computed the Kohn-Sham eigenvalues of several single particle states near the VBM of water, for various snapshots extracted from both PBE and PBE0 simulations (every 2 (0.5) ps in the case of PBE (PBE0) trajectories). We assigned single particle states to water molecules or anions by establishing where the states are localized in real space.

In both PBE and PBE0 simulations (see Fig. 7.4), the energy of the HOMO of the doubly charged anion SO_4^{2-} is well above the VBM of water (~ 0.8 eV with PBE and ~ 1 eV with PBE0), while that of the singly charged anion HSO_4^- fluctuates close to that of the VBM of water. Our results are also consistent with previous stability studies of the SO_4^{2-} anion [209, 210, 215]. It was shown that a bare SO_4^{2-} is unstable unless coordinated by at least three water molecules, while a bare HSO_4^- anion is stable [243].

At the PBE level of theory (see left panels of Fig. 7.3), we found that the SO_4^{2-} ion is the majority species in the solution and the HOMO is mostly above the water VBM; this result is in qualitative agreement with photoelectron emission experiments [232] placing the HOMO of SO_4^{2-} about 1.4 eV above the VBM of water. The energy difference $\Delta E_{\text{HOMO-VBM}}$ between the HOMO of SO_4^{2-} and the water VBM predicted by the PBE0 functional is ~ 1 eV, that is ~ 0.2 eV larger than that obtained with the PBE functional (see right panels of Fig. 7.3), and in better agreement with experiments. This difference between PBE and

PBE0 functionals can be understood from the different localization properties of the HOMO orbitals predicted by the two functionals, which can be characterized by computing the inverse participation ratio (IPR). The IPR of an orbital σ is defined as $\mathrm{IPR}_\sigma = \frac{1}{\int |\psi_\sigma(\mathbf{r})|^4 d\mathbf{r}}$. The IPR of a delocalized state tends to 1 and the larger the IPR, the more localized the orbital. The IPR of the HOMO orbital averaged over configurations extracted from the PBE0 simulation computed with PBE and PBE0 functionals are 712 and 774 respectively. Not surprisingly the PBE functional predicts a more delocalized HOMO orbital than PBE0. The discrepancies with experiments for the value of $\Delta E_{\text{HOMO-VBM}}$ are not surprising due to inaccuracies in describing band gaps and energy levels in water, both at the PBE and PBE0 levels of theory [65, 244]. However the error found here for the position of the anion SO_4^{2-} ($\sim$ 1 eV versus a measured value of 1.4 eV) is much smaller than the one reported for Cl^- [65]($\sim$ 0.2 eV versus 1.3 eV). This may be due to the fact that while SO_4^{2-} anions are preferably solvated in the bulk of liquid water, Cl^- may lie closer to the surface [30]; therefore the energies detected for Cl^- in photoelectron experiments are more likely probing the levels of the anions at the surface of water, not in the bulk [245].

The energy alignment between anions and water found in our simulations (Fig. 7.4) is helpful to understand the oxidation reactions occurring at photoanodes in water splitting reactions. As discussed by Mi *et al.* [51], sulfuric acid was used as the electrolyte for OER on tungsten oxide, which is a promising photoanode material for water splitting. Instead of oxidizing water to produce oxygen after electron-hole pairs were created by visible light illumination, holes were found to oxidize anions to produce $S_2O_8^{2-}$ [51]. This oxidation reaction could follow one of the following two competing paths:

$$2\mathrm{HSO_4^-(aq)} \rightarrow \mathrm{S_2O_8^{2-}(aq)} + 2\mathrm{H^+} + 2\mathrm{e^-},\ \mathrm{E^\circ} = 2.12\ \mathrm{V}, \tag{7.3}$$

$$2\mathrm{SO_4^{2-}(aq)} \rightarrow \mathrm{S_2O_8^{2-}(aq)} + 2\mathrm{e^-},\ \mathrm{E^\circ} = 2.01\ \mathrm{V}, \tag{7.4}$$

where $\mathrm{E^\circ}$ indicates the standard redox potential of each half reaction. The difference in

redox potentials between SO_4^{2-} and HSO_4^- is only 0.11 V. Because these two anions exist under equilibrium in aqueous sulfuric acid solutions, and reactions 7.3 and 7.4 describe the same equilibrium at different reaction quotients, they could both be directly involved in the charge transfer reaction that scavenges holes in the photoexcited tungsten oxide electrode. The competition between the two anions is determined ultimately by the kinetics of the charge transfer from the anion to the surface. Our calculations showed that the energy levels of SO_4^{2-} are better positioned to donate electrons. Therefore we expect reaction 7.4 to be favored over 7.3, although the details of the interface between tungsten oxide and the aqueous solution need to be considered before one can make definitive conclusions on the rates of charge transfer reactions. This analysis suggests that in future simulations of photocatalyst-aqueous solution interfaces, SO_4^{2-} should be considered the focus of the study in the electrolyte. We expect the same analysis to apply to triply charged phosphate anions in water, as they share similar oxidation rates [50] and ionization thresholds [232] with SO_4^{2-}, although the fraction of triply deprotonated PO_4^{3-} becomes substantial only at pH > 12.

7.4 Conclusions

In conclusion, we carried out AIMD simulations of sulfuric acid in water at a concentration ($\sim$ 1 mol/L) close to that of electrolyte solutions used in photocatalytic water splitting reactions. We presented the first *ab initio* simulation with hybrid functionals (PBE0) and we compared our results with those obtained with a semilocal functional (PBE). We found a qualitatively correct degree of dissociation of the HSO_4^- ion at the PBE0 level of theory, while PBE severely overestimated its dissociation degree. Consistent with the results of AIMD simulations of the halide ion Cl^- in water [65, 69], our findings show that the use of hybrid functionals is essential to obtain qualitatively correct results, when solvated anions are present in water. We also investigated the electronic structure of the solution, which previously had been studied only experimentally. We evaluated the relative energy alignment between the HOMO levels of the SO_4^{2-} and HSO_4^- anions and the VBM of liquid water.

Within PBE0 and PBE we found that the HOMO of SO_4^{2-} is about 1 eV and 0.8 eV above the water VBM, respectively, consistent with the 1.4 eV value reported by photoelectron emission experiments. The energy of the HOMO level of HSO_4^-, on the other hand, fluctuates around that of the water VBM. Our findings indicate that SO_4^{2-} is more susceptible to oxidation than HSO_4^- by photoanodes used in water oxidization experiments, and we expect the same results would hold for a triply charged phosphate anion in water. Work is in progress to improve on the electronic structure description of anions in water, by carrying out many body perturbation theory calculations with the GW approximation, on snapshots extracted from AIMD trajectories obtained using the PBE0 functional [65].

CHAPTER 8

CONCLUSIONS

The work presented in this dissertation encompasses two main aspects: (1) the development and implementation of first-principles methods to compute the response of ordered and disordered condensed systems to applied electric fields, and their use to simulate, from first principles, vibrational spectra of aqueous systems, including Raman and SFG spectra; (2) the investigation of the structural, electronic and vibrational properties of aqueous systems, including liquid water, ice surfaces and simple ions in solution.

In terms of method developments, I first implemented the density functional perturbation theory (DFPT) method [43] in the FPMD simulation software Qbox [46] to compute the linear response to electric fields in semiconductor and insulators, while carrying out *ab initio* simulations. I then proposed an accurate first-principles method to carry out electronic structure calculations in the presence of finite electric fields, by combining the refinement scheme [49] based on MLWF [48] and an electric enthalpy functional [246]. This method works equally well with GGA and hybrid functionals and can be used for the calculation of multipole moments, e.g. quadrupole moments. I showed that the multipole moments and polarizabilities computed using this method converge much faster with respect to cell size than conventional approaches based on Berry phase and MLWF techniques, and the results are in very good agreement with experiments.

The capability of computing multipole moments and their polarizabilities in the condensed phase allowed for the development of a robust computational framework to compute SFG spectra of semiconductors and insulators from first principles. For the first time, we were able to consistently take into account the varying electric field across the interface by introducing a local dielectric constant. Higher multipole contributions from both surface and bulk regions were also included, by simulating slab and bulk samples, respectively. In addition I provided a simple solution to the origin dependence problem in the calculation of quadrupole contributions, which had been an unsolved issue in previous studies.

Using these methods, I carried out a series of first-principles simulations to model vibrational spectra of aqueous systems. I computed Raman spectra of liquid heavy water, and the results obtained are in good agreement with experiments and previous simulations [143, 144]. By studying intra- and intermolecular contributions to the Raman spectra, as well as the power spectra of MLWF spread and of molecular polarizabilities, I identified the presence of intermolecular charge fluctuations at 270 cm^{-1}, accompanying intermolecular HB stretching modes; finally I showed the importance of using explicit electronic structure methods in obtaining the Raman spectra.

I carried out first principles simulations for SFG spectra of ice I_h basal surfaces and obtained good agreement with experiments. I found that including three ice bilayers (BL) in surface calculations is sufficient to converge the spectra, and the first BL contributes to only about 50% of the total SFG intensities. The results show that bulk vibrations yield contributions to the SFG spectra of the *ppp* polarization combinations. I found that failing to take into account the varying electric field at the surface can substantially change these spectra. These results provide guidance for future simulations of SFG spectra of aqueous systems and the interpretation of experimental results.

In addition to vibrational spectra of water and ice, I also studied structural, electronic and vibrational properties of simple ions in water. In particular, I investigated hydrated sulfate clusters with 12 and 13 water molecules, and interpreted infrared and photoelectron experiments. Our results showed that different configurations exhibit similar infrared and photoelectron spectral features and may be difficult to distinguish based on these measurements. Using hybrid functional calculations, we determined that a surface solvated configuration is energetically more stable than the internally solvated one. Since internally and surface solvated clusters behave differently as nucleation sites in water droplets, our results have implications in understanding aerosol processes that involve sulfate ions.

I also extended our study of ion-water interactions to the condensed phase and carried out FPMD simulations of sulfuric acid with both GGA PBE and hybrid PBE0 functionals.

We found a qualitatively correct degree of dissociation of the HSO_4^- ion at the PBE0 level of theory, while GGA severely overestimated its dissociation degree, due to the delocalization error [66]. With the PBE0 functional we found that the HOMO of SO_4^{2-} is about 1 eV above the water VBM; the HOMO of HSO_4^-, on the other hand, fluctuates around that of the water VBM. Our findings indicate that SO_4^{2-} may be more susceptible to oxidation than HSO_4^- by photoanodes used in water oxidization experiments.

In summary, this dissertation presented both method developments and applications of FPMD for the calculation of structural, electronic and vibrational properties of aqueous systems. The methods developed here for Raman and SFG spectroscopies are general and are applicable to ordered and disordered semiconductors and insulators. In addition they may be applied using GGA and hybrid functionals and they are easily generalizable to include van der Waals functionals as well [10, 20]. Future work includes the application of the methods developed here to simulate SFG spectra of surfaces of water and aqueous solutions, and the investigation of aqueous surfaces of relevance to atmospheric chemistry [30, 247].

REFERENCES

[1] D. E. Woon and T. H. Dunning Jr., J. Chem. Phys. **100**, 2975 (1994).

[2] Q. Wan, L. Spanu, G. Galli, and F. Gygi, J. Chem. Theory Comput. **9**, 4124 (2013).

[3] J. R. Scherer, M. K. Go, and S. Kint, J. Phys. Chem. **78**, 1304 (1974).

[4] V. Zhelyaskov, G. Georgiev, Z. Nickolov, and M. Miteva, J. Raman Spectrosc. **20**, 67 (1989).

[5] M. Paolantoni, P. Sassi, A. Morresi, and S. Santini, J. Chem. Phys. **127**, 024504 (2007).

[6] S. G. Warren and R. E. Brandt, J. Geophys. Res. **113**, D14220 (2008).

[7] X. Wei, P. Miranda, C. Zhang, and Y. Shen, Phys. Rev. B **66**, 085401 (2002).

[8] P. J. Bisson and M. J. Shultz, J. Phys. Chem. A **117**, 6116 (2013).

[9] S. Habershon, G. S. Fanourgakis, and D. E. Manolopoulos, J. Chem. Phys. **129**, 074501 (2008).

[10] C. Zhang, D. Donadio, G. Galli, and F. Gygi, J. Chem. Theory Comput. **7**, 1443 (2011).

[11] J. Zhou, G. Santambrogio, M. Brümmer, D. T. Moore, L. Wöste, G. Meijer, D. M. Neumark, and K. R. Asmis, J. Chem. Phys. **125**, 111102 (2006).

[12] Y.-K. Choe, E. Tsuchida, and T. Ikeshoji, J. Chem. Phys. **126**, 154510 (2007).

[13] L. B. Skinner, C. Huang, D. Schlesinger, L. G. M. Pettersson, A. Nilsson, and C. J. Benmore, J. Chem. Phys. **138**, 074506 (2013).

[14] A. P. Gaiduk, F. Gygi, and G. Galli, J. Phys. Chem. Lett. **6**, 2902 (2015).

[15] G. E. Walrafen, in *Water: A Comprehensive Treatise. The Physics and Physical Chemistry of Water*, v. 1, edited by F. Franks (Plenum Press, 1974).

[16] J. Eaves and J. Loparo, Proc. Natl. Acad. Sci. U. S. A. **102**, 13019 (2005).

[17] J. Morrone and R. Car, Phys. Rev. Lett. **101**, 017801 (2008).

[18] F. Paesani, S. S. Xantheas, and G. A. Voth, J. Phys. Chem. B **113**, 13118 (2009).

[19] T. Spura, C. John, S. Habershon, and D. K. Thomas, Mol. Phys. **113**, 808 (2014).

[20] C. Zhang, J. Wu, G. Galli, and F. F. Gygi, J. Chem. Theory Comput. **7**, 3054 (2011).

[21] S. Grimme, J. Comput. Chem. **27**, 1787 (2006).

[22] A. Pasquarello and R. Resta, Phys. Rev. B **68**, 174302 (2003).

[23] M. Sharma, R. Resta, and R. Car, Phys. Rev. Lett. **95**, 187401 (2005).

[24] H. Torii, J. Phys. Chem. B **115**, 6636 (2011).

[25] H. J. Bakker, Chem. Rev. **108**, 1456 (2008).

[26] H. J. Bakker and J. L. Skinner, Chem. Rev. **110**, 1498 (2010).

[27] Y. R. Shen and V. Ostroverkhov, Chem. Rev. **106**, 1140 (2006).

[28] X. Wei, P. Miranda, and Y. Shen, Phys. Rev. Lett. **86**, 1554 (2001).

[29] C. Tian and Y. Shen, Phys. Rev. Lett. **101**, 139401 (2008).

[30] C. Tian, S. J. Byrnes, H.-L. Han, and Y. R. Shen, J. Phys. Chem. Lett. **2**, 1946 (2011).

[31] D. A. McQuarrie, *Statistical Mechanics* (University Science Books, 2000) p. 484.

[32] R. Car and M. Parrinello, Phys. Rev. Lett. **55**, 2471 (1985).

[33] A. Hassanali, M. K. Prakash, H. Eshet, and M. Parrinello, Proc. Natl. Acad. Sci. **108**, 20410 (2011).

[34] A. Hassanali, F. Giberti, J. Cuny, T. D. Kühne, and M. Parrinello, Proc. Natl. Acad. Sci. U. S. A. **110**, 13723 (2013).

[35] C. Zhang, D. Donadio, and G. Galli, J. Phys. Chem. Lett. **1**, 1398 (2010).

[36] M. Sharma, R. Resta, and R. Car, Phys. Rev. Lett. **98**, 247401 (2007).

[37] D. Pan, L. Spanu, B. Harrison, D. A. Sverjensky, and G. Galli, Proc. Natl. Acad. Sci. U. S. A. **110**, 6646 (2013).

[38] L. Spanu, D. Donadio, D. Hohl, E. Schwegler, and G. Galli, Proc. Natl. Acad. Sci. U. S. A. **108**, 6843 (2011).

[39] D. Pan, Q. Wan, and G. Galli, Nat. Commun. **5**, 3919 (2014).

[40] A. Putrino and M. Parrinello, Phys. Rev. Lett. **88**, 176401 (2002).

[41] M. Pagliai, C. Cavazzoni, G. Cardini, G. Erbacci, M. Parrinello, and V. Schettino, J. Chem. Phys. **128**, 224514 (2008).

[42] M. Sulpizi, M. Salanne, M. Sprik, and M.-P. Gaigeot, J. Phys. Chem. Lett. **4**, 83 (2013).

[43] S. Baroni, S. De Gironcoli, A. Dal Corso, and P. Giannozzi, Rev. Mod. Phys. **73**, 515 (2001).

[44] R. Martin, *Electronic Structure: Basic Theory and Practical Methods* (Cambridge University Press, 2004).

[45] I. Souza, J. Íñiguez, and D. Vanderbilt, Phys. Rev. Lett. **89**, 117602 (2002).

[46] "Qbox Code," `http://eslab.ucdavis.edu/software/qbox/`, accessed: 2015-06-09.

[47] F. Gygi, IBM J. Res. Dev. **52**, 137 (2008).

[48] N. Marzari, A. Mostofi, J. Yates, I. Souza, and D. Vanderbilt, Rev. Mod. Phys. **84**, 1419 (2012).

[49] M. Stengel and N. Spaldin, Phys. Rev. B **73**, 075121 (2006).

[50] J. C. Hill and K.-S. Choi, J. Phys. Chem. C **116**, 7612 (2012).

[51] Q. Mi, A. Zhanaidarova, B. S. Brunschwig, H. B. Gray, and N. S. Lewis, Energy Environ. Sci. **2**, 5694 (2012).

[52] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).

[53] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).

[54] H. Nakashima and H. Nakatsuji, J. Chem. Phys. **127**, 224104 (2007).

[55] D. M. Ceperley and B. J. Alder, Phys. Rev. Lett. **45**, 566 (1980).

[56] J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).

[57] A. D. Becke, Phys. Rev. A **38**, 3098 (1988).

[58] C. Lee, W. Yang, and R. G. Parr, Phys. Rev. B **37**, 785 (1988).

[59] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).

[60] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **78**, 1396 (1997).

[61] P. L. Silvestrelli, N. Marzari, D. Vanderbilt, and M. Parrinello, Solid State Commun. **107**, 4 (1998).

[62] P. Silvestrelli and M. Parrinello, Phys. Rev. Lett. **82**, 3308 (1999).

[63] W. Chen, M. Sharma, R. Resta, G. Galli, and R. Car, Phys. Rev. B **77**, 245114 (2008).

[64] C. Zhang, L. Spanu, and G. Galli, J. Phys. Chem. B **115**, 14190 (2011).

[65] C. Zhang, T. A. Pham, F. Gygi, and G. Galli, J. Chem. Phys. **138**, 181102 (2013).

[66] A. J. Cohen, P. Mori-Sánchez, and W. Yang, Science **321**, 792 (2008).

[67] A. J. Cohen, P. Mori-Sánchez, and W. Yang, Chem. Rev. **112**, 289 (2012).

[68] C. Adamo and V. Barone, J. Chem. Phys. **110**, 6158 (1999).

[69] A. P. Gaiduk, C. Zhang, F. Gygi, and G. Galli, Chem. Phys. Lett. **604**, 89 (2014).

[70] B. Santra, J. Klime, A. Tkatchenko, D. Alf, B. Slater, A. Michaelides, R. Car, and M. Scheffler, J. Chem. Phys. **139**, 154702 (2013).

[71] D. Hamann, M. Schlüter, and C. Chiang, Phys. Rev. Lett. **43**, 1494 (1979).

[72] D. Vanderbilt, Phys. Rev. B **32**, 8412 (1985).

[73] E. D. Murray and G. Galli, Phys. Rev. Lett. **108**, 105502 (2012).

[74] F. Gygi, Journal of Physics: Conference Series **46**, 268 (2006).

[75] F. Gygi, Phys. Rev. Lett. **102**, 166406 (2009).

[76] F. Gygi and I. Duchemin, J. Chem. Theory Comput. **9**, 582 (2013).

[77] Q. Wan, L. Spanu, F. Gygi, and G. Galli, J. Phys. Chem. Lett. **5**, 2562 (2014).

[78] L. Verlet, Phys. Rev. **159**, 98 (1967).

[79] D. Marx and J. Hutter, *Ab Initio Molecular Dynamics: Basic Theory and Advanced Methods* (Cambridge University Press, 2009).

[80] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes 3rd Edition: The Art of Scientific Computing*, 3rd ed. (Cambridge University Press, 2007).

[81] M. Hybertsen and S. Louie, Phys. Rev. B **35**, 5585 (1987).

[82] A. Dal Corso, S. Baroni, and R. Resta, Phys. Rev. B **49**, 5323 (1994).

[83] D. Vanderbilt and R. King-Smith, Phys. Rev. B **48**, 4442 (1993).

[84] R. King-Smith and D. Vanderbilt, Phys. Rev. B **47**, 1651 (1993).

[85] R. Resta, Rev. Mod. Phys. **66**, 899 (1994).

[86] P. Umari and A. Pasquarello, Phys. Rev. B **68**, 1 (2003).

[87] R. Resta, Phys. Rev. Lett. **80**, 1800 (1998).

[88] A. Putrino, D. Sebastiani, and M. Parrinello, J. Chem. Phys. **113**, 7102 (2000).

[89] P. Umari and A. Pasquarello, Phys. Rev. Lett. **89**, 157602 (2002).

[90] N. Marzari and D. Vanderbilt, Phys. Rev. B **56**, 847 (1997).

[91] P. L. Silvestrelli and M. Parrinello, J. Chem. Phys. **111**, 3572 (1999).

[92] I. Souza, T. Wilkens, and R. Martin, Phys. Rev. B **62**, 1666 (2000).

[93] F. Gygi, J.-L. Fattebert, and E. Schwegler, Comput. Phys. Commun. **155**, 1 (2003).

[94] "NIST Computational Chemistry Comparison and Benchmark Database," http://cccbdb.nist.gov/.

[95] A. Hoy and P. Bunker, J. Mol. Spectrosc. **74**, 1 (1979).

[96] M. Valiev, E. J. Bylaska, N. Govind, K. Kowalski, T. P. Straatsma, H. J. J. Van Dam, D. Wang, J. Nieplocha, E. Apra, T. L. Windus, and W. A. De Jong, Comput. Phys. Commun. **181**, 1477 (2010).

[97] J. Verhoeven and A. Dymanus, J. Chem. Phys. **52**, 3222 (1970).

[98] K. Shiratori and A. Morita, Bull. Chem. Soc. Jpn. **85**, 1061 (2012).

[99] X. Xu and W. A. Goddard III, J. Phys. Chem. A **108**, 2305 (2004).

[100] S. A. Clough, Y. Beers, G. P. Klein, and L. S. Rothman, J. Chem. Phys. **59**, 2254 (1973).

[101] W. F. Murphy, J. Chem. Phys. **67**, 5877 (1977).

[102] A. Saitta, F. Saija, and P. Giaquinta, Phys. Rev. Lett. **108**, 207801 (2012).

[103] L. Delle Site, A. Alavi, and R. M. Lynden-Bell, Mol. Phys. **96**, 1683 (1999).

[104] K. Coutinho, R. Guedes, B. Costa Cabral, and S. Canuto, Chem. Phys. Lett. **369**, 345 (2003).

[105] A. V. Gubskaya and P. G. Kusalik, Mol. Phys. **99**, 1107 (2001).

[106] S. Niu, M.-L. Tan, and T. Ichiye, J. Chem. Phys. **134**, 134501 (2011).

[107] M. Salanne, R. Vuilleumier, P. A. Madden, C. Simon, P. Turq, and B. Guillot, J. Phys. Condens. Matter **20**, 494207 (2008).

[108] A. Buin and R. Iftimie, J. Chem. Phys. **131**, 234507 (2009).

[109] R. Ramírez, T. López-Ciudad, P. Kumar P, and D. Marx, J. Chem. Phys. **121**, 3973 (2004).

[110] M. Sharma, R. Resta, and R. Car, Phys. Rev. Lett. **100**, 019901 (2008).

[111] D. Porezag and M. Pederson, Phys. Rev. B **54**, 7830 (1996).

[112] G. E. Walrafen, M. R. Fisher, M. S. Hokmabadi, and W.-H. Yang, J. Chem. Phys. **85**, 6970 (1986).

[113] M. H. Brooker, G. Hancock, B. C. Rice, and J. Shapter, J. Raman Spectrosc. **20**, 683 (1989).

[114] D. E. Hare and C. M. Sorensen, J. Chem. Phys. **96**, 13 (1992).

[115] D. M. Carey and G. M. Korenowski, J. Chem. Phys. **108**, 2669 (1998).

[116] M. Galvin and D. Zerulla, Chemphyschem **12**, 913 (2011).

[117] T. Fukasawa, T. Sato, J. Watanabe, Y. Hama, W. Kunz, and R. Buchner, Phys. Rev. Lett. **95**, 1 (2005).

[118] Y. Amo and Y. Tominaga, Phys. A Stat. Mech. its Appl. **276**, 401 (2000).

[119] K. Winkler, J. Lindner, and P. Vöhringer, Phys. Chem. Chem. Phys. **4**, 2144 (2002).

[120] E. W. Castner, Y. J. Chang, Y. C. Chu, and G. E. Walrafen, J. Chem. Phys. **102**, 653 (1995).

[121] G. E. Walrafen, J. Phys. Chem. **94**, 2237 (1990).

[122] K. Mizoguchi, Y. Hori, and Y. Tominaga, J. Chem. Phys. **97**, 1961 (1992).

[123] G. E. Walrafen, M. S. Hokmabadi, W. H. Yang, Y. C. Chu, and B. Monosmith, J. Phys. Chem. **93**, 2909 (1989).

[124] A. Goncharov, N. Goldman, L. Fried, J. Crowhurst, I.-F. Kuo, C. Mundy, and J. Zaug, Phys. Rev. Lett. **94**, 1 (2005).

[125] J.-F. Lin, E. Gregoryanz, V. V. Struzhkin, M. Somayazulu, H.-k. Mao, and R. J. Hemley, Geophys. Res. Lett. **32**, L11306 (2005).

[126] T. Kawamoto, S. Ochiai, and H. Kagi, J. Chem. Phys. **120**, 5867 (2004).

[127] G. E. Walrafen, Y. C. Chu, and G. J. Piermarini, J. Phys. Chem. **100**, 10363 (1996).

[128] N. A. Chumaevskii and M. N. Rodnikova, J. Mol. Liq. **9697**, 31 (2002).

[129] M. Baumgartner and R. J. Bakker, Mineral. Petrol. **95**, 1 (2008).

[130] R. Li, Z. Jiang, Y. Guan, H. Yang, and B. Liu, J. Raman Spectrosc. **40**, 1200 (2009).

[131] I. A. Heisler and S. R. Meech, Science **327**, 857 (2010).

[132] L. Lupi, L. Comez, M. Paolantoni, D. Fioretto, and B. M. Ladanyi, J. Phys. Chem. B **116**, 7499 (2012).

[133] M. Paolantoni, N. F. Lago, M. Albertí, and a. Laganà, J. Phys. Chem. A **113**, 15100 (2009).

[134] F. G. Alabarse, J. Haines, O. Cambon, C. Levelut, D. Bourgogne, A. Haidoux, D. Granier, and B. Coasne, Phys. Rev. Lett. **109**, 035701 (2012).

[135] H. Suzuki, Y. Matsuzaki, A. Muraoka, and M. Tachikawa, J. Chem. Phys. **136**, 234508 (2012).

[136] P. A. Madden and R. W. Impey, Chem. Phys. Lett. **123**, 502 (1986).

[137] V. Mazzacurati, M. Ricci, G. Ruocco, and M. Sampoli, Chem. Phys. Lett. **159**, 383 (1989).

[138] B. D. Bursulaya and H. J. Kim, J. Chem. Phys. **109**, 4911 (1998).

[139] M. Skaf and M. Sonoda, Phys. Rev. Lett. **94**, 137802 (2005).

[140] M. T. Sonoda, S. M. Vechi, and M. S. Skaf, Phys. Chem. Chem. Phys. **7**, 1176 (2005).

[141] H. Torii, J. Phys. Chem. A **110**, 9469 (2006).

[142] J. J. Loparo, S. T. Roberts, R. A. Nicodemus, and A. Tokmakoff, Chem. Phys. **341**, 218 (2007).

[143] B. M. Auer and J. L. Skinner, J. Chem. Phys. **128**, 224511 (2008).

[144] M. Yang and J. L. Skinner, Phys. Chem. Chem. Phys. **12**, 982 (2010).

[145] J. A. Padro and J. Marti, J. Chem. Phys. **118**, 452 (2003).

[146] A. De Santis, A. Ercoli, and D. Rocca, J. Chem. Phys. **120**, 1657 (2004).

[147] J. A. Padro and J. Marti, J. Chem. Phys. **120**, 1659 (2004).

[148] P. L. Silvestrelli, M. Bernasconi, and M. Parrinello, Chem. Phys. Lett. **277**, 478 (1997).

[149] R. J. Heaton, P. A. Madden, S. J. Clark, and S. Jahn, J. Chem. Phys. **125**, 144104 (2006).

[150] J. J. Molina, S. Lectez, S. Tazi, M. Salanne, J.-F. Dufrêche, J. Roques, E. Simoni, P. A. Madden, and P. Turq, J. Chem. Phys. **134**, 014511 (2011).

[151] G. Bussi, D. Donadio, and M. Parrinello, J. Chem. Phys. **126**, 014101 (2007).

[152] O. F. Nielsen, Annu. Rep. Prog. Chem., Sect. C Phys. Chem. **93**, 57 (1996).

[153] T. Nakayama, Phys. Rev. Lett. **80**, 1244 (1998).

[154] H. Torii, Chem. Phys. Lett. **353**, 431 (2002).

[155] A. Aguado and P. A. Madden, J. Chem. Phys. **119**, 7471 (2003).

[156] T. Laino and J. Hutter, J. Chem. Phys. **129**, 074102 (2008).

[157] T. M. Nymand and P. Linse, J. Chem. Phys. **112**, 6152 (2000).

[158] J. Stenhammar, M. Trulsson, and P. Linse, J. Chem. Phys. **134**, 224104 (2011).

[159] D. Cappelletti, E. Ronca, L. Belpassi, F. Tarantelli, and F. Pirani, Acc. Chem. Res. **45**, 1571 (2012).

[160] T. K. Ghanty, V. N. Staroverov, P. R. Koren, and E. R. Davidson, J. Am. Chem. Soc. **122**, 1210 (2000).

[161] E. D. Glendening, J. Phys. Chem. A **109**, 11936 (2005).

[162] R. Z. Khaliullin, A. T. Bell, and M. Head-Gordon, Chem. Eur. J. **15**, 851 (2009).

[163] A. J. Stone and A. J. Misquitta, Chem. Phys. Lett. **473**, 201 (2009).

[164] R. Chelli, V. Schettino, and P. Procacci, J. Chem. Phys. **122**, 234107 (2005).

[165] H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma, J. Phys. Chem. **91**, 6269 (1987).

[166] S. Chatterjee, P. G. Debenedetti, F. H. Stillinger, and R. M. Lynden-Bell, J. Chem. Phys. **128**, 124511 (2008).

[167] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, J. Chem. Phys. **79**, 926 (1983).

[168] A. J. Lee and S. W. Rick, J. Chem. Phys. **134**, 184507 (2011).

[169] A. J. Lee and S. W. Rick, J. Phys. Chem. Lett. **3**, 3199 (2012).

[170] M. Soniat and S. W. Rick, J. Chem. Phys. **137**, 044511 (2012).

[171] B. Schropp and P. Tavan, J. Phys. Chem. B **112**, 6233 (2008).

[172] G. Lamoureux, A. D. MacKerell, and B. Roux, J. Chem. Phys. **119**, 5185 (2003).

[173] S. W. Rick, S. J. Stuart, and B. J. Berne, J. Chem. Phys. **101**, 6141 (1994).

[174] A. Perry, C. Neipert, B. Space, and P. B. Moore, Chem. Rev. **106**, 1234 (2006).

[175] J. J. Velasco-Velez, C. H. Wu, T. A. Pascal, L. F. Wan, J. Guo, D. Prendergast, and M. Salmeron, Science **346**, 831 (2014).

[176] D. Lis, E. H. G. Backus, J. Hunger, S. H. Parekh, and M. Bonn, Science **344**, 1138 (2014).

[177] M. Sovago, R. Campen, G. Wurpel, M. Müller, H. Bakker, and M. Bonn, Phys. Rev. Lett. **100**, 173901 (2008).

[178] M. Sovago, R. Campen, G. Wurpel, M. Müller, H. Bakker, and M. Bonn, Phys. Rev. Lett. **101**, 139402 (2008).

[179] T. Ishiyama, T. Imamura, and A. Morita, Chem. Rev. **114**, 8447 (2014).

[180] X. Wei, S.-C. Hong, A. I. Lvovsky, H. Held, and Y. R. Shen, J. Phys. Chem. B **104**, 3349 (2000).

[181] Y. R. Shen, J. Phys. Chem. C **116**, 15505 (2012).

[182] Y. R. Shen, J. Phys. Chem. C **117**, 11884 (2013).

[183] P. Guyot-Sionnest and Y. Shen, Phys. Rev. B **38**, 7985 (1988).

[184] Y. R. Shen, Appl. Phys. B **68**, 295 (1999).

[185] T. Kawaguchi, K. Shiratori, Y. Henmi, T. Ishiyama, and A. Morita, J. Phys. Chem. C **116**, 13169 (2012).

[186] S. J. Byrnes, P. L. Geissler, and Y. Shen, Chem. Phys. Lett. **516**, 115 (2011).

[187] V. Buch, H. Groenzin, I. Li, M. J. Shultz, and E. Tosatti, Proc. Natl. Acad. Sci. U. S. A. **105**, 5969 (2008).

[188] M. J. Shultz, P. Bisson, H. Groenzin, and I. Li, J. Chem. Phys. **133**, 054702 (2010).

[189] I. L. Barnett, H. Groenzin, and M. J. Shultz, J. Phys. Chem. A **115**, 6039 (2011).

[190] D. Pan, L.-M. Liu, G. Tribello, B. Slater, A. Michaelides, and E. Wang, Phys. Rev. Lett. **101**, 155703 (2008).

[191] R. Resta and S. Sorella, Phys. Rev. Lett. **82**, 370 (1999).

[192] D. Pan, L.-M. Liu, G. A. Tribello, B. Slater, A. Michaelides, and E. Wang, J. Phys. Condens. Matter **22**, 074209 (2010).

[193] B. Meyer and D. Vanderbilt, Phys. Rev. B **63**, 205426 (2001).

[194] P. Minary, M. E. Tuckerman, K. A. Pihakari, and G. J. Martyna, J. Chem. Phys. **116**, 5351 (2002).

[195] T. Ishiyama, H. Takahashi, and A. Morita, J. Phys. Chem. Lett. **3**, 3001 (2012).

[196] J. P. D. Abbatt, S. Benz, D. J. Cziczo, Z. Kanji, U. Lohmann, and O. Möhler, Science **313**, 1770 (2006).

[197] V. Ramanathan, P. Crutzen, J. Kiehl, and D. Rosenfeld, Science **294**, 2119 (2001).

[198] R. G. Harrison and K. S. Carslaw, Rev. Geophys. **41**, 1012 (2003).

[199] F. Yu and R. P. Turco, J. Geophys. Res. **106**, 4797 (2001).

[200] K. R. Asmis and D. M. Neumark, Acc. Chem. Res. **45**, 43 (2012).

[201] K. Tielrooij, N. Garcia-Araez, M. Bonn, and H. Bakker, Science **328**, 1006 (2010).

[202] X. Wang, X. Yang, J. Nicholas, and L. Wang, Science **294**, 1322 (2001).

[203] X. Yang, X. Wang, and L. Wang, J. Phys. Chem. A **106**, 7607 (2002).

[204] X.-B. Wang, A. P. Sergeeva, J. Yang, X.-P. Xing, A. I. Boldyrev, and L.-S. Wang, J. Phys. Chem. A **113**, 5567 (2009).

[205] M. Bush, R. Saykally, and E. Williams, J. Am. Chem. Soc. **129**, 2220 (2007).

[206] J. T. O'Brien, J. S. Prell, M. F. Bush, and E. R. Williams, J. Am. Chem. Soc. **132**, 8248 (2010).

[207] R. L. Wong and E. R. Williams, J. Phys. Chem. A **107**, 10976 (2003).

[208] A. T. Blades and P. Kebarle, J. Phys. Chem. A **109**, 8293 (2005).

[209] X.-B. Wang, J. B. Nicholas, and L.-S. Wang, J. Chem. Phys. **113**, 10837 (2000).

[210] C.-G. Zhan, F. Zheng, and D. A. Dixon, J. Chem. Phys. **119**, 781 (2003).

[211] P. Jungwirth and J. Curtis, Chem. Phys. Lett. **367**, 704 (2003).

[212] B. Gao and Z.-F. Liu, J. Chem. Phys. **121**, 8299 (2004).

[213] B. Gao and Z.-F. Liu, J. Chem. Phys. **123**, 224302 (2005).

[214] D. Lambrecht, G. Clark, T. Head-Gordon, and M. Head-Gordon, J. Phys. Chem. A **4**, 11438 (2011).

[215] A. Boldyrev and J. Simons, J. Phys. Chem. **98**, 2298 (1994).

[216] P. Jungwirth and D. J. Tobias, Chem. Rev. **106**, 1259 (2006).

[217] S. Gopalakrishnan, P. Jungwirth, D. J. Tobias, and H. C. Allen, J. Phys. Chem. B **109**, 8861 (2005).

[218] T. Ishiyama and A. Morita, J. Phys. Chem. C **115**, 13704 (2011).

[219] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, A. Dal Corso, S. de Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A. P. Seitsonen, A. Smogunov, P. Umari, and R. M. Wentzcovitch, J. Phys. Condens. Matter **21**, 395502 (2009).

[220] G. Makov and M. C. Payne, Phys. Rev. B **51**, 4014 (1995).

[221] J. Carrasco, A. Michaelides, M. Forster, S. Haq, R. Raval, and A. Hodgson, Nat. Mater. **8**, 427 (2009).

[222] S. A. Egorov, K. F. Everitt, and J. L. Skinner, J. Phys. Chem. A **103**, 9494 (1999).

[223] S. A. Egorov and J. L. Skinner, Chem. Phys. Lett. **293**, 469 (1998).

[224] M. Sipilä, T. Berndt, T. Petäjä, D. Brus, J. Vanhanen, F. Stratmann, J. Patokoski, R. L. Mauldin, A.-P. Hyvärinen, H. Lihavainen, and M. Kulmala, Science **327**, 1243 (2010).

[225] K. A. Mauritz and R. B. Moore, Chem. Rev. **104**, 4535 (2004).

[226] K.-D. Kreuer, S. J. Paddison, E. Spohr, and M. Schuster, Chem. Rev. **104**, 4637 (2004).

[227] D. Cavagnat and J. C. Lassegues, J. Phys. Condens. Matter **2**, SA189 (1990).

[228] D. Cavagnat and J. C. Lassegues, Solid State Ionics **46**, 11 (1991).

[229] H. Ohtaki and T. Radnai, Chem. Rev. **93**, 1157 (1993).

[230] Y. Kameda, K. Hosoya, S. Sakamoto, H. Suzuki, T. Usuki, and O. Uemura, J. Mol. Liq. **65-66**, 305 (1995).

[231] K. von Burg and P. Delahay, Chem. Phys. Lett. **78**, 287 (1981).

[232] P. Delahay, Acc. Chem. Res. **15**, 40 (1982).

[233] A. M. Margarella, K. A. Perrine, T. Lewis, M. Faubel, B. Winter, and J. C. Hemminger, J. Phys. Chem. C **117**, 8131 (2013).

[234] C. E. Lund Myhre, D. H. Christensen, F. M. Nicolaisen, and C. J. Nielsen, J. Phys. Chem. A **107**, 1979 (2003).

[235] D. A. Knopf, B. P. Luo, U. K. Krieger, and T. Koop, J. Phys. Chem. A **107**, 4322 (2003).

[236] A. D. Hammerich, V. Buch, and F. Mohamed, Chem. Phys. Lett. **460**, 423 (2008).

[237] Q. Wan, L. Spanu, and G. Galli, J. Phys. Chem. B **116**, 9460 (2012).

[238] M. Guidon, J. Hutter, and J. VandeVondele, J. Chem. Theory Comput. **6**, 2348 (2010).

[239] X. Wu, A. Selloni, and R. Car, Phys. Rev. B **79**, 085102 (2009).

[240] R. A. DiStasio, B. Santra, Z. Li, X. Wu, and R. Car, J. Chem. Phys. **141**, 084502 (2014).

[241] A. Musinu, G. Paschina, G. Piccaluga, and M. Magini, J. Appl. Crystallogr. **15**, 621 (1982).

[242] R. Mills, J. Phys. Chem. **77**, 685 (1973).

[243] X.-B. Wang, J. B. Nicholas, and L.-S. Wang, J. Phys. Chem. A **104**, 504 (2000).

[244] T. A. Pham, C. Zhang, E. Schwegler, and G. Galli, Phys. Rev. B **89**, 060202 (2014).

[245] B. Winter, R. Weber, I. V. Hertel, M. Faubel, P. Jungwirth, E. C. Brown, and S. E. Bradforth, J. Am. Chem. Soc. **127**, 7203 (2005).

[246] R. Nunes and X. Gonze, Phys. Rev. B **63**, 155107 (2001).

[247] C. Tian, N. Ji, G. Waychunas, and Y. R. Shen, J. Am. Chem. Soc. **130**, 13033 (2008).