Predictive Theory and Modelling of Heterogeneous Interfaces

By

Tuan Anh Pham
B.S. (Hanoi National University of Education) 2006

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

CHEMISTRY

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved by:

Professor Giulia Galli (Chair)

Professor C. William McCurdy

Professor Gergely Zimanyi

Committee in Charge
2014

For my mother, my dad and for my wife,
for all their support and encouragement along the way.
And for my awaited son for the great motivation.

Contents

Tuan Anh Pham
September 2014
Chemistry

Predictive Theory and Modelling of Heterogeneous Interfaces

Abstract

While much progress has been made in the microscopic characterization of bulk materials, understanding the behavior, at the atomic scale, of complex materials interfaces remains a challenging problem in condensed matter physics and materials science. Development of predictive theory and modeling for interfacial systems is not only important to aid experimental interpretations but is also crucial to accelerate materials design efforts for a broad range of technological applications from the semiconductor industry to the development of new sources of energy.

This dissertation presents a computational framework that combines first-principles molecular dynamics simulations for the determination of interfacial atomic structures, and advanced electronic structure methods for the description of electronic properties to understand, predict and design materials interfaces. In particular, the research presented in this thesis was carried out in two parallel directions:

- Predictions of the interfacial atomic structure of complex materials interfaces using first-principles molecular dynamics, and validation of the predictions by relating structural models to experiments, e.g., data from X-ray, infrared spectra and sum-frequency generation spectroscopy experiments.

  In particular, I investigated structural and dynamical properties of the ice Ih surface and of the Al_2O_3/water interface. While the study of the ice surface is the first step towards the understanding of complex semiconductor/water interfaces, the Al_2O_3/water interface represents a suitable prototype interface for extensive comparisons between theory and experiments.

- Development of advanced first-principles techniques to study electronic states at interfaces and application of these techniques to gain insights on the relationship between between local interface structure and electronic properties.

  I devised a new technique to study excited states and photoemission spectra based on many-body perturbation theory, within the so-called GW approximation, which improves both the computational efficiency and accuracy of existing methodologies, and that can be employed to study realistic systems.

  In addition, in the thesis I employed the new GW technique to investigate a variety of systems including molecules, nanostructures, semiconducting interfaces, liquid water and simple aqueous solutions. These studies are crucial to build a fundamental understanding of the electronic structure of semiconductor/liquid water interfaces and of, e.g., water with dissolved ions under different pH conditions, interfaced with a photoelectrode.

Information provided by these two parallel research directions helped establish a structure-electronic properties-chemical reactivity paradigm, that is general and applicable to a large class of materials. An example presented in the thesis is that of functionalized Si surfaces interfaced with liquid water, in which I studied the effect of surface functionalization on the alignment between Si band edges and water redox potentials, and I suggested a possible approach to engineer and design semiconductor surfaces for photoelectrochemical water splitting.

Acknowledgments and Thanks

I would like to express my deepest gratitude to a number of people without whom this thesis would not be completed.

First and foremost, I would like to thank the support and direction I received from my advisor, Prof. Giulia Galli. During the last five years of my PhD, Giulia has provided me with an excellent research environment, offering me the opportunity to work with outstanding researchers, and more importantly, the opportunity to explore ideas independently. Giulia's incredible passion for scientific research, her vision and optimism have had an enormous impact on my study as well as on my future career.

I would like to thank Dr. Eric Schwegler, my mentor at the Lawrence Livermore National Laboratory, for his priceless guidance, encouragement and constant support. Thanks to Eric, I have gradually developed my professonalism as well as improved my self-development.

I would like to thank my collaborators: Prof. Francois Gygi, Prof. Tianshu Li, Dr. Huy Viet Nguyen, Dr. Dario Rocca, Dr. Cui Zhang, Dr. Deyu Lu, Dr. Patrick Huang, Dr. Donghwa Lee, Dr. Daniel Opalka, Prof. Michiel Sprik and Dr. Sadasivan Shankar for the fruitful work and delightful collaboration we had together. Without their support, this thesis would not be possible.

I would like to thank all of the other Angstrom group members at the University of California Davis, and Quantum Simulations group members at the Lawrence Livermore

National Laboratory for their tremendous support and encouragement. I I have learned so much from each of them, and I appreciate the pleasant times we spent together.

Last but not least, my wife Linh Trinh holds all my gratitude for her patience, support and for her incredible faith in my talent. Her love has enriched everyday in my life with happiness and passion. This thesis is dedicated to her, my awaited son and my beloved parents.

List of Figures

List of Tables

Introduction

Global demand for energy is projected to more than double by 2050 and to more than triple by the end of the century. Finding sufficient supplies of clean, efficient and renewable energy sources is thus one of the most important scientific and technical challenges facing humanity in the 21st century. Renewable sources, such as solar and wind, offers enormous potential for meeting this future energy demand. For example, more energy from sunlight strikes the earth in 1 hour (4.3×10^{20} J) than all of the energy currently consumed on the planet in 1 year (4.1×10^{20} J in 2001) [1, 2].

The design of highly efficient "machines", such as photoelectrochemical (PEC) cells and chemical energy storage devices (batteries) to harness and store solar energy, is both a formidable challenge and an opportunity that, if realized, could have a revolutionary impact on our energy system. These are two representative devices, for which the understanding of structure and reactivity at interfaces between, e.g., solids (electrodes) and liquids (electrolytes) is the key to designing novel materials for the next generation of devices.

Heterogeneous interfaces are often difficult to access and characterize by experimental probes as they are typically hidden or buried within bulk materials. In complex systems, modeling and simulation have played an important role by aiding in the interpretation of the experiments and in some cases by providing information that cannot be directly accessed by experiment. In recent years, due to important advances in both high perfor-

mance computing resources and sophisticated electronic structure codes, our ability to use modeling and simulation as predictive tools has been rapidly increasing and it is now conceivable to use these approaches to guide and accelerate the discovery of new materials for a wide range of energy technologies.

As shown recently, e.g., Refs. [3, 4], it is now possible to use first-principles calculations to scan a large number of combinations of elements across the entire periodic table, and to suggest many new candidates as, e.g., promising photo-electrode materials for photo-electrochemical solar cells and anode/cathode materials for batteries. However, current computational screening schemes presented in the recent literature have mostly focused on the bulk properties of solids, without the explicit treatment of the microscopic structure and chemistry of heterogeneous interfaces. A focus on bulk materials properties is a reasonably simple starting point, as it usually only involves the calculation of relatively small system sizes that can be efficiently carried out with modest computational resources. Such simplicity, however, does not account for the complex physical and chemical processes that occur at surfaces and interfaces, such as the separation and recombination of charge carriers at PEC electrode/electrolyte interfaces. Without a detailed information of the microscopic structure and chemistry occurring at heterogeneous interfaces, the formulations of a consistent strategy to optimize materials for the next generation of, e.g., PEC cells and batteries, is an extremely difficult task, if not impossible.

A desirable theoretical approach for optimizing materials for, e.g., the next generation of photoelectrochemical cells and chemical energy storage devices should allow one to accurately predict the behavior and properties of heterogeneous interfaces. The development of such approach is very challenging as it requires:

1. The generation and validation of realistic structural models of complex materials with *ab-initio* simulations.

2. The development of advanced electronic structure methodologies which are capable

of providing accurate descriptions of the electronic properties of interfacial systems with possibly hundreds of atoms. In particular, calculations using advanced electronic structure methods such as many-body perturbation theory, e.g., within the *GW* approximation are rather demanding from a computational standpoint, and are usually limited to systems containing only a few tens of atoms with conventional implementations.

3. The applicability of *ab-initio* calculations to complex materials interfaces involving different phases, and to broad classes of materials with various compositions and morphology.

This dissertation focuses on devising new theoretical and computational techniques to meet these challenges. In this dissertation, I present a computational framework that couples first-principles molecular dynamics simulations for the determination of interfacial atomic structures, together with an *ab-initio* description of electronic properties based on many-body perturbation theory within the *GW* approximation, to establish simple relationships between local interface structure, electronic properties of semiconductors and their reactivity in aqueous solutions. The investigations carried out in this dissertation have proceeded in two parallel directions, with the following goals:

1. Predicting the interfacial atomic structure of solids and liquids, using first-principles molecular dynamics, and validating our predictions by relating the structural models to experiment, e.g., data from X-ray, infrared, Raman and sum-frequency generation spectroscopy experiments.

2. Devising advanced first-principles electronic structure techniques that improve the efficiency and accuracy of *GW* calculations and allow for the application of the *GW* approach to study realistic interfacial systems. Applying the new implementation to study electronic states at interfaces to derive insights into the relationship between

interfacial structure and reactivity.

The results of these two parallel research directions allowed me to establish a structure-electronic properties-chemical reactivity paradigm, that is general and applicable to a large class of materials. The rest of the thesis is organized as follows:

- Chapter 1 presents the theoretical background on which this thesis is built, with the emphasis on density functional theory and first-principles molecular dynamics for the description of ground state properties, and on many-body perturbation theory within the so-called *GW* approximation for the description of excited state properties of materials.

- Chapter 2 presents the development of a new technique to investigate excited states and photoemission spectra based on many-body perturbation theory within the *GW* approximation, which improves both the computational efficiency and accuracy of existing methodologies, and that can be employed to study realistic systems, directly comparable to experiments. The chapter also presents applications of the technique to study the electronic properties of a variety of systems including molecules, nanostructures, insulators and semiconductors.

- Chapter 3 presents applications of density functional theory and first-principles molecular dynamics simulations to study structural and dynamical properties of oxide surfaces and interfaces. In particular, two systems were considered, including the ice Ih surface and the Al_2O_3/water interface. While the study of the ice surface is the first step towards the understanding of complex semiconductor/water interfaces, the Al_2O_3/water interface represents a suitable prototype interface for extensive comparisons between theory and experiments.

- Chapter 4 combines density functional theory, first-principles molecular dynamics and the *GW* approach to investigate nitride surfaces and interfaces. In particular,

first-principles molecular dynamics was employed to obtain a realistic model of the Si_3N_4/Si(001) interface, density functional theory was used to investigate the dielectric properties of the nitride interface and surfaces, and the *GW* approach was utilized to compute the band offsets at the Si_3N_4/Si(001) interface.

- Chapter 5 is devoted to the electronic properties of simple aqueous solutions, i.e., liquid water and a chloride anion solution, using the newly developed *GW* approach presented in Chapter 2. These studies are crucial to build a fundamental understanding of the electronic structure of semiconductor/liquid water interfaces and of, e.g., water with dissolved ions under different pH conditions, interfaced with a photoelectrode, as employed in water splitting devices.

- Chapter 6 presents a new computational framework that combines first-principles molecular dynamics and the *GW* approach to investigate the electronic structure of semiconductor/water interfaces, with focus on the calculation of the energy alignment between photoelectrode band edges and water redox potentials. The computational framework is then applied to functionalized Si surfaces interfaced with water.

Finally the work is summarized in the Conclusion chapter, which also includes perspective on future work.

Chapter 1

Theoretical background

1.1 Schrödinger equation and Born-Oppenheimer approximation

Quantum mechanical simulation of electrons and ions is one of the most powerful tools for explaining and predicting a vast range of phenomena in physics, chemistry and materials science. In quantum theory, electrons and ions are described by wavefunctions and therefore, from a theoretical point of view, knowing the wavefunctions is enough to describe all the properties of a system. Although the wavefunction does not represent any physical quantity, its square modulus is interpreted as the probability density of the distribution of particles, which plays the key role in understanding the behaviors of a system.

In the most general form, the wavefunction is the solution of the Schrödinger equation:

$$\hat{H}\Psi = W\Psi, \tag{1.1.1}$$

where Ψ and W are the wavefunction and total energy of the system. The non-relativistic

Hamiltonian $\hat{H}$ can be written explicitly for a system of N electrons and M nuclei:

$$\begin{aligned}\hat{H} &= \hat{T}_e + \hat{U}_{e-e} + \hat{W}_{e-I} + \hat{T}_I + \hat{W}_{I-I}, \\ &= \sum_{i=1,N} (\frac{1}{2}\nabla_i^2 + v_c(\mathbf{r}_i)) + \frac{1}{2}\sum_{i\neq j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \hat{T}_I + \hat{W}_{I-I},\end{aligned} \tag{1.1.2}$$

where

$$v_c(\mathbf{r}_i) = -\sum_{I=1,M} \frac{Z_I}{|\mathbf{r}_i - \mathbf{R}_I|}; \hat{W}_{I-I} = \frac{1}{2}\sum_{I\neq J} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|}. \tag{1.1.3}$$

The operators in Eq. 1.1.2 are the kinetic energy $\hat{T}_e$ of electrons, the interaction energy $\hat{U}_{e-e}$ between electron-electron, the interaction energy $\hat{W}_{e-I}$ between electrons and nuclei of charge Z_I, the kinetic energy $\hat{T}_I$ of nuclei, and the Coulomb interaction energy $\hat{W}_{I-I}$ between nuclei. Atomic units ($\hbar = m = e = 1$) are used throughout the thesis.

Solving Eq. 1.1.1 yields the many-body wavefunction $\Psi(\mathbf{r}_1, ..., \mathbf{r}_N; \mathbf{R}_1, ..., \mathbf{R}_M)$ for a system of N-electron-M-nucleus. However, the Hamiltonian in Eq. 1.1.2 is too complicated for a direct handling, and therefore, approximations are needed for practical descriptions. To partially overcome this problem, the usual choice is to resort to the so-called adiabatic approximation, introduced by Born and Oppenheimer in 1927. The general idea of this method is the factorization of the full wavefunction $\Psi(\{\mathbf{r}_i\}, \{\mathbf{R}_I\})$ into the nuclear and electronic ones, based on the fact that the mass of electrons and those of ions are sufficiently different, i.e., $m_{nuclei}/m_e \geq 1836$. In this spirit, electrons can be described as following instantaneously the motion of nuclei, while their dynamics is characterized by the dynamically screened interaction potential. The total wavefunction $\Psi(\{\mathbf{r}_i\}, \{\mathbf{R}_I)\}$ can be separated into the product of a nuclei part, $\Phi(\{\mathbf{R}_I\})$, and an electronic part, $\psi(\{\mathbf{r}_i\}; \{\mathbf{R}_I)\}$, as:

$$\Psi(\{\mathbf{r}_i\}, \{\mathbf{R}_I\}) = \Phi(\{\mathbf{R}_I\})\psi(\{\mathbf{r}_i\}; \{\mathbf{R}_I\}), \tag{1.1.4}$$

and the Hamiltonian can be rewritten as two coupled problems for electrons and nuclei, respectively:

$$(-\frac{1}{2}\sum_{i=1,N}\nabla_i^2 - \sum_{i=1,N}\sum_{I=1,M}\frac{Z_I}{|\mathbf{r}_i-\mathbf{R}_I|} + \frac{1}{2}\sum_{i\neq j}\frac{1}{|\mathbf{r}_i-\mathbf{r}_j|})\psi(\mathbf{r}_i;\mathbf{R}_I) = E_{el}(\{\mathbf{R}_I\})\psi(\{\mathbf{r}_i;\mathbf{R}_I\}), \tag{1.1.5}$$

and

$$(-\frac{1}{2}\sum_{I=1,M}\frac{\nabla_I^2}{M_I} + E_{el}(\{\mathbf{R}_I\}))\Phi(\{\mathbf{R}_I\}) = \varepsilon\Phi(\{\mathbf{R}_I\}). \tag{1.1.6}$$

Although the Born-Oppenheimer approximation allows one to work with the much simpler problem of considering motions of electrons and nuclei separately, the problem of the electrons only itself is still a many-body problem. Solving the equation Eq. 1.1.5 for electrons is the main task of computational electronic structure theory.

1.2 Density functional theory

In practice, solving the equation Eq. 1.1.5 to find the electronic ground state wavefunction for a many-electron system is still an unsolvable mathematical problem because of its dependence on $3N_e$ degrees of freedom. Therefore, suitable approximations are further needed.

In 1964, Hohenberg and Kohn proposed and proved a theorem that allows to search for the ground state energy of a many-body electron system as a functional of its ground-state electron density, introducing the so-called Density Functional Theory (DFT) [5]. Employing the DFT approach to the electronic problem is, in principle, simpler than a direct handling of the equation Eq. 1.1.5 since it reduces the number of degrees of freedom from $3N_e$ to only 3 coordinate variables. The theoretical framework of DFT is based on two theorems, first proved by Hohenberg and Kohn:

1. For any system of interacting particles in an external potential $V_{ext}(\mathbf{r})$, the potential

$V_{ext}(\mathbf{r})$ is determined uniquely, except for a constant, by the ground state electronic density $n_0(\mathbf{r})$.

2. An universal functional of the electronic density $F[n]$ can be defined, such that the ground state energy, as a functional of the external potential, $E_0[V_{ext}]$, can be derived from the variational principle:

$$E_0[V_{ext}] = min_n(F[n] + \int d\mathbf{r} V_{ext}(\mathbf{r})n(\mathbf{r})) + W_{I-I}, \tag{1.2.1}$$

with the constraint that electron density is normalized to the total number of electrons:

$$\int d\mathbf{r} n(\mathbf{r}) = N. \tag{1.2.2}$$

The density that minimizes the Eq. 1.2.1 is the ground state density $n_0(\mathbf{r})$.

The Hohenberg-Kohn (HK) theorems demonstrate the existence of such a functional, $F[n]$, connecting the ground state of a many-body quantum system to its ground-state particle density, whose explicit form is however unknown. A practical recipe for computing this functional is provided by the Kohn-Sham (KS) ansatz [6], which maps the many-body system onto an auxiliary system of non-interacting particles. The ground state density of N non-interacting electrons is defined as:

$$n(\mathbf{r}) = \sum_v 2|\psi_v(\mathbf{r})|^2, \tag{1.2.3}$$

where the summation runs over the occupied states, and $\psi_v(\mathbf{r})$ are single particle wavefunctions.

The kinetic energy of the system is:

$$T_0 = -\frac{1}{2}\sum_v \langle\psi_v|\nabla^2|\psi_v\rangle. \tag{1.2.4}$$

The Hartree energy, defined as the classical electrostatic energy of a system with electron density $n(\mathbf{r})$ is:

$$E_H[n] = \frac{1}{2}\int d\mathbf{r}d\mathbf{r}'\frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}. \tag{1.2.5}$$

The KS ansatz splits the HK functionals into the sum of T_0, E_H and the unknown term called the exchange-correlation energy:

$$F[n] = T_0[n] + E_H[n] + E_{xc}[n]. \tag{1.2.6}$$

Applying the variational equation of Theorem II, we obtain the Euler equation:

$$\frac{\delta T_0[n]}{\delta n(\mathbf{r})} + V_{KS}(\mathbf{r}) = \mu \tag{1.2.7}$$

where μ is the Lagrange multiplier accounting for the constraint on the total number of electrons, and V_{KS} is defined as:

$$V_{KS}(\mathbf{r}) = \frac{\delta E_H[n]}{\delta n(\mathbf{r})} + \frac{\delta E_{xc}[n]}{\delta n(\mathbf{r})} + V_{ext}(\mathbf{r}) \tag{1.2.8}$$

$$= V_H(\mathbf{r}) + V_{xc}(\mathbf{r}) + V_{ext}(\mathbf{r}) \tag{1.2.9}$$

Equation 1.2.7 is the same to that used to describe a system of non-electrons, subject to an effective potential $V_{KS}(\mathbf{r})$. Therefore we can define the Hamiltonian of an auxiliary non-interacting system as:

$$H_{KS} = \frac{1}{2}\nabla^2 + V_{KS}(\mathbf{r}). \tag{1.2.10}$$

The corresponding Schrödinger equation of this non-interacting system is:

$$H_{KS}\psi_i(\mathbf{r}) = \epsilon_i\psi_i(\mathbf{r}), \tag{1.2.11}$$

which is so called the KS equation. Eigenvalues and eigenfunctions are KS orbital energies and KS orbitals, respectively.

Given the exact form of the exchange-correlation functional, DFT would provide exact ground state properties of an many-body interacting system. In practice, the exact form of $E_{xc}[n]$ is unknown, and different approximations are employed to account for this term. The most commonly used approximations are the local and semi-local approximations which assume that the exchange-correlation energy depends only locally and semi-locally on the electron density. The Local Density Approximation (LDA) for the exchange-correlation energy is written as:

$$E_{xc}^{LDA}[n(\mathbf{r})] = \int \epsilon_{xc}^{LDA}(n(\mathbf{r}))n(\mathbf{r})d\mathbf{r} \tag{1.2.12}$$

where ϵ_{xc}^{LDA} is the exchange-correlation energy per electron of the homogeneous electron gas with a density n. The Generalized Gradient Approximation (GGA), in constrast, take into account the dependence of the exchange-correlation energy on the local density gradient [6]:

$$E_{xc}^{GGA}[n(\mathbf{r})] = \int \epsilon_{xc}(n(\mathbf{r}), \nabla n(\mathbf{r}))n(\mathbf{r})d\mathbf{r}. \tag{1.2.13}$$

One of the most popular GGA functional is the Perdew, Burke, and Enzerhof (PBE) [7].

1.3 First-principles molecular dynamics

Numerical simulations based on first-principles molecular dynamics (MD) have proven to be a powerful tool to investigate structural, dynamical and electronic properties of condensed systems in a wide range of temperatures. MD simulations generate the ionic trajectories which allow to calculate both equilibrium and nonequilibrium statistical averages, and access to microscopic details of atomic dynamics, which may not be available

from experiments.

Let us consider N atoms with coordinates $\{\mathbf{R}_{I,N}\}$, the main ingredient of MD is the calculation of the atomic force:

$$\mathbf{F}_I = -\frac{\partial E_0[\{\mathbf{R}_I\}]}{\partial \mathbf{R}_I}, \tag{1.3.1}$$

where the ground state energy $E_0[\{\mathbf{R}_i\}]$ can be obtained from DFT with an approximation, e.g. LDA, for the exchange-correlation energy. Because the mass of electrons is much smaller than that of nuclei, the time evolution of electrons is much more rapid than nuclear motion. In the Born-Oppenheimer (BO) molecular dynamics, these two freedoms of motion are considered separately. For any molecular configuration occurring at each step, the ground-state density is first calculated by solving KS equations, then the ionic coordinates are updated by integrating the equations of motion:

$$M_I \ddot{\mathbf{R}}_I = \mathbf{F}_I, \tag{1.3.2}$$

for a small time step. Although this approach is straightforward and usually stable, a drawback is that the electronic configuration has to be converged for every single ionic step. The so-called Car-Parrinello (CP) algorithm avoids this problem by letting both ionic and electronic degrees of freedom evolve under the action of a fictious Lagrangian:

$$L = \sum_i 2\mu \int d\mathbf{r} |\dot{\psi}_i(\mathbf{r})|^2 + \frac{1}{2}\sum_I M_I |\dot{\mathbf{R}}_I|^2 - E[\{\psi_i\}, \{\mathbf{R}_I\}] + \sum_{ij} \Lambda_{ij}[\int d\mathbf{r}\psi_i(\mathbf{r})\psi_j(\mathbf{r}) - \delta ij], \tag{1.3.3}$$

where μ is an arbitrary parameter of appropriate units which serves to define the fictious kinetic energy of the KS orbitals, and Λ_{ij} is a Lagrangian multiplier matrix to comply with the orthonormality constrain. This Lagrangian generates a dynamics in the coupled

electron-ion parameter space through the equation of motion:

$$\mu\ddot{\mu}_i(\mathbf{r},t) = -\hat{H}_{KS}\psi_i(\mathbf{r},t) + \sum_j \Lambda_{ij}\psi_i(\mathbf{r},t)$$
$$M_I\ddot{\mathbf{R}}_I = \mathbf{F}_I = -\frac{\partial E_0[\{\mathbf{R}_I\}]}{\partial \mathbf{R}_I}. \tag{1.3.4}$$

In the presence of a large gap in the electronic spectrum, a wise choice of the fictitious mass makes it possible to limit the energy transfer from the nuclear to the electronic degrees of freedom, thus reducing the need of reoptimizing the electronic wavefunctions at each ionic step. In the formal limit where $\mu \to 0$, the Car-Parrinello equations of motion approach those of Born-Oppenheimer molecular dynamics.

1.4 Many-body perturbation theory

Although DFT has been proved very successful for describing the ground state properties, e.g. structural and vibrational properties, of a wide range of materials, it is usually not a method of choice to study electronic excitations. The solution of the KS equations (Eq. 1.2.11) yields single particle energies that are often identified as excitation energies. This has no theoretical foundation as KS eigenvalues must be considered as mathematical tools and should not be interpreted as electron removal and addition energies. The only exception is the DFT highest occupied molecular orbital energy, which equals the exact *minus* of the ionization potential given an exact-exchange correlation functional. As a consequence, while often qualitatively correct, the DFT calculations fails to provide reliable quantitative values for band gaps and band structures of semiconductors and insulators.

Band gaps and band structures are experimentally measured by photoelectron spectroscopy. In the direct photoelectron spectroscopy, a photon impinges on the sample

and ejects an electron, whose the binding energy ϵ_i is measured and equals the difference $\epsilon_i = E_0^N - E_i^{N-1}$ between the total energy E_0^N of the N-particle ground state and the energy E_i^{N-1} of the $N-1$-particle excited state. In the inverse photoelectron spectroscopy, an electron is injected into the sample, and the energy of the photon is measured which equals the difference $\epsilon_i = E_i^{N+1} - E_0^N$ between the total energy E_i^{N+1} of the $N+1$-particle excited state and the energy E_0^N of the N-particle ground state.

A state of the art method to describe processes involving the ejection and injection of electrons in photoelectron experiments is based on many-body perturbation theory (MBPT). The central variable in this approach is the single particle Green's function, whose the basic concept is explained in the following section.

1.4.1 Green's function

Single particle Green's functions are useful concepts to compute electronic excitations involving the additional or the removal of electrons, and thus to describe photoemission and inverse photoemission experiments. The time-ordered single particle Green's function is defined as:

$$G^{(1)}(1,2) = -i\langle \Psi_0^N | \hat{T}\hat{\psi}(1)\hat{\psi}^+(2) | \Psi_0^N \rangle, \tag{1.4.1}$$

where $1 \equiv (\mathbf{r}_1, t_1)$ and $2 \equiv (\mathbf{r}_2, t_2)$ are compact notations for space and time cordinates; Ψ_0^N is the normalized many body ground state wavefunction of a system containing N electrons, $\hat{T}$ is the time ordering operator, and $\hat{\psi}^+(\mathbf{r}, t)$ ($\hat{\psi}(\mathbf{r}, t)$) is an operator in the Heisenberg representation that creates (destroys) an electron at position $\mathbf{r}$ at time t. If $t_1 > t_2$ the Green's function $G^{(1)}$ gives the probability amplitude that an electron added to the system at time t_2 at position $\mathbf{r}_2$ be detected at position $\mathbf{r}_1$ at time t_1. If $t_2 > t_1$ an electron is removed from the system, and $G^{(1)}$ describes the time evolution of the corresponding hole.

Considering the time-ordered Green's function of a stationary system with $\tau = t_1 - t_2$,

by inserting a complete set of eigenstates of the system $\{\Psi_j^{N\pm1}\}$ with $(N \pm 1)$ particles, and applying a Fourier transform one obtains the so-called Lehmann representation of the Green's function:

$$G^{(1)}(\mathbf{r}_1, \mathbf{r}_2, \omega) = \sum_j \frac{\psi_j^{N+1}(\mathbf{r}_1)\psi_j^{N+1^*}(\mathbf{r}_2)}{\omega - \epsilon_j^{N+1} + i\eta} + \sum_j \frac{\psi_j^{N-1}(\mathbf{r}_1)\psi_j^{N-1^*}(\mathbf{r}_2)}{\omega - \epsilon_j^{N-1} - i\eta}, \tag{1.4.2}$$

where η is an infinitesimal positive number, and:

$$\begin{aligned} \psi_j^{N-1}(\mathbf{r}) &= \langle \Psi_j^{N-1} | \hat{\psi}(\mathbf{r}) | \Psi_0^N \rangle, \\ \psi_j^{N+1}(\mathbf{r}) &= \langle \Psi_0^N | \hat{\psi}(\mathbf{r}) | \Psi_j^{N+1} \rangle. \end{aligned} \tag{1.4.3}$$

together with the excitation energies:

$$\begin{aligned} \epsilon_j^{N-1} &= E_0^N - E_j^{N-1}, \\ \epsilon_j^{N+1} &= E_i^{N+1} - E_0^N. \end{aligned} \tag{1.4.4}$$

Within the Lehmann representation, it is seen that the poles of the time-ordered single-particle Greens function represent the energies necessary to add or remove an electron from a N-electron system. The physical interpretation of the Green's function indicates that it can be used to describe photoemission or inverse photoemission experiments.

1.4.2 Dyson's equation and the GW approximation

Calculations of the single particle Green's function for any interacting electrons system is a difficult task. The Feynman-Dyson perturbation theory can be used to to approximate the many-body Green's function by considering the Coulomb interaction between electrons as a perturbation acting on the non-interacting system. Within this framework the single particle Green's function can be defined in terms of the non-interacting single particle

Green's function $G_0^{(1)}$ through the Dyson's equation:

$$G^{(1)}(1,2) = G_0^{(1)}(1,2) + \int G_0^{(1)}(1,3)\Sigma(3,4)G^{(1)}(4,2)d(34), \tag{1.4.5}$$

where Σ is a nonlocal, frequency-dependent and non-Hermitian self-energy operator, which contains all many-body exchange and correlation effects beyond the electrostatic Hartree potential. If an accurate approximation of the self-energy Σ was known, the Dyson's equation would provide a way to compute the interacting Green's function $G^{(1)}$ from $G_0^{(1)}$, the Green function of a mean-field system defined by wavefunctions φ_j and energies ϵ_j^{ind} obtained from the independent electron equation:

$$\hat{H}_0\varphi_j^0 = [-\frac{1}{2}\Delta + \hat{V}_{e-I} + \hat{V}_H]\varphi_j^0 = \epsilon_j^0\varphi_j^0, \tag{1.4.6}$$

where $\hat{V}_{e-I}$ is the Coulomb potential due to nuclei, $\hat{V}_H$ is the Hartree potential. By inserting the Lehmann representation (Eq. 1.4.2) into Eq. 1.4.5, we find that the wavefunctions ψ_j and eigenvalues ϵ_j obey the quasiparticle equation:

$$\hat{H}_0(\mathbf{r}_1)\psi_j(\mathbf{r}_1) + \int \Sigma(\mathbf{r}_1, \mathbf{r}_2, \epsilon_j)\psi_j(\mathbf{r}_2)d\mathbf{r}_2 = \epsilon_j\psi_j(\mathbf{r}_1). \tag{1.4.7}$$

Although Eq. 1.4.7 is similar to single particle equations of mean-field approaches such as DFT, it does not constitute a mean-field formulation as the self-energy takes all dynamic many-electron processes into account.

In practice, the self-energy must be approximated in order to solve the quasiparticle equation. One of the most common used approximation for Σ is so-called the GW approximation [8] which reads:

$$\Sigma(\mathbf{r}_1, \mathbf{r}_2, \omega) = \frac{i}{2\pi}\int G_0^{(1)}(\mathbf{r}_1, \mathbf{r}_2, \omega + \omega')W(\mathbf{r}_1, \mathbf{r}_2, \omega')e^{i\omega'\eta}d\omega', \tag{1.4.8}$$

where $G_0^{(1)}$ is the Green's function of the non-interacting system defined by Eq. 1.4.6, and W is the dynamically screened interaction, which is related to the bare Coulomb potential through the inverse of the dielectric function:

$$W(\mathbf{r}_1, \mathbf{r}_2, \omega) = \int \epsilon^{-1}(\mathbf{r}_1, \mathbf{r}_3, \omega) v_c(\mathbf{r}_3, \mathbf{r}_2) d\mathbf{r}_3 \tag{1.4.9}$$

It has been shown that using the GW approximation, the quasiparticle energies ϵ_j obtained from the quasiparticle equation can represent very well electron removal and addition energies and thus photoelectron spectroscopy experiments. Details of practical implementations of the GW approach will be discussed in chapter 2.

Chapter 2

Improving the calculation of photoelectron spectra within many-body perturbation theory

2.1 Introduction

Calculations based on DFT are widely used in condensed matter physics and chemistry to study the structural and electronic properties of molecules, nanostructures and materials. However, in principle DFT calculations with approximate exchange correlation functionals are limited to ground state properties; often, when used to describe electronic excitation processes, these calculations do not yield good, quantitative agreement with experiments, although they may account for trends in specific properties within given classes of materials or molecules [9].

A formal basis for studying electronic excitation processes in materials is provided by MBPT [10, 11]. In particular, in the mid-60s an approach was introduced by Hedin, to describe the direct and inverse photoemission spectra of solids and to compute ionization

potentials and electron affinities of molecules, the so called GW approach [8]. It then took two decades to develop efficient techniques to apply such a method to solids, e.g., crystalline semiconductors. Since the pioneering work of Hybertsen and Louie [12], the GW approximation to quasiparticle energies has been widely used to study fundamental band gaps [13], band offsets at interfaces [14], and quantum transport in molecular contacts [15]. In addition, quasiparticle energies obtained within the GW approximation may be used as input for the solution of the Bethe-Salpeter equation, an alternative method to time-dependent density functional theory for the calculation of absorption spectra [11].

Despite its success over the past two and a half decades, the use of the GW approximation to study materials and molecules still faces several numerical challenges. One of the difficulties in computing quasiparticle energies is the evaluation of the dielectric matrix (ϵ) used to describe the electronic polarization and screening of a system subjected to an external electromagnetic perturbation. Several approximations have been developed to reduce the computational workload required to evaluate ϵ, including the use of plasmon-pole models to approximate the frequency dependence of ϵ, and the use of modest plane-wave cutoffs to represent the dielectric matrix with a plane-wave basis set [12]. Furthermore, the direct evaluation of both the Green's function (G) and the dielectric matrix entering the expression of the screened Coulomb interaction (W) requires summations over an infinite number of unoccupied electronic states. In practice, such summations are truncated at a finite number of unoccupied states, and the convergence of the computed quasiparticle energies as a function of this number needs to be carefully checked for each specific system [16, 17, 18, 19].

In this chapter we present a technique for the evaluation of quasiparticle energies aimed at improving both the computational efficiency and the control of numerical errors of existing methodologies. Our approach does not require the calculation of unoccupied electronic states, nor the direct diagonalization of large dielectric matrices, and it avoids

the use of plasmon pole models. Most importantly, its numerical accuracy is controlled by a single parameter, i.e. the number of eigenvalues and eigenvectors used in the spectral decomposition of the static dielectric matrix. In the last few years, several other techniques have appeared in the literature, which overcome or alleviate the problem of summing over large numbers of unoccupied states [20, 21, 22, 18].

In particular we discuss in detail the accuracy and efficiency of the approach. We compare results for ionization potentials, electron affinities and band structures with those of existing GW calculations, and we address the convergence of quasiparticle energies with respect to the number of unoccupied states, as well as the use of plasmon-pole models for the frequency dependence of the dielectric matrix. We provide well-converged numerical values for excitation energies of molecules and semiconducting solids, which may serve as benchmarks for other numerical techniques, and which allow one to establish the performance of many-body perturbation theory at the non self-consistent GW level, with respect to available experiments.

2.2 Implementations of the GW approach

Within the framework of many-body perturbation theory, quasiparticle energies, E_n^{qp}, and wavefunctions, $\psi_n^{qp}(r)$, are obtained by solving the following quasiparticle equation:

$$[T + V_{ion}(\mathbf{r}) + V_H(\mathbf{r})]\psi_n^{qp}(\mathbf{r}) \quad + \quad \int d\mathbf{r}' \Sigma(\mathbf{r}, \mathbf{r}'; E_n^{qp})\psi_n^{qp}(\mathbf{r}') = E_n^{qp}\psi_n^{qp}(\mathbf{r}), \tag{2.2.1}$$

where T is the kinetic energy operator, $V_{ion}(\mathbf{r})$ is the external potential of the nuclei, $V_H(\mathbf{r})$ is the Hartree potential, and Σ is the non-local, energy-dependent, and non-Hermitian self-energy operator, which describes many-body electronic interactions.

In the GW approximation [8], the self-energy is expressed in terms of the interacting

one-electron Green's function G and the screened Coulomb interaction W:

$$\Sigma(\mathbf{r}, \mathbf{r}'; i\omega) = \frac{1}{2\pi} \int d\omega' G(\mathbf{r}, \mathbf{r}'; i(\omega - \omega')) W(\mathbf{r}, \mathbf{r}'; i\omega'), \tag{2.2.2}$$

where $W = \epsilon^{-1} \cdot v_c$, ϵ is the dielectric matrix, and v_c is the bare Coulomb potential.

In most practical implementations of the GW method, the one-electron Green's function is approximated by the non-interacting one, G_0, evaluated using eigenvalues, ε_n, and wavefunctions, ψ_n, obtained from Kohn-Sham (KS) Hamiltonians:

$$G_0(\mathbf{r}, \mathbf{r}', i\omega) = \sum_n \frac{\psi_n(\mathbf{r}) \psi_n^*(\mathbf{r}')}{i\omega - \varepsilon_n}. \tag{2.2.3}$$

The screened Coulomb interaction W_0 is computed within the random phase approximation (RPA):

$$W_0 = \epsilon^{-1} \cdot v_c = v_c + v_c \cdot \chi \cdot v_c, \tag{2.2.4}$$

where ϵ^{-1} denotes the inverse RPA dielectric matrix, and χ is the interacting density-density response function. In the current notation $v_c \cdot \chi = \int d\mathbf{r}'' v_c(\mathbf{r}, \mathbf{r}'') \chi(\mathbf{r}'', \mathbf{r}'; i\omega)$, and similarly for all other quantities. Within the RPA, χ is related to the non-interacting density-density response function, χ_0, by the equation: $\chi = (1 - \chi_0 \cdot v_c)^{-1} \cdot \chi_0$, and χ_0 is written in terms of KS eigenvalues and orbitals:

$$\chi_0(\mathbf{r}, \mathbf{r}'; i\omega) = 4 \ \mathrm{Re} \sum_{cv} \frac{\psi_v^*(\mathbf{r}) \psi_c(\mathbf{r}) \psi_c^*(\mathbf{r}') \psi_v(\mathbf{r}')}{i\omega - (\varepsilon_c - \varepsilon_v)}, \tag{2.2.5}$$

where the subscripts v and c indicate valence and conduction states, respectively (or occupied and virtual states in the case of molecules). Within first-order perturbation theory, quasiparticle energies (QPEs) are obtained as corrections to the unperturbed KS eigenvalues:

$$E_n^{qp} = \varepsilon_n + \langle \psi_n | \Sigma_{G_0 W_0}(E_n^{qp}) | \psi_n \rangle - \langle \psi_n | V_{xc} | \psi_n \rangle, \tag{2.2.6}$$

where V_{xc} is the exchange-correction potential entering the chosen KS Hamiltonian.

In conventional GW calculations [12], one computes the non-interacting Green's function $G_0(\mathbf{r}, \mathbf{r}'; i\omega)$ and the density-density response function $\chi_0(\mathbf{r}, \mathbf{r}'; i\omega)$ directly from the expressions of Eq. (2.2.3) and Eq. (2.2.5), respectively. This approach thus requires a summation over both occupied and unoccupied orbitals, hereafter refered to as the sum-over-states (SOS) approach. Despite being widely used, there are serious technical difficulties that limit its applicability to systems with more than a few tens of atoms.

First, a large number of unoccupied states is usually required to converge QPEs [16, 17, 18, 19]. Even for molecules with a few atoms, several thousands of unoccupied states might be required. Due to this slow convergence, an extrapolation scheme is often employed to obtain QPEs in the limit of infinite number of unoccupied states [19, 23]. However, the final result may depend on the specific function used in the extrapolation procedure.

Second, since the inverse dielectric matrix ϵ^{-1} enters the calculation of the screened Coulomb interaction W_0 (Eq. 2.2.4), if one evaluate ϵ directly from the expression of χ_0 in Eq. 2.2.5 then ϵ must be stored and inverted. As a result, large computational resources, in terms of both memory and CPU time, are usually required in practical applications. In many cases, the size of the dielectric matrix is truncated to make the calculation numerically affordable. Often, a cutoff energy of a few Ry is employed (much smaller than that adopted for the KS wavefunctions) when using a plane-wave basis set to represent the dielectric matrix, leading to a significant reduction of the computational resources. However, this approximation may lead to incorrect converged results of QPEs as shown in the case of ZnO [17].

Furthermore, in the conventional GW approach, the frequency dependence of the dielectric matrix is often approximated by various plasmon-pole models [12, 24, 25, 26]. This approximation not only reduces the computational workload, but also provides an analytical expression for the self-energy as a function of frequency. However, it has

been shown that results obtained with different plasmon-pole models may be significantly different [27, 14], and thus it is important to develop efficient techniques to explicitly take into account the frequency dependence of W. In the following, we present in detail the scheme to compute QPEs, which avoids all the difficulties mentioned above.

2.3 GW approach using the spectral decomposition of the dielectric matrix

2.3.1 Iterative calculation of dielectric eigenpotentials

The concept of dielectric band structure was introduced to investigate the screening properties of solids more than two decades ago [28]. This idea is based on the eigenvalue decomposition of the dielectric matrix, computed within the traditional SOS approach. Therefore it is not surprising to see a moderate number of dielectric band structure calculations in literature, mostly for crystalline semiconductors and insulators [29, 30, 31].

The renewed interest in the calculation of RPA correlation energy in the last few years has created a demand for efficient calculations of dielectric eigenvalues. In this research direction, we and our collaborators have developed an efficient computational approach that allows for calculation of dielectric eigenvalues and eigenpotentials without computing any empty states. This is done by combining iterative diagonalization algorithms with the well-established density-functional perturbation theory technique for the calculation of charge density response to a given external perturbation [32, 33, 34].

To go into some detail of our approach, we first note that standard iterative diagonalization algorithms, such as conjugate-gradient, Davidson, or orthogonal iteration (with Ritz acceleration), do not require an explicit representation of the matrix to be diagonalized; instead only the result of the action of the matrix on a generic set of trial vectors

is required. In the linear regime, the action of the RPA dielectric matrix ϵ_{RPA} on a trial potential ΔV can be written as:

$$\epsilon_{RPA}\Delta V = (1 - v_c \cdot \chi_0)\Delta V, \tag{2.3.1}$$

and therefore, the practical problem of computing the action of ϵ_{RPA} on a trial potential ΔV is basically equivalent to the problem of applying χ_0 on ΔV. By definition the product of χ_0 and ΔV gives the charge density response Δn of the system to the application of the trial potential ΔV, that in this context has to be intended as an internal screened potential. The charge density response is obtained within the first-order perturbation theory:

$$\Delta n(\mathbf{r}) = 4 \ \mathrm{Re} \sum_v \psi_v^*(\mathbf{r})\Delta\psi_v(\mathbf{r}), \tag{2.3.2}$$

where $\Delta\psi_v(\mathbf{r})$ is the first-order variation of KS orbital $\psi_v(\mathbf{r})$. To avoid the computation of empty states in the calculation of the charge density response, we employ the DFPT technique to determine $\Delta\psi_v(\mathbf{r})$ from the solution of a linear equation [35]:

$$[\hat{H}_0 + \alpha\widehat{\mathrm{P}}_v - (\varepsilon_v + i\omega)]|\Delta\psi_v\rangle = -(1 - \widehat{\mathrm{P}}_v)\Delta V|\psi_v\rangle, \tag{2.3.3}$$

where $\hat{H}_0$ is the unperturbed KS Hamiltonian and $\widehat{\mathrm{P}}_v = \sum_v |\psi_v\rangle\langle\psi_v|$ is the projector onto the occupied (valence) manifold. The value of the positive constant α is chosen to be larger than the valence bandwidth so that the linear equation is not singular when $i\omega \to 0$. Details of the solution of Eq. (2.3.3) in plane-wave pseudopotential approach for extended systems are discussed throughly in Ref. [35]. Since the full dielectric matrix is non-Hermitian, for practical purposes it is useful to define the symmetric form $\tilde{\epsilon}_{RPA} = v_c^{-\frac{1}{2}}\epsilon_{RPA}v_c^{\frac{1}{2}}$. This form can be diagonalized by standard iterative algorithms for Hermitian problems and, if necessary, the final non-Hermitian form can be obtained by a simple

similarity transformation.

2.3.2 The self-energy operator

The self-energy in Eq. (2.2.2) can be written as a sum of an exchange (Σ_x) and a correlation (Σ_c) term, where $\Sigma_x = G_0 v_c$ and $\Sigma_c = G_0(v_c \cdot \chi \cdot v_c)$. The integral on ω defining Σ_x leads to the well-known expression of the Hartree-Fock exchange energy and can be evaluated with the techniques of Ref. [33, 36]. In the following we focus on the calculation of the correlation term of the self-energy, Σ_c.

We suppress the subscript RPA for convenience and we consider the dielectric matrix in the Hermitian form: $\tilde{\epsilon} = v_c^{-\frac{1}{2}} \cdot \epsilon \cdot v_c^{\frac{1}{2}} = 1 - v_c^{\frac{1}{2}} \cdot \chi_0 \cdot v_c^{\frac{1}{2}} \equiv 1 - \overline{\chi}_0$. Analogously, the inverse of $\tilde{\epsilon}$ is given by $\tilde{\epsilon}^{-1} = v_c^{-\frac{1}{2}} \cdot \epsilon^{-1} \cdot v_c^{\frac{1}{2}} = 1 + v_c^{\frac{1}{2}} \cdot \chi \cdot v_c^{\frac{1}{2}} \equiv 1 + \overline{\chi}$, where $\overline{\chi}$ and $\overline{\chi}_0$ are related by the equation:

$$\overline{\chi} = (1 - \overline{\chi}_0)^{-1} \cdot \overline{\chi}_0. \tag{2.3.4}$$

Using the definition of $\overline{\chi}$ the correlation part of the self-energy is written as:

$$\Sigma_c = G_0(v_c^{\frac{1}{2}} \cdot \overline{\chi} \cdot v_c^{\frac{1}{2}}). \tag{2.3.5}$$

Following Refs. [37, 34, 32], we use a spectral decomposition to represent the inverse of the Hermitian dielectric matrix and $\overline{\chi}$:

$$\tilde{\epsilon}^{-1}(\mathbf{r}, \mathbf{r}') - 1 \equiv \overline{\chi}(\mathbf{r}, \mathbf{r}') = \sum_{i}^{N_{eig}} (\lambda_i^{-1} - 1) \Phi_i^*(\mathbf{r}) \Phi_i(\mathbf{r}'), \tag{2.3.6}$$

where λ_i and $\Phi_i(\mathbf{r})$ denote eigenvalues and eigenpotentials of $\tilde{\epsilon}$, respectively. It has been shown that dielectric eigenvalue spectra of non-metallic solids, nanostructures, and molecular systems exhibit a rapid decay toward one [37, 34, 32, 28, 29, 30]. One therefore expects a relatively small number of eigenvalues N_{eig} to be necessary to numerically con-

verge the summation of Eq. (2.3.6). In principle, the eigenvalue decomposition of $\overline{\chi}$ can be computed for each value of the imaginary frequency $i\omega$. However, it is computationally more convenient to expand $\overline{\chi}$ in terms of the eigenvectors of the static $\overline{\chi}(\mathbf{r}, \mathbf{r}'; i\omega = 0)$:

$$\overline{\chi}(\mathbf{r}, \mathbf{r}'; i\omega) = \sum_{i,j=1}^{N_{eig}} \overline{c}_{ij}(i\omega)\Phi_i^*(\mathbf{r})\Phi_j(\mathbf{r}'), \tag{2.3.7}$$

where $\overline{c}_{ij}$ are expansion coefficients. In this representation $\overline{\chi}(i\omega \neq 0)$ is not diagonal but the number N_{eig} (Eq. 2.3.7) necessary in practical calculations is small and, correspondingly, the size of the matrix of the coefficients $\overline{c}_{ij}$ can be easily managed.

The explicit calculation of the coefficients $\overline{c}_{ij}$ is carried out starting from the analogous definition of $\overline{c}^0_{ij}(i\omega) = \int \Phi_i^*(\mathbf{r})\overline{\chi}_0(\mathbf{r}, \mathbf{r}'; i\omega)\Phi_j(\mathbf{r}')d\mathbf{r}d\mathbf{r}'$, corresponding to the expansion coefficients of $\overline{\chi}_0$:

$$\overline{c}^0_{ij}(i\omega) = 2\sum_v \left\{ \langle \psi_v(v_c^{\frac{1}{2}}\Phi_i)|\widehat{\mathrm{P}}_c(\widehat{\mathrm{H}}_0 - \varepsilon_v + i\omega)^{-1}\widehat{\mathrm{P}}_c|\psi_v(v_c^{\frac{1}{2}}\Phi_j)\rangle + c.c. \right\}, \tag{2.3.8}$$

here $\widehat{\mathrm{P}}_c = 1 - \widehat{\mathrm{P}}_v$ is the projection operator on the unoccupied (conduction) electronic state manifold. Once $\overline{c}^0_{ij}$ and thus the matrix $\overline{\chi}_0$ are computed, the matrix $\overline{\chi}$ is obtained from Eq. (2.3.4) by simple inversion. As already mentioned, the dimension of these matrices is given by N_{eig} and it is thus much smaller than that of the dielectric matrix expanded in plane-waves.

By inserting the expansion of $\overline{\chi}$ (Eq. 2.3.7) and the non-interacting density functional theory Green's function (Eq. 2.2.3) into the definition of the correlation part of self-energy (Eq. 2.3.5) one gets:

$$\langle \psi_n|\Sigma_c(i\omega)|\psi_n\rangle = \frac{1}{2\pi}\sum_{i,j=1}^{N_{eig}} \int d\omega'\ \overline{c}_{ij}(i\omega')\langle \psi_n(v_c^{\frac{1}{2}}\Phi_i)|(\widehat{\mathrm{H}}_0 - i(\omega - \omega'))^{-1}|\psi_n(v_c^{\frac{1}{2}}\Phi_j)\rangle, \tag{2.3.9}$$

where $|\psi_n(v_c^{\frac{1}{2}}\Phi_j)\rangle$ is a vector whose coordinate representation is $\langle \mathbf{r}|\psi_n(v_c^{\frac{1}{2}}\Phi_j)\rangle = \psi_n(\mathbf{r})\int d\mathbf{r}' v_c^{\frac{1}{2}}(\mathbf{r},\mathbf{r}')\Phi_j(\mathbf{r}')$, and $\widehat{\mathrm{H}}_0$ is the unperturbed KS Hamiltonian. The computation of the matrix elements of the inverse, shifted Hamiltonian entering Eq. (2.3.9) and Eq. (2.3.8) may be carried out by using the Lanczos algorithm described in the next section.

2.3.3 Lanczos algorithm

As shown in Eq. 2.3.8 and Eq. 2.3.9, the evaluation of *GW* quasi-particle corrections require the evaluation of several diagonal and off-diagonal elements of the resolvent of the Hamiltonian $\widehat{\mathrm{H}}_0$ for imaginary frequencies $i\omega$. The general mathematical problem that needs to be solved can be written as:

$$g(i\omega) = \langle u|(\mathcal{H} - i\omega)^{-1}|v\rangle, \tag{2.3.10}$$

where $\mathcal{H}$ is a generic Hermitian operator and u and v are generic vectors. The calculation of $g(i\omega)$ can be obtained by using standard iterative techniques to solve Hermitian linear systems (e.g. the conjugate gradient algorithm) [38]. In this case the linear systems $(H - i\omega)|z\rangle = |v\rangle$ is solved and $g(i\omega)$ is computed as $\langle u|z\rangle$. The disadvantage of this approach is that a different linear system has to be solved for each different value of $i\omega$.

The Lanczos algorithm is known to be an efficient method to compute the resolvent of Hermitian operators by performing a single iterative recursion independent of ω [39]. The standard Lanczos algorithm is limited to the case $u = v$, however for the purpose of this work it is necessary to compute also the off-diagonal elements of $g(i\omega)$. For this task we use the Lanczos algorithm proposed in Ref. [40] and applied for the first time to *GW* calculations in Ref. [20]. This algorithm generates iteratively a series of vectors

$\{q_1, q_2, ...\}$ by using the following procedure [38]:

$$
\begin{aligned}
& q_0 = 0 \\
& q_1 = v/\sqrt{\langle v|v\rangle} \\
& \beta_{n+1}|q_{n+1}\rangle = \mathcal{H}|q_n\rangle - \alpha_n|q_n\rangle - \beta_n|q_{n-1}\rangle,
\end{aligned}
\tag{2.3.11}
$$

where β_{n+1} is determined in order to impose the normalisation condition $\langle q_{n+1}|q_{n+1}\rangle = 1$ and $\alpha_n = \langle q_n|\mathcal{H}|q_n\rangle$. In the orthonormal basis set of the vectors $\{q_1, q_2, ...\}$ the matrix $\mathcal{H}$ has tridiagonal form:

$$
T^j = \begin{pmatrix}
\alpha_1 & \beta_2 & 0 & \cdots & 0 \\
\beta_2 & \alpha_2 & \beta_3 & 0 & \vdots \\
0 & \beta_3 & \alpha_3 & \ddots & 0 \\
\vdots & 0 & \ddots & \ddots & \beta_j \\
0 & \cdots & 0 & \beta_j & \alpha_j
\end{pmatrix},
\tag{2.3.12}
$$

where α and β are the coefficents of the Lanczos recursion Eq. 2.3.11 and j is the maximum number of iterations performed. Analogously to Ref. [40] the value of $g(\omega)$ can be approximated as:

$$
g(i\omega) \approx \langle \zeta^j | (T^j - i\omega)^{-1} | e_1^j \rangle,
\tag{2.3.13}
$$

where ζ^{jT} is a j-dimensional vector defined as $(\langle u|q_1\rangle, \langle u|q_2\rangle), \ldots, \langle u|q_j\rangle)$ and e_1^{jT} is the j-dimensional unit vector $(1, 0, \ldots, 0)$. As we will discuss later the dimension j (number of Lanczos iterations) of the matrix T^j necessary to obtain an accurate approximation of $g(i\omega)$ is much smaller than the dimension of the full matrix $\mathcal{H}$.

There are two important advantages in using Eq. 2.3.13. First, the matrix T^j and the corresponding Lanczos iterative recursion do not depend on $i\omega$. Once T^j is generated, the value of $g(i\omega)$ can be computed for many values of $i\omega$ by simple linear algebra operations

in a small j-dimensional space (Eq. 2.3.13). Second, the Lanczos recursion (Eq. 2.3.11) depends only on the right vector $|v\rangle$ and different ζ^j vectors can be generated "on the fly" during the Lanczos chain in order to compute matrix elements of $g(i\omega)$ for different left vectors $\langle u|$. This feature is particularly convenient to evaluate Eq. 2.3.8 and Eq. 2.3.9, where the calculation of several off-diagonal elements ($u \neq v$) is required.

2.3.4 Computational cost

We now turn to a discussion of the efficiency of the our approach in comparison with that of the SOS. We denote by N_v and N_c the valence and conduction bands, and by $N_{pw\psi}$ and $N_{pw\chi_0}$ the number of planewaves used to represent wave functions and the response function χ_0, as well as the dielectric matrix, respectively. It has been shown that [32, 34, 33] the the computational effort for constructing dielectric matrix with N_{eig} eigenpotentials scales as $N_{eig} \times N_{pw\psi} \times N_v^2$. In addition, the generation of Lanczos recursion scales as $N_{eig} \times N_{pw\psi} \times N_v^2 \times N_{Lanczos}$ with $N_{Lanczos}$ is the number of Lanczos iteration performed. Therefore the total workload of our approach scales as $N_{eig} \times N_{pw\psi} \times N_v^2 \times N_{Lanczos}$.

This scaling represents a substantial improvement over the scaling of conventional approaches, $N_{pw\chi_0}^2 \times N_v \times N_c$. In particular, $N_{pw\psi}$ is typically smaller than $N_{pw\chi_0}$ by an order of magnitude since $N_{pw\chi_0}$ relates to density responses and perturbing potentials whose kinetic-energy cutoff is four times larger than the one needed for wavefunctions in the case of norm conserving pseudopotentials and even more than that in the case of ultrasoft pseudopotentials. Furthermore, the number of occupied states N_v for a give system is typically much smaller than the number of empty states N_c required to converge summations in the dielectric matrix and Green's function while the number of Lanczos iteration needed to converge calculations is unlikey to exceed 100. Finally, as we show later, the number of eigenpotentials N_{eig} that needs to be computed is expected to be

three order of magnitude smaller than the size of the response function $N_{pw\chi_0}$. The favorable scaling of our approach over the SOS allows us to compute QPEs of systems with hundreds of valence electrons.

2.4 Benchmark for the benzene molecule

In this section, we illustrate our scheme by considering the benzene molecule, which has been widely investigated using the *GW* approximation [20, 41, 42, 18]. When using the conventional *GW* approach, calculations of ionization potential and electron affinity of this system are known to be challenging due to the slow convergence of QPEs as functions of empty electronic states.

2.4.1 Convergence properties

We demonstrate the convergence properties of our scheme by considering the convergence of the correlation self-energy (Σ_c) as a function of: i) the number of eigenpotentials (N_{eig}) included in the decomposition of the dielectric matrix; ii) the number of Lanczos iteration ($N_{Lanczos}$) in calculations of matrix elements of Greens function and $\overline{\chi}_0(\mathbf{r}, \mathbf{r}'; i\omega)$; iii) the size of the supercell used to simulate the benzene molecule.

Our calculations were performed using the LDA exchange-correlation functional, with norm-conserving pseudopotentials taken from the QUANTUM ESPRESSO distribution, and a PW basis set with a kinetic energy cutoff of 40 Ry to represent wavefunctions. We use a cubic cell with periodic boundary conditions to simulate the benzene molecule, and for the calculations of the self-energy, a cutoff for the Coulomb potential suggested in Ref. [43] was employed to avoid spurious periodic image interactions. Fig. 2.4.1b and Fig. 2.4.1c show the real and imaginary part of the correlation self-energy (Σ_c) of the HOMO level as functions of imaginary frequency, computed for different sizes of the

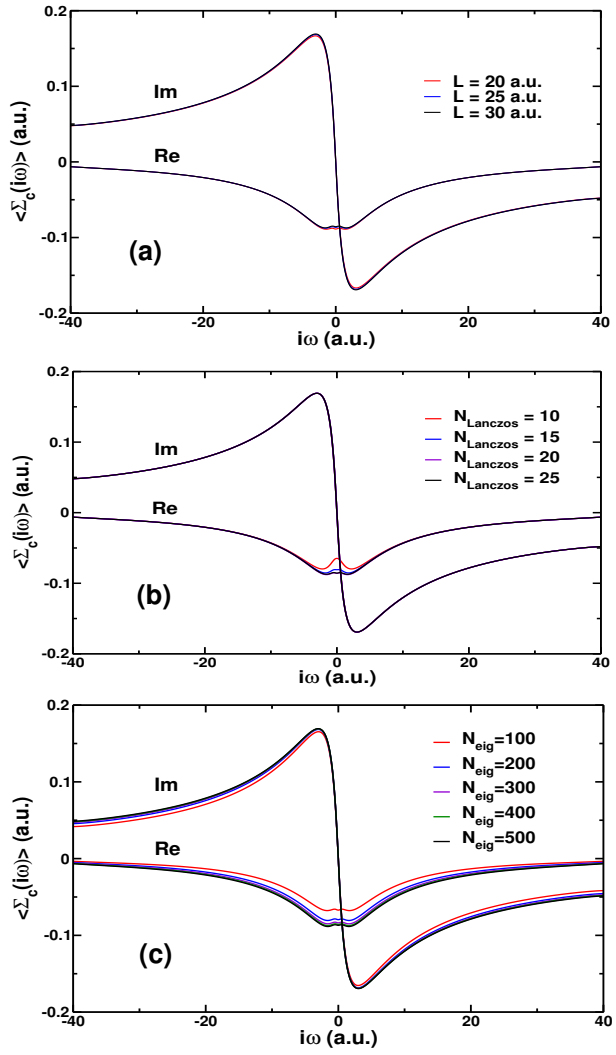

Figure 2.4.1: Imaginary (Im) and Real (Re) part of the correlation self-energy (Σ_c) computed for the HOMO level of the benzene molecule as a function of imaginary frequency $i\omega$: a) for different number of eigenpotentials (N_{eig}) in the definition of the dielectric matrix [see Eq. 2.3.6]; b) for different number of Lanczos iteration in calculations of matrix elements of Greens function and $\overline{\chi}_0(\mathbf{r}, \mathbf{r}'; i\omega)$; and c) for different unit-cell lengths (L).

simulation cell and for different number of the Lanczos iteration (we use a frequency grid with 1000 points and maximum frequency of 40 Ry). It is seen that the correlation self-energy shows a fast convergence as a function of L and $N_{Lanczos}$, with both $\mathrm{Re}(\Sigma_c)$ and $\mathrm{Im}(\Sigma_c)$ are well converged for $L = 25$ a.u. length and $N_{Lanczos} = 25$. These parameters were then employed in all calculations presented below for benzene molecule. We note the convergence of the correlation self-energy with respect to L and $N_{Lanczos}$ can be easily checked using a relatively small number of eigenvalues and eigenvectors in the spectral decomposition of the DM, N_{eig}. Once the values of L and $N_{Lanczos}$ were choosen, the accuracy of QPEs is controlled by a single parameter: N_{eig}, as discussed below.

Fig. 2.4.1a shows the convergence of the real and imaginary part of the correlation self-energy computed for the HOMO level as functions of N_{eig}. It is seen that both $\mathrm{Re}(\Sigma_c)$ and $\mathrm{Im}(\Sigma_c)$ are well converged for $N_{eig} = 300$. We note that using a kinetic energy cutoff of 40 Ry for wavefunctions and $L = 25$ a.u., the number of basis function is $\sim 5.3 \times 10^5$ if it were represented using a plane-wave basis set, more than three order of magnitude larger than N_{eig} and leads to a dielectric matrix of a size at least six order of magnitude larger. Once the correlation self-energy in imaginary frequencies is obtained, its values at real frequencies are computed by the analytic continuation technique with a two-pole formula [44] (the accuracy of the analytic continuation technique will be discussed later) and QPEs are obtained by solving Eq. 2.2.6. In Fig. 2.4.2 we present the convergence of vertical ionization potentials (VIPs) of C_6H_6 as a function of N_{eig}; our computed values are already well converged (within 0.05 eV) for $N_{eig} = 300$. The value of the first VIP, 9.23 eV, is in excellent agreement with experiment[45] (9.3 eV) and results in the literature [20, 42], but slightly higher than the one recently reported in Ref. [18] (9.05 eV). The vertical electron affinity (VEA) exhibits a convergence behavior similar to the VIP, and we obtain a value of -0.81 eV, in fair agreement with experiment [46] (-1.12 eV). A fast convergence with respect to N_{eig} is also observed for all VIPs and the largest error

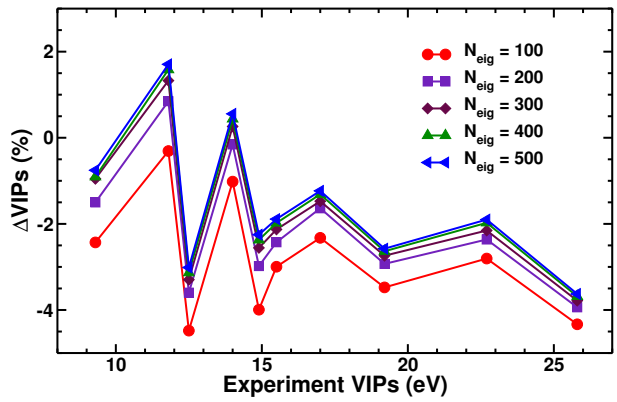

Figure 2.4.2: Differences between the calculated vertical ionization potentials (VIPs) of benzene and experimental results as functions of the number of eigenpotentials N_{eig} included in the definition of the dielectric matrix [see Eq. 2.3.6]: circle, $N_{eig} = 100$; square, $N_{eig} = 200$; diamond, $N_{eig} = 300$; triangle up, $N_{eig} = 400$; and triangle left, $N_{eig} = 500$.

of the computed VIPs is found to be less than 4% compared to experiment [45].

Our computed quasiparticle gap (10.03 eV) is slightly smaller than those reported by Tiago *et al.* [42] (10.29 eV), Neaton *et al.* [41] (10.51 eV), and Samsonidze *et al.* [18] (10.56 eV). We note that in Ref. [42] vertex corrections were included in the calculation, which is beyond the scope of this work. A plasmon pole model and a relatively small kinetic energy cutoff to represent the dielectric matrix (6.0 Ry) were used in Ref. [18] and Ref. [41]. As pointed out in Ref. [27], the use of a plasmon pole model and a small kinetic energy cutoff to represent the dielectric matrix migh introduce substantial errors in the quasiparticle gap and might be the sources of the discrepancies with the results presented here. Finally, we discuss the favorable scaling of our scheme over the traditional SOS approach. As pointed out above, the scaling of the approach presented here is $N_{eig} \times N_{pw\psi} \times N_v^2 \times N_{Lanczos}$ while it is $N_{pw\chi_0}^2 \times N_v \times N_c$ for the traditional SOS approach. For a G_0W_0 calculation for benzene molecule using a cubic supercell of 25 a.u. length and a kinetic energy cutoff of 40 Ry for the wavefunction, the numbers of planewave used to represent the wavefunction and the non-interacting density-density response function are $N_{pw\psi} \sim 6.6 \times 10^4$ and $N_{pw\chi_0} \sim 5.3 \times 10^5$, respectively. Furthermore,

it has been shown that the number of empty states needed to converge the QPEs of benzene molecule of the order of $N_v \sim 2.6 \times 10^3$ in the traditional SOS approach [18]. Taking $N_v = 15$ for benzene molecule and $N_{eig} = 300$, $N_{Lanczos} = 25$ to converge QPEs in our calculation, we estimate the computational cost of our approach is about four order of magnitude smaller than that of the conventional SOS approach. This clearly highlight the advantage of our scheme over the conventional SOS appproach.

2.4.2 Basis functions

An important step in our calculations is to construct basis functions $\{\Phi_i(\mathbf{r})\}$ to represent the inverse of the Hermitian dielectric matrix (see Eq. 2.3.6). As in principle, different choices of basis function can be employed, one might expect it is the linear space spanned by the basis functions that is important, not the detail of each basis function. The accuracy of each basis function $\{\Phi_i(\mathbf{r})\}$ in our calculations is controlled by the convergence threshold in the iteractive procedure. The iterative procedure is considered to be converged when the error in eigenvalues λ_i in two successive iterations is less than this convergence threshold. Fig. 2.4.3 shows the correlation self-energy of the HOMO level of benzene molecule, computed using basis sets $\{\Phi_i(\mathbf{r})\}$ constructed with different convergence threshold. It is seen that there is no significant different in the computed correlation self-energy when the threshold increases from 10^{-6} to 10^{-1}. In addition, as given in Table. 2.1, the computed values of VIP and VEA show a variation of only 0.02 eV when the threshold increases from 10^{-6} to 10^{-1}. These results indicate that the linear spaces spanned by the sets of eigenpotentials computed with different convergence threshold are very similar, if not the same.

This observation shows a significant advantage of our approach. One may save considerably on CPU time by using the eigenpotentials obtained with looser convergence parameters without loosing the accuracy of the calculated QPEs. For example, calcula-

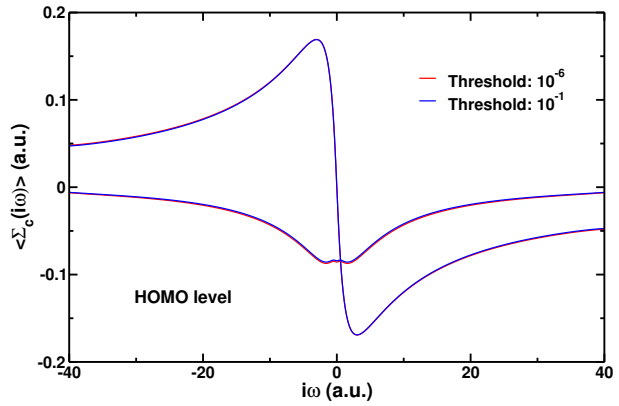

Figure 2.4.3: Imaginary (Im) and Real (Re) part of the correlation self-energy (Σ_c) for the HOMO level of the benzene molecule, computed for different threshold used in the iterative diagonalization procedure.

Table 2.1: First vertical ionization potential (VIP) and electron affinity (VEA) (eV) of the benzene molecule computed for different convergence threshold in the iteractive procedure.

Threshold	10^{-6}	10^{-3}	10^{-2}	10^{-1}
VIP	9.22	9.22	9.21	9.20
VEA	-0.81	-0.81	-0.81	-0.82

tion of 400 eigenpotentials that yields a well-converged VIPs of benzene molecule took about 1.6 h using 32 AMD Opteron processors (2.3 GHz).

2.4.3 Analytic continuation

In this section we demonstrate in detail the analytic continuation technique. As discussed above, after the correlation self-energy at the imaginary frequencies is converged, its values at real frequencies were computed by analytic continuation using a multipole model function:

$$f(z) = a_0 + \sum_{i}^{N_{pole}} \frac{a_i}{z - b_i}, \tag{2.4.1}$$

where a_i and b_i are complex parameters that are determined during the fit along the imaginary axis. It is thefore important to understand the error contributed by the fitting procedure.

Fig. 2.4.4 shows the values of the first VIP and VEA of the benzene molecule computed for different number of pole in the multipole model function. It is seen that the the absolute error introduced in the fitting procedure from the analytic continuation is within 0.1 eV. We stress that the error introduced from the analytic continuation procedure does not affect the way that the convergence of QPEs are controlled. The quantity that needs to be converged is the correlation self-energy in the imaginary frequencies, which can be easily checked for convergence before the analytic continuation procedure is applied. And even when the correlation self-energy is fully converged at the imaginary frequencies, the uncertainty of the computed values of QPEs introduced by the analytic continuation technique is within 0.1 eV. In all calculations below, a two pole formula is employed in the analytic continuation.

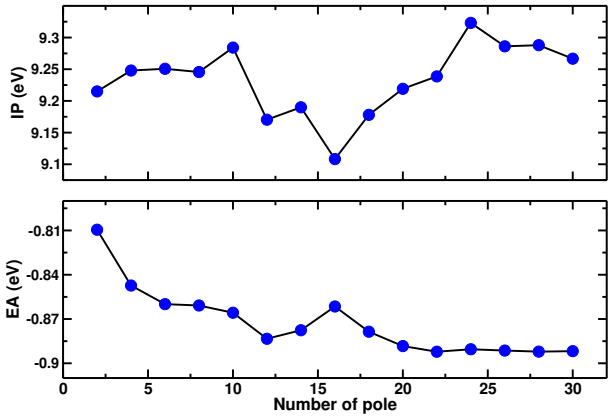

Figure 2.4.4: Vertical ionization potential (VIP) and electron affinity (VEA) of benzene molecule (eV), computed for different number of pole included in the multipole model function. Calculations were performed using $N_{eig} = 400$.

2.5 Comparison with other GW implementations

2.5.1 Comparison with plane-wave basis set implementations

In order to illustrate the importance of converging the number of unoccupied states in GW calculations, in this section we compare QPEs obtained with the current method to those using the conventional plane-wave based SOS approach. Two systems were selected, the benzene diamine-$C_6H_8N_2$ (BDA) molecule and C_{60}, for which SOS's results of the IP and EA have been reported in the literature [18, 19].

Benzene diamine molecule

The conventional G_0W_0 approach that requires sums over unoccupied states was employed in Ref. [19] to investigate the BDA molecule. In order to directly compare with this work, we adopted the Perdew-Burke-Ernzerhof (PBE) exchange-correlation functional [7] with a plane-wave cutoff of 60 Ry, and periodic boundary conditions with a cell of 30 a.u. The equilibrium geometry of the gas-phase BDA was determined using DFT with the PBE functional. We used $N_{eig} = 300$, which yields values of QPEs converged within 0.02 eV.

Table 2.2 summarizes the IP and EA for the gas phase BDA, computed at different

Table 2.2: Ionization potential (IP) and electron affinity (EA) (eV) of the benzene diamine (BDA) molecule, as obtained from Kohn-Sham eigenvalues (PBE-eig), and from GW calculations at different levels of theory. G_1W_1 corresponds to calculations where both G and W are updated using the QPEs from the previous cycle, while only W is updated in the G_0W^1 calculation. The number of unoccupied states (N_c) included in conventional GW calculations is also given.

		No. of unocc. orb	IP	EA
Ref. [19]				
	PBE-eig	$\cdots$	4.22	1.03
	G_0W_0	N_c=1024	5.89	-2.18
	G_0W_0	$N_c \to \infty$	6.64	-1.42
	G_0W_1	$N_c \to \infty$	6.71	-1.65
	G_1W_1	$N_c \to \infty$	6.85	-1.59
This work				
	PBE-eig	$\cdots$	4.21	1.02
	G_0W_0	$\cdots$	6.88	-1.02
Expt.[47]		$\cdots$	7.34	$\cdots$

levels of theory. Consistent with previous work [19], the IP computed at the DFT/PBE level of theory (PBE-eig) underestimates the experimental value by 3.12 eV. Our G_0W_0 calculations significantly improve the IP value, with the error relative to experiment reduced to 0.46 eV.

The slow convergence of QPEs of the BDA molecule with respect to the number of unoccupied states has been discussed in Ref. [19]. In particular, at the G_0W_0 level of theory, when 1024 unoccupied states are included, the computed IP shows a deviation as large as 0.75 eV compared with the $N_c \to \infty$ extrapolated value. Extrapolating the calculated values of IP to infinite N_c reduces the error with respect to experiment from 1.45 eV to 0.7 eV, and depending on whether G or W is updated using the QPEs from the previous cycle, this error may be further reduced to 0.5 eV.

At the G_0W_0 level of theory, our result for the IP is higher and closer to experiment than the extrapolated value reported in Ref. [19]. The discrepancy may be attributed to

Table 2.3: Ionization potential (IP) and electron affinity (EA) (eV) for buckminsterfullerene (C_{60}) as obtained from Kohn-Sham eigenvalues (PBE-eig), and from G_0W_0 calculations. The number of unoccupied states (N_c) included in conventional GW calculations is also given. SAPO stands for simple approximate physical orbitals [18].

		No. of unocc. orb	IP	EA
Ref. [18]				
	PBE-eig	$\cdots$	5.84	4.19
	G_0W_0	N_c=5370	6.76	2.19
	G_0W_0	27387 SAPOs	7.21	2.62
This work				
	PBE-eig	$\cdots$	5.81	4.13
	G_0W_0	$\cdots$	7.31	2.74
Expt.[47]		$\cdots$	7.64	2.69

the use of a relatively small cutoff for the dielectric matrix (6.0 Ry), as well as the use of a plasmon-pole model [19]. Additional errors may stem from the choice of the analytical form of the function used to extrapolate the value of IP to infinite N_c. Similar to the IP, our G_0W_0 result for EA is higher than the extrapolated value reported in Ref. [19]; unfortunately no experimental value for EA of the BDA molecule is available.

Buckminsterfullerene C_{60}

For buckminsterfullerene C_{60}, we employed the PBE exchange-correlation functional with a plane-wave cutoff of 40 Ry, and periodic boundary conditions with a cell of 40 a.u.. Our computed values of IP and EA at the DFT/PBE level of theory are in excellent agreement with previous studies [18] (see Table 2.3). QPEs of C_{60} were calculated using $N_{eig} = 700$, that yields converged values of IP and EA within 0.03 eV.

Calculations of QPEs of C_{60} using the conventional G_0W_0 approach showed a slow convergence with respect to the number of unoccupied states [18]. In particular, using 5370 unoccupied states, the values of the IP and EA are strongly underestimated compared to experiment, with errors of 0.88 eV and 0.51 eV, respectively. In contrast, with

the current approach that avoids the direct evaluation of unoccupied states, we found values of 7.31 eV and 2.74 eV for the IP and EA, respectively, in good agreement with experiment (within 0.30 eV).

The authors of Ref. [18] also addressed the problem of the computational cost associated with the construction of unoccupied states in G_0W_0 calculations, and proposed to replace the DFT unoccupied states with simple approximate physical orbitals (SAPOs) that are computationally cheaper to generate. As shown in Table 2.3, our results for EA and IP are in good agreement with those reported in Ref. [18] when a large number of SAPOs (27387) is included in G_0W_0 calculations. We note that the approach of Ref. [18] still requires to carry out explicit sums over SAPOs, and the use of a plasmon-pole model together with a relatively small energy cutoff (6.0 Ry) to represent the dielectric matrix.

2.5.2 Comparison with localized basis set implementations

Recently QPEs of several molecules of interest for photovoltaic applications were investigated within the *GW* approximation [48]. These calculations were based on the conventional implementation of the *GW* method with a localized basis-set and a direct sum over unoccupied states. In order to further validate our scheme, and to compare our results with those obtained with localized basis-sets, we computed QPEs for anthracene ($C_{14}H_{10}$), tetracene ($C_{18}H_{12}$), pentacene ($C_{22}H_{14}$), C_{60} and phthalocyanine-$C_{32}H_{18}N_8$ (H_2Pc) molecules, whose values were reported in Ref. [48].

Similar to Ref. [48], our calculations were performed with the local density approximation (LDA) for the exchange-correlation functional. A kinetic energy cutoff of 40 Ry for wavefunctions was employed for all systems, except for the H_2Pc molecule where we used 70 Ry due to the presence of the nitrogen atom. The size of the supercell was chosen in such a way that the distance between periodic images was at least 10 Å. The QPEs of C_{60}, whose number of valence electrons is the largest in this set of molecule, were

computed using N_{eig}=700, similar to the calculations with the PBE functional (Table III). The number of dielectric eigenvectors needed to converge QPEs of other molecules was also carefully checked, e.g., using N_{eig}=600 yields values of the IP and EA converged within 0.05 eV for the case of H_2Pc.

Table 2.4: Ionization potentials (IP), electron affinities (EA), and quasiparticle gaps (eV) for a set of molecules (see text) as obtained from Kohn-Sham eigenvalues (LDA-eig), and from LDA based G_0W_0 calculations.

	LDA-eig		G_0W_0		Expt.[47]
	Present	Ref. [48]	Present	Ref. [48]	
$C_{14}H_{10}$					
IP	5.18	5.47	7.25	6.89	7.40
EA	2.81	3.22	1.05	0.74	0.50
Gap	2.37	2.25	6.20	6.15	6.90
$C_{18}H_{12}$					
IP	4.85	5.15	7.04	6.37	6.97
EA	3.19	3.58	1.41	1.34	1.07
Gap	1.66	1.57	5.63	5.03	5.90
$C_{22}H_{14}$					
IP	4.63	4.94	6.30	5.98	6.60
EA	3.47	3.84	1.96	1.77	1.40
Gap	1.16	1.10	4.34	4.21	5.20
C_{60}					
IP	6.03	6.37	7.45	7.28	7.64
EA	4.35	4.79	3.05	2.88	2.69
Gap	1.68	1.58	4.40	4.40	4.95
H_2Pc					
IP	5.21	5.56	6.11	6.08	6.40
EA	3.76	4.14	2.35	2.41	···
Gap	1.45	1.42	3.76	3.67	···

The computed values of the IP, EA and quasiparticle gap obtained at different levels of theory are summarized in Table. 2.4. The IPs computed within the DFT/LDA (LDA-eig) significantly underestimate experimental values. In contrast, for molecules whose experimental EAs are available, we found that DFT-LDA overestimates experiments by $\sim$ 2.0 eV. Calculations at the G_0W_0 level of theory show a significant improvement,

bringing the errors of IP and EA to average values of 0.20 and 0.45 eV, respectively, in satisfactory agreement with experiments.

As seen in Table 2.4, our G_0W_0 results for IP and EA are overall larger than those obtained with localized basis-set, leading to better (worse) agreement of IP (EA) with experiment than those reported in Ref. [48]. Quasiparticle gaps for this set of molecule, turned out to be in excellent agreement with the results of Ref. [48], yielding an average error of 0.60 eV compared to experiment.

Direct comparison with the G_0W_0 results reported in Ref. [48] is not straightforward, due to several technical differences. As shown in Table 2.4, there are already significant differences in the results of ground state calculations using localized basis sets and plane-waves. For all molecules studied here, we found that absolute values of LDA eigenvalues obtained with plane-waves are systematically smaller than those reported with a localized basis set by 0.30-0.40 eV. Such large differences cannot be attributed to the sole change in the relaxed geometries of the molecules. More importantly, these discrepancies are not negligible when compared, for example, to the quasiparticle corrections of 0.90 eV (this work) and 0.52 eV (Ref. [48]) for IPs of H_2Pc, respectively. We note that the ground state calculations of Ref. [48] were performed starting from a double-ζ+polarization (DZP) basis; then the numerical radial part was fitted by up to five contracted Gaussians in order to exploit the computational efficiency of Gaussian basis sets to facilitate *GW* calculations.

2.6 Ionization potentials of the test set G2/97

In order to provide highly converged ionization potentials that can be used as a benchmark for future calculations, we considered a subset of 80 molecules from the G2/97 test set [49]. The open-shell molecules belonging to the G2/97 test set were not considered in this work. To minimize possible effects of molecular geometry on G_0W_0 calculations, we

employed the relaxed geometries at the MP2 level of theory with the 6-31G(d) basis set as published in Ref. [49], without any further structural relaxation. In addition, we used PBE Troullier-Martins norm-conserving pseudopotentials taken from the Abinit distribution [50]. Therefore our results are straightforward to reproduce.

Together with the G_0W_0 method, we considered other approaches to evaluate the IP. First, IPs were obtained from the KS energies corresponding to the highest occupied molecular orbital in DFT calculations. We considered not only the PBE approximation to the exchange-correlation potential but also fully nonlocal approximations, including Hartree-Fock (HF) and hybrid functionals. We choose two popular hybrid functionals, i.e., the PBE0 [51, 52] and the HSE06 [53] ones, both of which contain 25% Fock exchange. All G_0W_0 calculations were performed using ground state PBE eigenvalues and orbitals.

We further evaluated IPs from total energy differences in a so-called delta self-consitent field (ΔSCF) procedure:

$$I = E_0^{N-1} - E_0^N, \tag{2.6.1}$$

where E_0^N is the total energy of the N electron system, that can be computed, e.g., with the PBE or hybrid exchange-correlation functionals.

The mean error (ME) and mean absolute error (MAE) of the computed IPs with respect to experimental data of 80 molecules are summarized in Fig. 2.6.1. We also report ME and MAE for three specific classes of molecules of the test set, specifically non-hydrogen systems, hydrocarbons, and substituted hydrocarbons. As well known, IPs obtained from the PBE-HOMO energies suffer from the self-interaction error [54], and they thus strongly underestimate experiments (negative ME) with a MAE as large as 4.16 eV. Part of this self-interaction error is removed using the PBE0 and HSE06 functionals due to the inclusion of 25% Fock exchange. However the IPs obtained from the HSE06 and PBE0 eigenvalues are still far from experiment with MAEs of 3.13 eV and 2.73 eV, respectively.

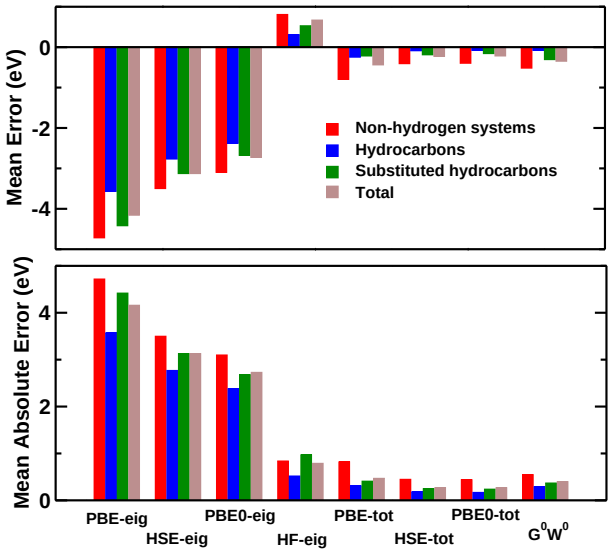

Figure 2.6.1: Deviation of computed ionization potentials, obtained within different approaches (see text), with respect to experimental data. The upper panel shows the mean error (ME), and the lower panel the mean absolute error (MAE) for a 80 molecule set (see text). The grey rectangle corresponds to the MAE and ME computed for all the molecules of the set. The red, blue and green rectangles correspond to the results computed for non-hydrogen systems, hydrocarbons, and substituted hydrocarbons only, respectively.

Since screening is weak in small molecules and HF is self-interaction free, HF results are much closer to experiment, with a MAE of 0.79 eV. In general, IPs computed within HF overestimate experimental values (positive ME). The G_0W_0 approximation based on ground state PBE eigenvalues and orbitals performs better than HF, yielding a MAE of 0.40 eV with respect to experiments. This value is in agreement with those reported in Refs. [55, 56]. Therefore it appears that irrespective of the implementation, IPs obtained at the G_0W_0 level of theory, with PBE ground state eigenvalues and orbitals, show a MAE in the range of 0.4-0.5 eV, with respect to experiment. Not surprisingly, the ΔSCF procedure provides good quality IPs. We find that ΔSCF results with the PBE0 and HSE06 hybrid functionals perform equally well, with a MAE of 0.27 eV. They are superior to those based on the semilocal PBE exchange-correlation functional, with a MAE of 0.47 eV. We note that for some molecules, the ΔSCF procedure based on the PBE functional severely underestimates experiment, e.g., the deviation is as large as 1.58 eV for the BF_3

molecule. In such cases, the G_0W_0 approximation shows significant improvement over the PBE based ΔSCF, leading to very good agreement with the hybrid functionals based ΔSCF calculations.

For all approaches presented here, the hybrid functionals based ΔSCF performs the best for IPs when compared to experimental results. However, unlike the G_0W_0 approach, it does not allow one to access the full energy spectrum of the system and its generalization to extended systems is not straightforward, although approximate forms of ΔSCF for solids have been recently proposed in the literature [57].

2.7 Band gaps of semiconductors

In this section, we turn to extended systems and we consider different semiconductors whose QPEs have been reported in literature. In addition we present a study of the band gap of a disordered system represented by a 56 atoms supercell (256 valence electrons). Note that for the purpose of treating systems with large supercells, in the present work we sample the Brillouin zone using the Γ point only.

In extended systems, an accurate sampling of the Brillouin zone is necessary to compute the long-wavelength component of the dielectric matrix [58]. In the present work, the "head" of the dielectric matrix ($\mathbf{G} = \mathbf{G}' = 0$) at frequency $i\omega$ is computed separately using the Lanczos chain algorithm with dense k-point meshes. Note that when sampling the Brillouin zone with the Γ point only, the "wings" of the dielectric matrix ($\mathbf{G} = 0, \mathbf{G}' = 0$) vanish [59].

2.7.1 Quasiparticle energies of Si, AlAs and SiC

We considered Si first, which has been extensively studied by the electronic structure community, and for which several *GW* results are available. In addition, we present

results for AlAs, a medium-gap semiconductor, and for SiC, a wide-gap semiconductor. We used a kinetic energy cutoff of 25 Ry for Si; 40 Ry for AlAs and SiC. Experimental lattice constants of 10.26 a.u. for Si; and 10.67 a.u. and 8.24 a.u. for AlAs and SiC were used, respectively. For all systems we employed a 64-atom cubic supercell and we sampled the corresponding Brilloun zone using the Γ point. The "head" of the dielectric matrix at the frequency $i\omega$ was computed with a 4×4×4 k-point grid. In order to compare to available *GW* calculations, we used the LDA exchange-correlation functional. Similar to calculations for molecules, we find that QPEs are converged with a relatively small number of Lanczos iterations (∼20-25). For all systems considered here, the QPEs were converged within 0.05 eV when using 600 dielectric eigenpotentials.

Table 2.5 presents the calculated QPEs at points of high-symmetry for Si, together with available theoretical and experimental results. For all energy bands, our results are in excellent agreement with those of other calculations using pseudopotentials, and plane-wave basis set, and full frequency integration (no plasmon-pole models) [16, 60, 61]. In the vicinity of the valence band maximum (VBM) our results are also consistent with those obtained with plasmon-pole models [62, 63]. However, as one moves away from the VBM, significant differences are found, e.g., our computed valence-band width is 11.64 eV, while it is 11.90-11.95 eV when using plasmon-pole models. These large deviations are not surprising, in fact, for Si the same behavior was observed with conventional *GW* approaches, when comparing QPEs obtained with plasmon-pole models and those with the contour deformation method that takes into account the full frequency dependence of the screened Coulomb interaction [64]. Our computed value of the valence-band width is in excellent agreement with that reported in Ref. [64] using the contour deformation method (11.7 eV).

Except for the energy band X_{1c}, our results are also in good agreement with those using the all-electron full-potential projector augmented wave method (PAW) [65]. In

addition, the consistency between our results and all other pseudopotential based calculations indicates that the large deviation ($\sim$0.35 eV for X_{1c}) from Ref. [65] may stem from the contribution of core-electrons that are explicitly taken into account in PAW calculations [65].

Computed QPEs for AlAs and SiC are shown in Table 2.6 and 2.7, respectively. For AlAs our results are in good agreement with those using pseudopotentials with a plasmon-pole model, indicating that at least for AlAs, plasmon-pole models are a good approximation for the frequency dependence of the dielectric matrix. In constrast, for SiC, when comparing to results obtained with plasmon-pole models, we find small deviations in the vicinity of the VBM, but large variations away from the VBM. Similar to the case of Si, when compared to all-electron PAW approach, large variations ($\sim$0.5 eV) are observed for the energy band X_{1c} of both AlAs and SiC.

Based on our observations, we conclude that using plasmon-pole models may be a good approximation only for energy levels close to the band gap. For all systems considered here, the convergence of QPEs with respect to the number of unoccupied states, as well as to the cutoff of the dielectric matrix does not appear to be as severe as, e.g. in the case of ZnO [17].

2.7.2 Band gaps of amorphous Si_3N_4

As a final example, we present calculations of the quasiparticle band gap of amorphous Si_3N_4. We considered a model containing 56 atoms (256 valence electrons), taken from Ref. [66], whose structural and dielectric properties are in good agreement with experiments [67, 66]. For G_0W_0 calculations, we used the PBE exchange-correlation functional with a kinetic energy cutoff of 60 Ry. The Brillouin zone was sampled with the Γ point only. We used 600 dielectric eigenpotentials to compute the quasiparticle band gap.

The band gap of amorphous Si_3N_4 measured in experiment is sensitive to the prepa-

Table 2.5: Calculated quasiparticle energies at points of high-symmetry for Si (in eV), together with available theoretical and experimental results (as quoted in Ref. [60]). Unless noted otherwise in the "Comment" row, all calculations use pseudopotentials, plane-waves, and no plasmon-pole models (PPM). GO stands for Gaussian orbitals and PAW for all electron full-potential projector augmented wave. Energies are measured relative to the valence band maximum.

	Present	Ref. [16]	Ref. [60]	Ref. [61]	Ref. [62]	Ref. [63]	Ref. [65]	Expt.
Γ_{1v}	-11.64	-11.49	-11.57	-11.57	-11.90	-11.95	-11.85	-12.5±0.6
Γ'_{25c}	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Γ_{15c}	3.25	3.24	3.24	3.23	3.25	3.36	3.09	3.40, 3.05
Γ'_{2c}	3.92	3.89	3.94	3.96	3.86	3.89	4.05	4.23, 4.1
X_{1v}	-7.75	-7.58	-7.67	-7.57	-7.90	-7.95	-7.74	
X_{4v}	-2.88	-2.80	-2.80	-2.83	-2.96	-2.93	-2.90	-2.90, -3.3±0.2
X_{1c}	1.36	1.41	1.34	1.35	1.31	1.43	1.01	1.25
L'_{2v}	-9.38		-9.39	-9.35	-9.65	-9.70	-9.57	-9.3±0.4
L_{1v}	-6.93		-6.86	-6.78	-7.13	-7.14	-6.97	-6.7±0.2
L'_{3v}	-1.23		-1.17	-1.20	-1.25	-1.25	-1.16	-1.2±0.2
L_{1c}	2.21		2.14	2.18	2.13	2.19	2.05	2.1, 2.4±0.1
L_{3c}	4.00		4.05	4.06	4.13	4.08	3.83	4.15±0.1
Comment					PPM	GO/PPM	PAW	

Table 2.6: Calculated quasiparticle energies at points of high-symmetry for AlAs (in eV), together with available theoretical and experimental results (as quoted in Ref. [62]). Energies are measured relative to the valence band maximum. Notations are the same as in Table 2.5.

	Present	Ref. [62]	Ref. [65]	Expt.
Γ_{1v}	-11.66	-11.51		
Γ'_{15v}	0.0	0.0	0.0	
Γ_{1c}	2.96	2.74	2.72	3.13
Γ_{15c}	5.07	5.06		
X_{1v}	-9.77	-9.67		
X_{2v}	-5.37	-5.55		
X_{5v}	-2.20	-2.27		-2.41
X_{1c}	2.13	2.16	1.57	2.23
X_{3c}	3.08	3.04		
L_{1v}	-10.27	-10.19		
L_{1v}	-5.82	-5.69		
L_{3v}	-0.90	-0.87		
L_{1c}	3.02	2.84	2.99	2.36
L_{1c}	5.63	5.52		
Comment		PPM	PAW	

ration procedure, with values ranging from 4.5 to 5.3 eV [68]. Within the G_0W_0 approximation, we obtained 4.87 eV for the band gap of amorphous Si_3N_4, that falls within the range of the experimental results. As expected, the band gap computed at the PBE level of theory is a strong underestimation of experiments (3.17 eV).

We also computed the band gap of our model with the HSE06 and PBE0 hybrid functionals and obtained values of 4.55 and 5.32 eV, respectively. It turns out that due to the large variation in the measured values, both G_0W_0 and hybrid functional approaches give results in resonable agreement with experiments. However, we note that in general, even if HSE06 and PBE0 hybrid functionals may provide good results for band gap calculations, they can fail for the interface or nanostructures calculations of the same materials [13]. In such cases, the GW approximation is required to provide reliable

Table 2.7: Calculated quasiparticle energies at points of high-symmetry for SiC (in eV), together with available theoretical and experimental results (as quoted in Ref. [62]). Energies are measured relative to the valence band maximum. Notations are the same as in Table 2.5.

	Present	Ref. [62]	Ref. [65]	Expt.
Γ_{1v}	-15.54	-16.08		
Γ'_{15v}	0.0	0.0	0.0	
Γ_{1c}	7.26	7.19	7.23	7.4
Γ_{15c}	8.10	8.18		7.75
X_{1v}	-10.46	-10.96		
X_{3v}	-8.17	-8.44		
X_{5v}	-3.47	-3.53		
X_{1c}	2.31	2.19	1.80	2.39, 2.42
X_{3c}	5.41	5.23		5.2
L_{1v}	-12.06	-12.46		
L_{1v}	-8.92	-9.19		
L_{3v}	-1.10	-1.21		-1.15
L_{1c}	6.43	6.30	6.45	6.35
L_{3c}	8.32	8.25		8.55
Comment		PPM	PAW	

quantities such as band offsets or band alignments.

2.8 Quasiparticle energies of diamondoids

In order to demonstrate the ability to apply our scheme to investigate large systems, in this section we present calculations of QPEs of relatively large diamond clusters that contain up to 424 valence electrons.

We studied electronic properties of small diamondoids constructed from adamantane cages $C_{4n+6}H_{4n+12}$ (n = 1,4) for which both experimental and quantum Monte Carlo (QMC) results are available. Furthermore, to demonstrate the efficiency of our method, we computed QPEs of three H-terminated, spherical diamond clusters, $C_{29}H_{36}$, $C_{66}H_{64}$,

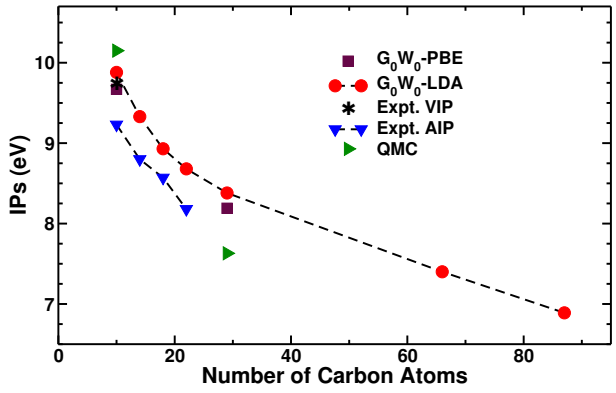

Figure 2.8.1: Computed and measured (Expt.) ionization potentials of diamondoids and diamond clusters as a function of the number of constituent carbon atoms. All computed values refer to vertical ionization potentials (VIP), and were obtained using quantum Monte Carlo (QMC) and many-body perturbation theory with the G_0W_0 approximation. Measured adiabatic ionization potentials (AIPs) are also shown.

and $C_{87}H_{76}$ that has 424 valence electrons. Calculations were performed with the LDA exchange-correlation functional, the parameters used in QPEs calculations were the same as those of the benzene molecule except that a larger cell of 40 a.u. length was used. For the largest system, i.e., the spherical diamond cluster $C_{87}H_{76}$, we have used up to 1000 eigenpotentials in the calculations of QPEs. We note that the size of the dielectric matrix would be at least four order of magnitude larger if it were represented by a planewave basis set.

In Fig. 2.8.1 our calculated values of VIP of diamondoids and the diamond cluster are compared to the experimental adiabatic IP [69] and vertical IP [47], and to the previous theoretical study using QMC [70]. We found that the calculated VIPs at the G_0W_0 level are consistent with experiments, being systematically larger than the measured adiabatic IPs (0.34−0.66 eV), and decreasing as a function of the cluster size.

In Table. 2.8 we present detailed results for the two diamond clusters, $C_{10}H_{16}$ and $C_{29}H_{32}$, for which the QMC results are available. For the smallest diamondoid, i.e., $C_{10}H_{16}$ whose the measured vertical IP is available, our G_0W_0 result based on LDA eigenvalues and wavefunctions is in a very good agreement with experiment with an error less than

Table 2.8: Vertical ionization energies (VIP), electron affinities (VEA) for diamond cluster $C_{10}H_{16}$, $C_{29}H_{36}$ obtained from different level of theory: DFT and G_0W_0 with both the LDA and PBE functional, and from Quantum Monte Carlo (QMC).

	LDA	PBE	G_0W_0-LDA	G_0W_0-PBE	QMC-PBE	Expt.
$C_{10}H_{16}$						
IP	6.58	6.42	9.86	9.67	10.15	9.75
EA	0.86	0.83	-0.55	-0.54	-0.13	
$C_{29}H_{36}$						
IP	5.81	5.60	8.37	8.19	7.63	
EA	0.86	0.81	-0.54	-0.47	-0.29	

0.1 eV. When compared to QMC calculations, we find differences of 0.3 eV and 0.7 eV for $C_{10}H_{16}$ and $C_{29}H_{36}$ using LDA-based G_0W_0, respectively.

Since the PBE exchange-correlation functional was employed in QMC calculations, we also computed the QPEs of $C_{10}H_{16}$ and $C_{29}H_{36}$ using this density functional. When PBE wavefunctions and eigenvalues are used, we obtain VIPs which are smaller by 0.16 and 0.2 eV than that obtained at the LDA level for $C_{10}H_{16}$ and $C_{29}H_{36}$, respectively, and our computed VIPs are in fair agreement with QMC calculations, with differences of 0.46-0.50 eV, as shown in Table. 2.8.

Irrespective of the functional used for ground-state calculations, we find that the computed VEA within the G_0W_0 approximation is almost constant as a function of the cluster size (-0.55 eV with the use of the LDA functional). Our observation is in agreement with x-ray absorption experiment [71] where the LUMO energy of diamondoids show no quantum confinement effect. The origin of the anomalous size dependence of the LUMO energy has been explained due to the localization of LUMO states of diamond clusters on the surface as the size of the clusters increases [70]. In addition, our computed values of the VEA are negative, consistent with QMC results, and the G_0W_0 results for VEAs of $C_{10}H_{16}$ and $C_{29}H_{36}$ obtained with LDA wavefunctions and eigenvalues are 0.42 and 0.26 eV smaller than those obtained in QMC calculations, respectively. We note that

in the absence of quasiparticle corrections to LDA eigenvalues, the VEA is very weakly dependent on cluster size but it is positive (0.86 eV). Using the PBE density functional for ground-state calculations does not change significantly the values of VEA, as shown in Table. 2.8.

2.9 Summary

In summary, we have presented an approach to perform G_0W_0 quasiparticle energy calculations using the spectral decomposition of the static dielectric matrix and we have systematically investigated its performance for a series of molecules and extended systems. The approach presented here offers several advantages: i) it does not require explicit summations over unoccupied states; ii) the dielectric matrix is represented by a small number of dielectric eigenvectors, and the storage and inversion of large dielectric matrices are avoided; iii) the use of plasmon-pole models is not necessary; iv) numerical accuracy is controlled by a single parameter, i.e., the number of dielectric eigenvalues and eigenvectors in the spectral decomposition of the dielectric matrix. These advantages allowed us to carry out GW calculations for systems with several hundreds of valence electrons, and to converge quasiparticle energies in a systematic way.

For all molecules and semiconductors (amorphous and crystalline) considered in this work, the computed quasiparticle energies show good agreement with experiments. We note that although all results presented here are for systems with less than 300 valence electrons, applications of our G_0W_0 approach to large systems with several hundred atoms and thousand of valence electrons are straightforward [72].

Chapter 3

Structural and dynamical properties of oxides surfaces and interfaces

3.1 Ice Ih (0001) surface

3.1.1 Introduction

Ice Ih is the most common crystalline form of ice on Earth and the surfaces of ice Ih crystals play a fundamental role in a variety of chemical and physical processes. For example, chemical reactions on the surfaces of ice particles in the polar stratosphere are an important factor in the seasonal depletion of the ozone layer [73].

Bulk ice Ih crystals are built from a regular hexagonal arrangement of oxygen atoms with disordered protons arranged within the constraints of the Pauling ice rules [74]. At the ideal (0001) basal surface of ice Ih, the water molecules in the interior side of the outermost ice bilayer retain four hydrogen bonds while each molecule in the outer side of the surface bilayer lacks a hydrogen bond and is thus three-fold coordinated. Each of these three-fold coordinated molecules contribute to the surface of ice either a hydrogen atom with a dangling bond (d-H) or an oxygen atom (d-O) with a dangling lone

pair. Contrary to the case of bulkice Ih, whose structure has been widely studied, the microscopic structure of the ice Ih basal surface is not as well understood, inparticular the preferred arrangements of d-H and, to a lesser extent, d-O atoms are yet unclear.

Several experimental studies have addressed the question of the arrangement of d-O atoms. Low energy electron diffraction experiment [75, 76] reveals that oxygen atoms form an ordered-hexagonal (0001) ice bilayer without reconstruction at 90 K. In addition, He scattering studies [77, 78, 79, 80] show that at temperatures of tens of Kelvin, the diffraction patterns of the ice basal surface match those of bulk ice, which indicates an oxygen-ordered bilayer is formed. At higher temperatures the loss of oxygen atom order is observed, with the order-disorder transition at 180 K as seen in He scattering experiments [81], and at 200 K using sum frequency generation (SFG) experiment [82].

While an ordered pattern of oxygen atoms at temperatures below 200 K is widely confirmed by experiments, the understanding of the arrangement of protons at the basal ice surface as a function of temperature is in its infancy. Proton disorder was first studied theoretically by Fletcher [83], who suggested that at low temperatures the basal surface of ice is proton ordered and becomes disordered at 30 K. According to Ref. [83], alternating rows of dangling OH bonds (those with so-called d-H atom) is the most stable surface structure (the Fletcher phase) at temperatures below 30 K. This idea was later supported by Ref. [84], in which total energy calculations using empirical potentials were carried out for surfaces with different proton arrangements. DFT calculations [85, 86, 87] also found a stable proton ordered basal ice surface. However, by considering an order parameter based on the number of dangling OH bonds on the surface, the authors concluded that the Fletcher phase is only one of the lowest energy geometries of the ice surface [85, 86, 87]. Therefore the geometrical arrangement of protons at the basal ice surface remains an open question.

In order to provide a better understanding of the d-H patterns of ice surfaces and to

identify the possible signatures of the Fletcher phase and of proton disordered phases, we computed IR spectra of several ordered and disordered models of the (0001) surface of ice, and we investigated the sensitivity of these spectra to the microscopic details of the surface structure. In particular we discuss possible signatures of disorder in the computed spectra, together with a comparison with available experiments.

3.1.2 Methods

Bulk ice Ih was modeled by a periodically repeated orthorhombic cell with 96 D_2O molecules taken from previous *ab-initio* calculations [88]. Starting from the model of bulk ice Ih, we generated different structures for the basal surface. Following Ref. [85] we define the order parameter:

$$C_{OD} = \frac{1}{N_{OD}} \sum_{i=1}^{N_{OD}} c_i, \tag{3.1.1}$$

where N_{OD} is the number of dangling OD bonds at the surfaces and c_i is the number of nearest neighbor dangling OD bonds around the ith dangling OD bond. Fully proton disordered surfaces yield $C_{OD} = 3.0$ while $C_{OD} = 2.0$ for the proton ordered Fletcher phase; this is the smallest possible value of the order parameter for a nonpolar surface.

Our surface structures were simulated by 96 D_2O-molecule supercells arranged in four bilayers. A vacuum region of ~ 10 Å was introduced into the supercells to reduce spurious interaction between images. We considered two slab models: (i) a proton ordered surface slab with the Fletcher configuration on both sides, i.e., $C_{OD} = 2.0$ (model A in Fig. 3.1.1); (ii) a proton disordered surface slab with proton disordered on both sides (model B and C in Fig. 3.1.1), yielding the same value for the order parameter, i.e., $C_{OD} = 3.0$. The two interior ice bilayers were kept in the same geometry in the slab models i) and ii).

To compute the IR spectra, we carried out *ab-initio* molecular dynamics simulation (AIMD) using the Qbox code [89]. The interatomic forces were computed by solving the electronic structure within DFT using the generalized gradient approximation by Perdew,

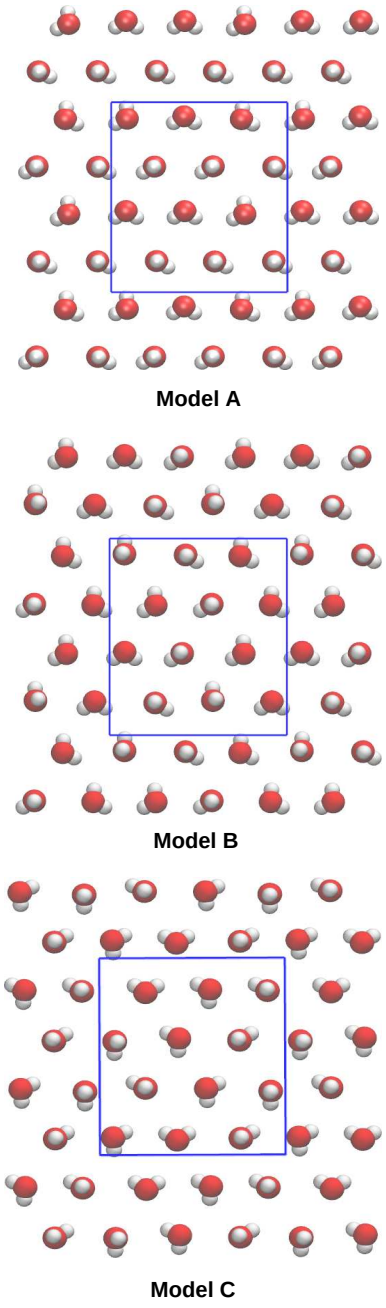

Figure 3.1.1: Models of the ice surface: (A) the top and the bottom surfaces of the proton ordered surface slab (see text), where the Fletcher phase, i.e., $C_{OD} = 2.0$ is employed; (B) and (C) the top and the bottom surfaces of the proton disordered surface slab (see text) with $C_{OD} = 3.0$ for both cases, but with different OD arrangements. White and red spheres are deuterium and oxygen atoms, respectively. Blue line indicates the boundary of the unit cell in the surface plane.

Burke and Ernzerhof (PBE) [7]. We adopted norm-conserving pseudo-potentials [90], and a plane-wave basis set with a kinetic energy cutoff of 85 Ry. At each ionic step the electronic ground state was obtained by solving self-consistently the Kohn-Sham (KS) equation within a Born-Oppenheimer molecular dynamics (BOMD) scheme. We used a time step of 10 a.u. to integrate the ionic equation of motion. Simulations were carried out in the NVE ensemble and trajectories to compute IR spectra were collected over 12 ps, a simulation time similar to that used for bulk ice calculations [91]. We have compared results obtained from averaging over 6.5 and 12 ps trajectories and found no noticeable differences. To avoid surface melting of ice, which appears to set in around 200 K [82], our production runs were performed after equilibrating the systems at 100 K (with a weakly coupled Berendsen thermostat) for 5 ps. We also carried out some runs at 150 K. We found that the hydrogen-bond topology does not change during the equilibrium and production runs.

Within linear response theory, the IR absorption coefficient per unit length is given by the power spectrum of the time-correlation function of the dipole operator. We approximated the quantum time-correlation function with the classical one [92, 91], and we employed the harmonic approximation to obtain the quantum correction factor [93], Within these approximations, the IR absorption coefficient per unit path length of a sample of volume V is given by:

$$\alpha(\omega)n(\omega) = \frac{2\pi\omega^2\beta}{3cV}\int dt e^{-i\omega t}\langle\sum_{ij}\vec{\mu}_i(0)\vec{\mu}_j(t)\rangle, \tag{3.1.2}$$

here $n(\omega)$ is the refractive index, and $\beta = 1/k_BT$ is the inverse temperature. In our simulations, the dipole moment of a water molecule $\vec{\mu}_i$ is defined as $\vec{\mu}_i = e(6\vec{r}_O+\vec{r}_{D_1}+\vec{r}_{D_2}-2\sum_{l=1,4}\vec{r}_{W_l})$, where $\vec{r}_{D_1}$, $\vec{r}_{D_2}$, and $\vec{r}_O$ are the coordinates of the deuterium and oxygen atoms, respectively, and $\vec{r}_{W_l}$ are the centers of the four maximally localized Wannier

functions (MLWC) associated with each molecule [92, 94], which are evaluated at each MD step with the algorithm proposed in Ref. [95].

In order to identify the role of surface water molecules on the computed IR spectra, we define the top and the bottom ice bilayers of the slabs as surface regions while the two interior ice bilayers are considered as the bulk ice region. Since the volumes of surface and bulk regions are not well defined, instead of computing $\alpha(\omega)n(\omega)$ as in Eq. 3.1.2, we computed the quantity $\alpha(\omega)n(\omega) \times \frac{3cV}{2\pi\beta}$. This choice is not expected to change the relative size and width of the peaks and their locations in the IR spectra.

3.1.3 Results and Discussion

In this section we present the IR spectra of the bulk and surface regions of our slab models, and we identify the types of water molecules associated with specific features of the computed IR spectra.

IR spectra of bulk ice regions

We first analyze the IR spectra of the two interior ice bilayers, i.e., the bulk region, obtained by considering all D_2O molecules of the two interior ice bilayers. In Fig. 3.1.2 we show the IR spectra calculated for the bulk regions of our two slab models. Since these regions are identical in both cases, the computed IR spectra of the two models exhibit the same line shape, relative intensity and positions of the band maxima. In addition, our results overall reproduce the main features of the bulk D_2O ice spectrum that was reported in Ref. [91] and obtained using the PBE functional and Car-Parrinello molecular dynamics. The positions of the band maxima corresponding to hindered translations, librations, bending, and oxygen-deuterium (OD) stretching modes are 210 (222), 715 (640), 1190 (1210), and 2190 (2425) cm^{-1}, respectively, where wavenumbers in parentheses are the corresponding experimental values [96].

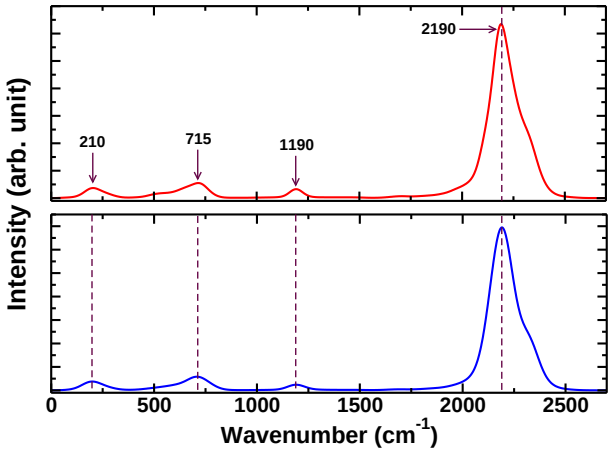

Figure 3.1.2: Calculated IR spectra of bulk regions of the two slab models with proton disordered surfaces (upper panel) and with proton ordered surfaces (lower panel). The sum in the dipole-dipole correlation function is restricted to the D_2O molecules from the two interior ice bilayers. The positions of the band maxima corresponding to hindered translations, librations, bending, and OD stretching modes are 210, 715, 1190, and 2190 cm^{-1}, respectively.

In agreement with Ref. [97], we found that the use of BO-AIMD yields a stretching band maximum (2190 cm^{-1}) in slightly better agreement with experiments than Car-Parrinello approaches (2120 cm^{-1}) [91]. Nevertheless the computed spectrum is still red shifted by $\sim$ 230 cm^{-1} with respect to the experimental value. This discrepancy is mainly due to the use of the approximate PBE functional. In particular, based on calculations carried out for the water monomer, the PBE error on stretching frequencies was estimated to be of the order of $\sim$ 150 cm^{-1} [91, 98]. The system size considered in our calculations is another source of error; it was shown in Ref. [99] that the spread of dipoles is around 0.2 Debye in the crystal interior for a 4-bilayer slab model, an order of magnitude larger than the value of 0.03 Debye obtained in the bulk simulation. This value is reduced to 0.1 Debye when using a 6-bilayer slab model. This overestimate of the variation of the dipole in bulk ice is likely to contribute to the red shift of our computed IR spectra with respect to experiment. The error originated from the use of small thickness slabs was estimated by computing vibrational frequencies of 4-bilayer and 6-bilayer slab models based on density

functional perturbation theory (DFPT). [35] We found that in the OD stretching region, all vibrational frequencies computed using a 4-bilayer slab model are systematically red shifted by $\sim$ 20 cm^{-1} compared to those obtained with a 6-bilayer slab. Therefore we conclude that the error due to the use of a 4-bilayer slab model is of the order of 20 cm^{-1}. Additional inaccuracies in the calculated spectra come from the use of a finite basis set and of pseudopotentials, as discussed in the study of D_2O liquid water [98, 100]. Such inaccuracies are estimated to be of the order of 50 cm^{-1} [98, 100]. Finally we note that the use of hybrid functionals, i.e., the PBE0 functional [100] or van der Waals density functionals [101] greatly reduces the error in the computed OD stretching band. However for reasons of computational efficiency, these functional were not considered in the present study, whose main goal is to compare the spectra of different surface structures and to focus on trends in IR spectra, and not on absolute values of peak positions.

IR spectra of the ice surface and subsurface regions

In this section we discuss the computed IR spectra of our ice surface models. We first analyze the spectra of a proton disordered surface, i.e., model B with an order parameter $C_{OD} = 3.0$ (see Fig. 3.1.1). In this model (see Fig. 3.1.3), the outer side of the surface ice bilayer consists of three-fold coordinated water molecules; each of these molecules has either a dangling deuterium atom (d-D molecule) that contributes a dangling OD bond to the surface, or a dangling oxygen atom (d-O molecule). At the interior side of the surface ice bilayer there are four-fold coordinated water molecules, that form hydrogen bonds to either the four-fold coordinated water molecules from the subsurface ice bilayer or three-fold coordinated water molecules.

To distinguish the role of surface and bulk water molecules on the computed spectra, we separated the contributions to the surface spectra in two parts, by restricting the sum in the dipole-dipole correlation function (Eq. 3.1.2) to: (i) D_2O molecules from the surface

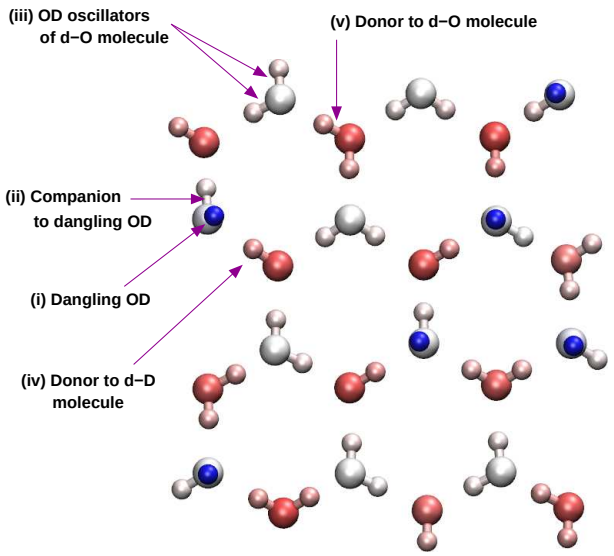

Figure 3.1.3: Surface ice bilayer of ice Ih: oxygen atoms belonging to the three-fold coordinated molecules in the outermost surface layer are represented in white; oxygen atoms belonging to the four-fold coordinated molecules in the interior of the ice bilayer are represented in red; dangling deuterium atoms are represented as blue spheres. Hydrogen bond configurations of the surface ice bilayer: (i) dangling OD or dangling-OD bond of a d-D water molecule; (ii) OD companion to the dangling OD of a d-D water molecule; (iii) OD oscillators of a d-O water molecule; (iv) OD of a water molecule hydrogen bonded to a d-D water molecule, and (v) OD of a water molecule hydrogen bonded to a d-O water molecule.

region, i.e., the surface ice bilayer; and (ii) D_2O molecules from the surface region and the first ice bilayer below the surface region (subsurface ice bilayer).

In Fig. 3.1.4 we show the spectra (black line) of the region contributed by the D_2O molecules from the surface ice bilayer. While the position of the band maximum corresponding to the bending mode (1190 cm^{-1}) is identical to that of the bulk region, variations up to 60 and 95 wavenumbers are observed for the hindered translations (270 cm^{-1} compared to 210 cm^{-1}) and librations (620 cm^{-1} compared to 715 cm^{-1}) modes, respectively. The most significant features of the IR spectra of the surface are observed in the high frequency region, which is characterized by the presence of three peaks: an intense one at $\sim$ 2330 cm^{-1}, a broad peak around $\sim$ 2000 cm^{-1} and a peak at 2666 cm^{-1}.

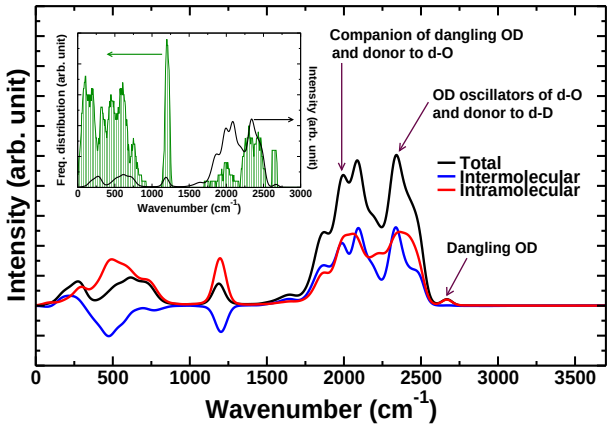

Figure 3.1.4: Calculated IR spectra (black line) of the surface region in a proton disordered model, model B (see Fig. 3.1.1). The sum in the dipole-dipole correlation function is restricted to D_2O molecules from the surface ice bilayer. Blue and red lines represent the inter- and intramolecular contribution to the IR spectra [91]. The peak at 2666 cm^{-1} is associated with the dangling OD group. The donors OD to d-D water molecules and OD oscillators of d-O molecules give rise to the peak at 2330 cm^{-1}. A band around 2000 cm^{-1} is associated with the donors OD to d-O molecules and the companions to the dangling OD groups. Inset: distribution of vibrational frequencies of water molecules in the surface region obtained from density functional perturbation theory calculations (green, to the left), and the corresponding IR spectra (black, to the right).

At variance, the IR spectra of the bulk has one main peak at 2198 cm^{-1}, corresponding to the OD stretch.

To identify the the vibrational modes associated with the band maxima of the surface spectra, we collected a snapshot along the MD trajectory and we extracted the geometry of the surface ice bilayer from this snapshot then the vibrational density of states of the chosen surface ice bilayer was computed based on DFPT. The inset of Fig. 3.1.4 shows the vibrational density of states (green line), which exhibits peaks in correspondence with those found in the IR spectrum (black line). Our analysis based on DFPT calculations showed that:

- The band at 2666 cm^{-1} is associated with the dangling OD groups of water molecules with a d-D.

- The band at ~ 2330 cm^{-1} originates from OD vibrations associated with three-fold coordinated molecules with a d-O, and by the donors OD of four-fold coordinated water molecules that are hydrogen bonded to molecules with a d-D.

- The band around 2000 cm^{-1} is associated with the companions to the dangling OD groups and the donors OD of fourfold-coordinate water molecules, hydrogen bonded to molecules with a d-O.

In Fig. 3.1.4 we also show the inter- (blue line) and intramolecular (red line) contribution to the IR spectra. The intramolecular contribution corresponds to retaining only the terms with $i = j$ in Eq. 3.1.2 and the remaining terms define the intermolecular contribution [91]. It is seen that the sharp peak at 2666 cm^{-1} is dominated by the intramolecular component, clearly indicating that this band maximum stems from the dangling OD groups that are present only on the surface [82].

Our results on the types of water molecule associated with specific features of the computed IR spectra are in agreement with those of the vibrational analysis reported in Refs. [102, 103] where an interatomic potential model was employed. We note that a combined experimental and theoretical study of the IR spectra of water clusters [104] drew similar conclusions. Similar conclusions were reached in Refs. [105, 106] for water clusters when using interatomic potential models and in Ref. [107] using DFT calculations and the B3LYP functional.

Next we consider the role of the subsurface bilayer in the IR spectra. In Fig. 3.1.5 we show the spectrum (black line) evaluated by restricting the sum in the dipole-dipole correlation function to the D_2O molecules of the surface and subsurface ice bilayers. The contribution to the IR spectrum from the subsurface bilayer is also extracted by including only D_2O molecules of the subsurface bilayer (red line). When the contribution of the sub-surface bilayer are added to those of the surface layer, one observes the appearance of a peak at 2190 cm^{-1} in the spectrum, characteristic of the bulk ice (see Fig. 3.1.2).

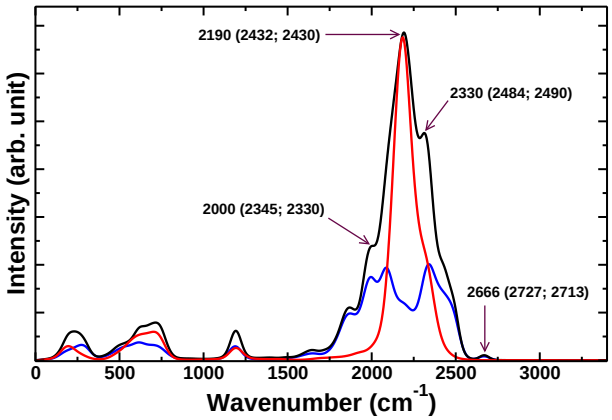

Figure 3.1.5: Calculated IR spectra of a proton disordered ice surface, model B (see Fig. 3.1.1). The sum in the dipole-dipole correlation function is restricted to the D_2O molecules from:(i) the surface ice bilayer (blue line); (ii) the subsurface ice bilayer (red line); and (iii) the surface and the subsurface bilayers (black line). Wavenumbers in parentheses are experimental values for D_2O ice films [108, 109, 110, 111] and D_2O clusters [112, 113].

This peak originates from four-fold coordinated D_2O molecules.

Effect of proton disordering on IR spectra

We now turn to the comparison of IR spectra obtained for different structural models. We focus on the contribution of the surface bilayer for the three surface models presented in Fig. 3.1.1.

In Fig. 3.1.6 we show the computed IR spectra of these three surfaces: (i) model A, a proton ordered surface, with $C_{OD} = 2.0$; (ii) model B and model C, proton disordered surfaces, with $C_{OD} = 3.0$. The results obtained for the three surface models exhibit similar features in the whole range of frequency, with only slight differences in the positions of the band maxima. As pointed out in the previous section, each feature of the computed IR spectra is associated with a specific type of bonding configuration on the surface. The similarities of the IR spectra computed for different surface models shown in Fig. 3.1.6 are likely due to the fact that all surface models share the same types of bonding config-

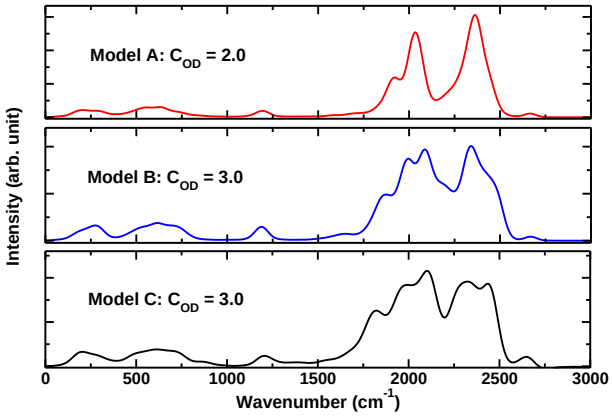

Figure 3.1.6: Calculated IR spectra of the different surfaces of ice with different order parameters: model A with $C_{OD} = 2.0$, model B and model C with $C_{OD} = 3.0$ (see Fig. 3.1.1).

urations; and therefore we expect that expanding the surface area of the simulation cell is unlikely to yield quantitatively different IR spectrum.

Our findings about the insensitivity of IR spectra to the details of the surface structure are consistent with experiments, showing very similar spectra for D_2O ice thin films [108, 109, 110, 111] and ice clusters [112, 113]. In addition, a sum frequency generation study[114] revealed that vibrational modes of the basal and prism ice surfaces are similar, with differences of a few wavenumbers. Given the accuracy of theory used here, it is not possible to establish how robust the small difference in the spectra of structural models A, B and C are. Therefore based on our results it is not yet possible to comment on the findings of Ref. [84] and Ref. [85] on the stability of the proton ordered surface.

We further addressed the effect of temperature on the computed IR spectra and carried out simulations at 150 K, with a protocol similar to the one used at 100 K. The proton ordered surface slab was used as input, and simulations were carried out in the NVE ensemble, and trajectories to compute IR spectra were collected over $\sim$ 10 ps after equilibrating the system at 150 K for $\sim$ 5 ps. As shown in Fig. 3.1.7 and Fig. 3.1.8 the computed IR spectra for these two temperatures are similar in line shape, relative inten-

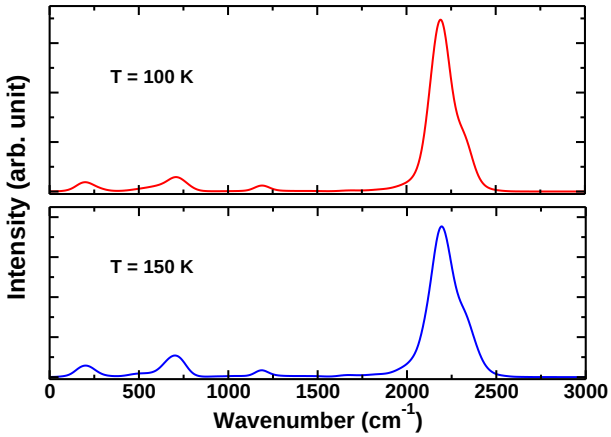

Figure 3.1.7: Calculated IR spectra of the bulk region of the proton order surface slab for temperatures of 100 K (upper panel) and 150 K (lower panel).

sity, and they exhibit the same positions of band maxima. This observation indicates that the surface of ice Ih is likely to retain its crystal structure for T below 150 K. Our results are in agreement with He scattering and SFG experiments where the loss of oxygen atomic order is observed at 180-200 K [81, 82]. and with calculations that found even a higher transition temperature than experiments [87, 115].

Comparison to experiment

In this section we present a comparison of our computed spectra with available experiments for D_2O ice films [108, 109, 110, 111] and D_2O ice clusters [112, 113].

Ultrathin deuterium ice films on Pt(111) have been studied by Fourier transform infrared reflection absorption spectroscopy [108, 109, 110, 111]. The IR spectra of the crystalline deuterium ice films obtained at 150-155 K exhibits a sharp peak at 2727 cm^{-1} due to surface dangling OD groups and three distinct features at 2484, 2432 and 2345 cm^{-1} in the OD stretching region. Similarly, Fourier transform infrared spectroscopy experiments of large D_2O clusters [112, 113] at 100 K show prominent features at 2713, 2490, 2430, and 2330 cm^{-1}.

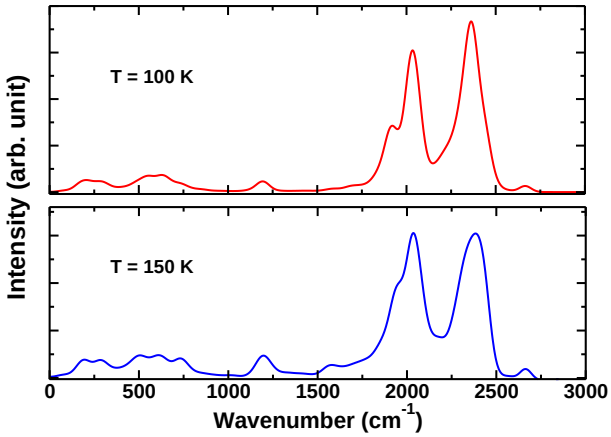

Figure 3.1.8: Calculated IR spectra of the proton ordered surface (Fletcher phase) for temperatures of 100 K (upper panel) and 150 K (lower panel).

Our computed IR spectra of surface and subsurface bilayers (black line) shown in Fig. 3.1.5 reproduce well all the features observed in the experimental IR spectra. The four prominent peaks at 2666, 2330, 2190 and at 2000 cm^{-1} in computed spectra are associated with the stretching mode of the dangling OD group; the OD oscillators of d-O water molecules and the donors OD of four-fold coordinated water molecules hydrogen bonded to d-D molecules; the OD oscillators of four-fold coordinated water molecules from the subsurface ice bilayer; the companion to the dangling OD group and the donor OD of fourfold-coordinated molecules hydrogen bonded to d-O water molecules, respectively. While we qualitatively reproduce all features of measured spectra, we find that the band maxima corresponding to the computed OD stretching mode contributed by the four-fold coordinated molecules from the subsurface bilayer (2190 cm^{-1}) is red shifted by $\sim$ 240 cm^{-1} compared to experiment on ice clusters (2430 cm^{-1}) [112, 113] and ice thin films (2432 cm^{-1}) [108, 109, 110, 111]. Furthermore, all the three band maxima in the computed spectra show a systematic red shift. The red shift with respect to experimental measurements can be explained mainly as a result of the use of the approximate PBE functional, similar to what observed in bulk ice regions.

3.1.4 Summary

In summary, we have computed the IR spectra of different ice surface models using first-principles molecular dynamics simulations and semi-local gradient corrected functionals (PBE). Our results reproduce well the features obtained in experiments for ice thin films and ice clusters. We have provided a detailed interpretation of measured spectra and we have identified the types of D_2O molecules contributing to the various features of the measured spectrum. We found that ice surfaces with different order parameters share similar IR spectra. Although recent studies showed that the ordered ice surface is energetic favorable, it is not yet possible to conclude if the ice surface is proton-ordered from the computed IR spectra, within the accuracy of our methodology.

3.2 Al_2O_3/water interface

3.2.1 Introduction

Understanding the chemical environment at the interface between a solid oxide and liquid water is important for a broad range of applications in natural sciences and engineering [116, 117, 118]. While there is general agreement that the behavior of interfacial water differs from that of bulk water, the relationship between the oxide surface structure and reactivity, and the degree of water ordering remain open questions. Unfortunately, no experimental tools are yet available which can directly probe the microscopic structure of solid/liquid interfaces. Few techniques can access buried interfaces with sufficient surface selectivity to achieve atomic-scale resolution, although much progress has been made in recent years. [119] These techniques include synchrotron X-ray reflectivity and scattering, which have yielded valuable information on the oxide surface termination and the first layer of adsorbed water for several oxide/water interfaces [120]. However X-ray techniques are insensitive to light nuclei such as hydrogen, and thus do not provide direct

probes of hydrogen bonding in water. Interface-selective vibrational spectroscopy based on sum-frequency generation (SFG) is instead better suited to probe hydrogen bonds and it has played an important role in probing the dynamics of protons of interfacial water molecules [121], whose vibrational motion may be directly correlated with hydrogen bond strength. However, the vibrational spectra of liquid water, both in the bulk and at interfaces, are composed of multiple broad, overlapping bands, rendering their interpretation difficult.

Atomic-scale, first-principles simulations have the potential to provide a detailed picture of water ordering and hydrogen bonding at oxide/water interfaces. In such simulations, interatomic forces are usually obtained from density functional theory (DFT) within local or semilocal approximations for the exchange-correlation energy functional. While the use of hybrid functionals for electronic structure calculations of condensed phases is an area of active research, to date only few first-principles molecular dynamics (MD) simulations with hybrid functionals have been applied to aqueous systems composed of several tens of atoms [122, 123, 124]. In the case of bulk liquid water, extensive comparisons between experiments and first-principles simulations have been made, and the errors introduced by semilocal approximations within DFT, as well as that of neglecting proton quantum effects [125] are relatively well understood [126, 127].

Here we focus on the alumina (α-Al_2O_3) (0001) surface in contact with liquid water, which we investigated using first-principles MD and the semilocal exchange-correlation energy functional of Perdew, Burke, and Ernzerhof (PBE) [128, 129]. The large supercells used here (with 1632 and 2016 electrons) prevented us from using hybrid functionals, and thus our results were interpreted taking into account the known drawbacks and errors of PBE for bulk liquid water. Experimentally, the alumina/water interface has been studied with both X-ray and SFG techniques, thus representing a suitable prototype interface for extensive comparisons between theory and experiments. Early studies employed syn-

chrotron X-ray reflectivity and scattering to probe the alumina(0001) surface in contact with water vapor [130], yielding information on the alumina(0001) surface structure and adsorbed water molecules. More recent experiments have examined a bulk-like layer of water [131]. The various X-ray measurements [130, 131] generally agree on the alumina(0001) surface termination and relaxation, reporting a single layer of adsorbed water molecules at a distance of about 2.3–2.5 Å above the alumina terminal oxygen layer. In contrast, similar experiments for water on the (012) and (110) surfaces of alumina found two hydration layers [132, 133], reflecting the more corrugated nature of these surfaces as compared to the (0001) surface.

Further information on interfacial water ordering and hydrogen bonding at the alumina(0001) interface was provided by SFG vibrational spectroscopy experiments [134, 135, 136]. Of notable interest is the presence of two broad bands in the OH stretching region of the spectrum: i) a $\sim$ 3200 cm^{-1} peak which is referred to as "ice-like" due to its proximity to the OH stretching peak in bulk ice Ih, and ii) a "liquid-like" peak at $\sim$ 3400 cm^{-1} whose position coincides with the bulk OH stretching mode in liquid water. These features have been observed in a variety of solid/water and water/vapor interfaces; in the case of alumina/water, their relative intensities also exhibit a pH dependence [137, 138].

The (0001) surface of alumina has been the subject of extensive theoretical studies employing first-principles techniques. Most of these investigations have focused on the clean surface under ultra-high vacuum conditions, where significant relaxation of the outer surface layers is known to occur [139]. The molecular mechanisms for the initial hydroxylation steps of the clean surface have also been elucidated using first-principles simulations at sub-monolayer water coverages [140, 141]. In addition, alumina/water interface models with 32 water molecules representing the fully hydroxylated (0001) interface have recently been employed to study the interfacial water structure and its relation to the computed

vibrational density of states (VDOS) [142].

In this work, we investigated the ideal hydroxylated alumina(0001) surface [143] in contact with a large (96 and 144 waters) liquid water layer and we studied both structural and vibrational properties of the interface. We show that the first layer of adsorbed water molecules is at a distance of about 2.5 Å above the alumina surface, in agreement with X-ray experiments [130, 131], and that the interfacial water region is localized to within 5 Å from the surface. As discussed further below, our findings for the first water layer structure also agree with results of previous MD simulations employing classical potentials [144]. We obtained the infrared (IR) spectrum of water at the interface using maximally localized Wannier functions to compute molecular dipole moments. Although the experimental SFG vibrational spectra contains both IR and Raman contributions, we could relate the main features of the simulated IR spectrum to those of the experimental SFG spectra, finding a good, qualitative agreement. Our results support the interpretation that the "ice-like" and "liquid-like" peaks originate from distinct types of interfacial water molecules with different strengths of hydrogen bonds at the interface, as concluded in Ref. 142 based on analysis of the VDOS computed from first-principles MD simulations.

3.2.2 Methods

All first-principles simulations were carried out with interatomic forces derived from DFT using the PBE approximation for the exchange-correlation energy functional [128, 129]. The interaction between valence electrons and ionic cores were represented using norm-conserving pseudopotentials (with one valence electron for deuterium, six for oxygen, and three for aluminum), and the electronic wavefunctions were expanded in a plane-wave basis set truncated at a cutoff energy of 85 Ry.

Static calculations for bulk alumina and the hydroxylated alumina(0001) slab were

carried out with the QUANTUM ESPRESSO package [145]. First-principles MD simulations of the hydroxylated alumina(0001)/water interface were performed using the Qbox code [89]. At each ionic step, the total energy of the sample was converged to 10^{-7} eV/atom and the centers of the maximally-localized Wannier wavefunctions were computed on-the-fly [146]. The Brillouin zone of the supercell was sampled using only the Γ point. Hydrogen atoms were replaced with deuterium in order to maximize the allowable MD time step; we adopted a time step of 10 a.u. = 0.242 fs. The alumina/water interfaces were initially prepared and equilibrated for $>$ 200 ps using classical MD with the CLAYFF force field [147], followed by another $\sim$ 7 ps of equilibration with first-principles MD. The first-principles equilibration runs were carried out under constant temperature (NVT) conditions at an average temperature of $T = 400$ K, maintained by a Berendsen thermostat with a coupling time of 0.1 ps [148]. The neglect of quantum zero-point effects [125] from light nuclei and the use of the PBE approximation [126, 127] have been shown to yield an overstructured liquid water, and an elevated simulation temperature of $T \sim 400$ K is necessary to recover the experimental water structure and diffusion coefficients at $T \sim 300$ K [126, 127]. Ensemble averages were computed from first-principles trajectories obtained under constant energy (NVE) conditions over $\sim$ 9 ps.

We computed the infrared absorption coefficient per unit path length:

$$\alpha(\omega)n(\omega) = \frac{2\pi\omega^2\beta}{3cV}\int dt\, e^{-i\omega t}\left\langle \mathbf{M}(0)\cdot\mathbf{M}(t)\right\rangle, \tag{3.2.1}$$

as describted in Ref. 149; $n(\omega)$ is the refractive index, c is the speed of light, $\beta = 1/k_BT$ is the inverse temperature, and $\mathbf{M}$ is the instantaneous total dipole moment. For a given MD snapshot, $\mathbf{M}$ is computed as a vector sum over all charges of the sample of volume V,

$$\mathbf{M} = \sum \mathbf{r}_{\mathrm{D}} + 6\sum \mathbf{r}_{\mathrm{O}} + 3\sum \mathbf{r}_{\mathrm{Al}} - 2\sum \mathbf{r}_{\mathrm{w}}, \tag{3.2.2}$$

where $\mathbf{r}_{\mathrm{D}}$, $\mathbf{r}_{\mathrm{O}}$, $\mathbf{r}_{\mathrm{Al}}$ are the coordinates of the D, O, Al nuclei, respectively, and $\mathbf{r}_{\mathrm{w}}$ are the positions of the Wannier centers. In water, Wannier centers correspond to the center of charge of Boys orbitals [150], *i.e.*, to the positions of bonds and lone pairs. By restricting the summation in Eq. 3.2.2 over charges associated with specific molecules and surface functional groups, it is possible to isolate individual contributions to the overall spectrum.

3.2.3 Interface model

The bulk structure of α-Al_2O_3 may be described using either a primitive, 10-atom rhombohedral unit cell or a 30-atom hexagonal cell. We carried out the optimization of the bulk structure in the rhombohedral cell and we sampled the Brillouin zone with a symmetrized 7×7×7 k-point mesh. The optimized rhombohedral lattice parameters are $a_R = 5.147$ Å and angle $\alpha_R = 55.31°$, which correspond to hexagonal parameters $a_H = 2a_R \sin(\alpha_R/2) = 4.777$ Å and $c_H = \sqrt{9a_R^2 - 3a_H^2} = 13.04$ Å. These are slightly larger than the corresponding experimental values of $a_H = 4.760$ Å and $c_H = 12.989$ Å [151]. The small overestimation of the equilibrium volume found in our calculations is attributed to the use of the PBE functional, and similar errors were reported for a variety of other covalently bonded and ionic systems.

To build the 2D-periodic alumina(0001) slab, we chose an orthorhombic unit cell with lateral dimensions $a = 3a_H = 14.33$ Å and $b = 2\sqrt{3}\,a_H = 16.55$ Å. Along the [0001]-direction, the layers are arranged in a repeating sequence of stoichiometric units consisting of an oxygen layer sandwiched between two aluminum layers, –Al–O_3–Al–. It is known from experiment [139] that the lowest energy termination of the alumina(0001) surface under ultra-high vacuum conditions is the non-polar Al-termination, which exhibits a large inward relaxation of the outermost Al atoms. In the presence of reactions with water, the surface structure and composition is expected to be significantly modified. We adopted an ideal hydroxylated alumina(0001) model with a non-polar termination

appropriate for high water coverages [143]. Starting with a bulk-terminated slab consisting of three stoichiometric layers, Al–O_3–Al–Al–O_3–Al–Al–O_3–Al, each outermost Al atom was replaced with three deuterium atoms to form an OD-terminated slab: $(OD)_3$–Al–Al–O_3–Al–Al–$(OD)_3$. In this model, every surface OD group bridges two subsurface Al atoms. Such a three oxygen layer, hydroxylated slab was also previously employed in Car-Parrinello MD simulations of the alumina(0001) surface in Refs. 140 and 141. As discussed further in Section 3.2.4, the relaxations observed in our simulations are in reasonable agreement with the structure inferred from synchrotron X-ray experiments for the hydrated surface [130].

We studied two alumina/water interface models, consisting of alumina and water layers alternating along the c-axis: a 96 D_2O model with $c = 20.365$ Å, and a 144 D_2O model with $c = 26.953$ Å. These values for c were chosen so as to obtain a water sample with the D_2O experimental density of ~ 1.1 g/cm^3.

3.2.4 Results and discussion

Structure

The computed number densities as a function of distance z from the surface oxygens are shown in Fig. 3.2.1. The 96 and 144 water models yield densities that are essentially indistinguishable. We found only minor relaxations of the alumina oxygen layers at the interface, relative to the bulk-terminated slab. In the interface model, the average distance between alumina oxygen layers is 2.18 Å, close to the corresponding value of 2.173 Å in the optimized bulk structure. On the other hand, we observed a larger relaxation of the subsurface Al layers: the average distances between adjacent Al layers in the interface model and in the bulk are 0.27 Å versus 0.497 Å, respectively; this corresponds to a 46% contraction at the interface relative to the bulk. Our results are consistent with those of synchrotron X-ray measurements of the hydrated alumina(0001) surface at $T = 300$ K by

Eng *et al.*, who reported a negligible relaxation of the alumina oxygen layers and a larger (53%) contraction between the two subsurface Al layers [130].

The peaks in the O and D atom densities at $z = 0 - 1$ Å are clearly identifiable with the OD groups on the alumina surface. The peaks at $z = 0$ and $z = 0.9$ Å in the D atom density arise from surface OD groups with two distinct orientations: i) approximately parallel to the slab and forming in-plane hydrogen bonds with nearest neighbor surface oxygens, and ii) pointing out of the slab plane and towards the liquid water region. The distribution of surface OD angles relative to the direction perpendicular to the surface is shown in Fig. 3.2.2. The peak near 90° corresponds to in-plane OD bonds, while the one at 20° to out-of-plane OD. These two types of surface OD bonds have been observed in previous first-principles simulations of the alumina(0001) surface [140, 141, 142], and their presence was also inferred from the SFG vibrational spectra [136]. These bonds are expected to exhibit different chemical reactivity and as we show below, they are responsible for different features in the vibrational spectra: indeed, the in-plane ODs are well positioned to accept hydrogen bonds, while the out-of-plane ones can more easily donate hydrogen bonds.

The features of the O and D atom distributions seen in the region $z = 1 - 4$ Å are due to the first water layer in contact with the surface. This layer is centered around a O atom peak at $z = 2.6$ Å and is sandwiched between two well-separated D atom peaks at 1.7 Å and 3.1 Å. Its structure is qualitatively different from that inferred from density profiles of, *e.g.*, the hydrophobic graphene/water interface [152]. In graphene/water, the first water layer O and D atom density maxima coincide due to in-plane hydrogen bonding between water molecules of the first layer. In contrast, in the alumina/water interface, molecules of the first water layer form hydrogen bonds with the alumina surface hydroxyl groups. Our observations for the structure of the first water layer are very similar to findings from previous alumina(0001)/water simulations employing classical potentials [144].

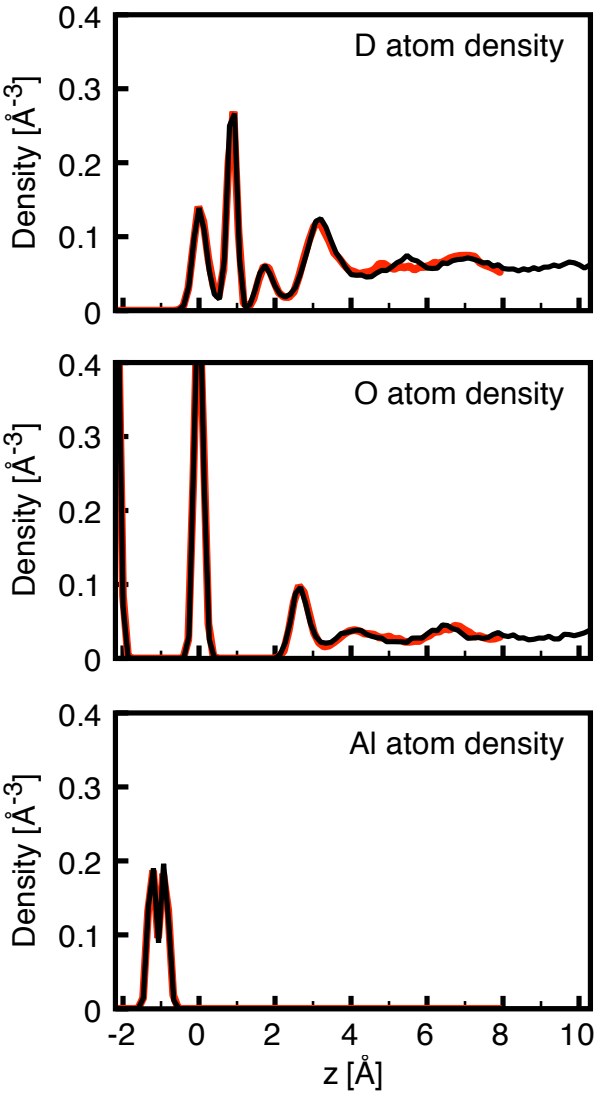

Figure 3.2.1: D atom (top), O atom (middle), and Al atom (bottom) number densities as a function of distance z from the outermost oxygens of the alumina surface. Black (red) lines are densities obtained from the alumina/144 (96)-water interface model.

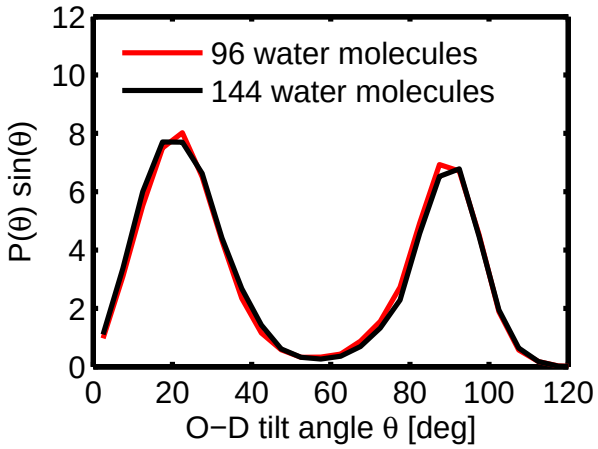

Figure 3.2.2: Distribution of surface hydroxyl (OD) tilt angles relative to the direction perpendicular to surface. The distribution is normalized such that $\int P(\theta)\sin(\theta)d\theta = N_{\mathrm{OD}}$, where N_{OD} is the number of surface hydroxyl groups in the simulation cell. The black (red) line denotes results obtained with the alumina/144 (96)-water model. The peak at 20° is due to out-of-plane surface OD bonds, while that at ∼90° is due to in-plane, hydrogen-bonded surface OD bonds.

Hydrogen bonding with the substrate was also seen in previous theoretical studies of water on hydrophilic surfaces such as classical MD simulations of silica/water [153] and first-principles MD simulations of SiC(100)/water [154]. Finally, our results show that the density resumes bulk-like values for z larger than 5.0 Å.

Infrared spectra

The IR spectra were computed using the procedure described in Section 3.2.2. Given the excellent agreement between the interfacial structural properties obtained with 96 and 144 water samples, we only computed the IR spectrum for the smaller alumina/96-water model. In all the following analyses, the statistical uncertainty on the peak positions of the computed IR spectra was estimated to be 25 cm^{-1}, based on previous results presented in Ref. [149].

Fig. 3.2.3 shows the total spectrum (black), together with contributions originating from the 96 D_2O molecules (blue), and from the surface OD bonds (red). The peak

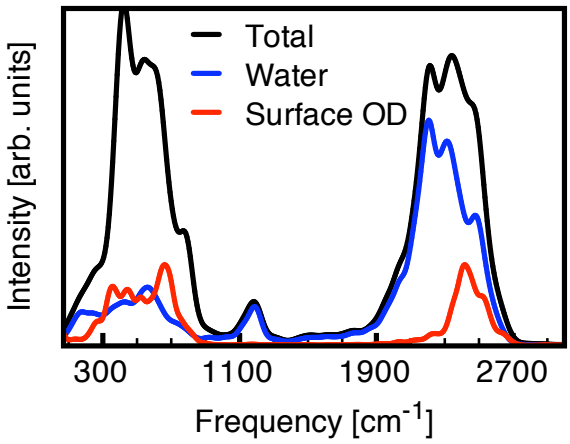

Figure 3.2.3: Computed IR spectrum (black line) of the alumina/water interface (96 water model), together with contributions from liquid water molecules (blue line) and surface hydroxyl (OD) groups (red line).

at 1185 cm^{-1} corresponds to the water bending mode, and it is present in the total signal and in that of water, but not in the OD surface contribution. Its position is in reasonable agreement with that of the corresponding bulk water value obtained from first-principles MD [149] (1173 cm^{-1}) and experiment [155] (1209 cm^{-1}). The broad band at low frequencies (< 1000 cm^{-1}) is due to hindered translational and librational motion in water and lattice vibrational motion in the alumina slab, whose detailed analysis is outside the scope of this work. Here, we focus on the interpretation of the OD stretching band centered at 1800–2800 cm^{-1}.

A detailed view of the surface OD spectrum over the 1800–2800 cm^{-1} range is shown in Fig. 3.2.4, and it is decomposed into contributions from in-plane (blue) and out-of-plane (red) surface OD bonds. The out-of-plane OD band exhibits two distinct features: a higher-frequency peak at 2630 cm^{-1} that is readily assigned to dangling, non-hydrogen-bonded OD, and a lower-frequency peak at 2453 cm^{-1} from surface OD that are hydrogen-bonded to water. The IR spectrum of the in-plane species shows a strong peak at 2410 cm^{-1} that is red-shifted from the dangling OD peak by 220 cm^{-1} due to

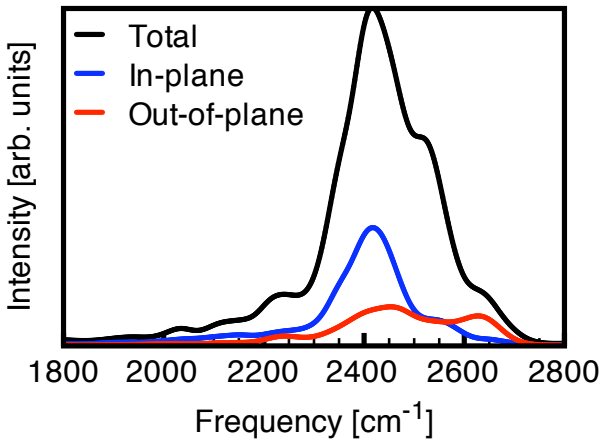

Figure 3.2.4: Computed IR spectrum of surface hydroxyl (OD) groups obtained with the alumina/96-water model. The black line represents the total OD contribution, the blue line the in-plane OD (tilt angles $> 65°$ relative to surface normal), and the red line the out-of-plane OD (tilt angles $< 25°$ relative to surface normal) contributions, respectively. In the out-of-plane OD spectrum, the 2453 cm^{-1} peak is due to surface OD groups that are hydrogen-bonded to water, and the 2630 cm^{-1} peak corresponds to a dangling surface OD. The in-plane OD peak at 2410 cm^{-1} is red-shifted from the dangling OD peak by 220 cm^{-1}, due to hydrogen bonding with neighboring surface oxygen atoms.

hydrogen bonding with neighboring surface oxygens.

Further insight into the nature of interfacial water ordering and bonding may be achieved by analyzing the water vibrational dynamics. We decomposed the IR spectrum due to the 96 D_2O molecules of the alumina/96-water interface model into contributions originating from an interfacial water region immediately adjacent to the surface, defined as $z < 5$ Å, and an interior water region at $z > 5$ Å. The interior and interface water spectra are shown in the upper panels of Fig. 3.2.5 and Fig. 3.2.6, respectively. These are compared to previously reported bulk liquid D_2O spectrum [149] (lower panel of Fig. 3.2.5) and bulk ice Ih spectrum [156] (lower panel of Fig. 3.2.6), both derived from first-principles MD simulations within the PBE approximation. The bulk water simulations of Ref. 149 were also carried out at the same elevated simulation temperature of $T \sim 400$ K chosen here for the alumina/water interface simulations, while the bulk ice simulations in Ref. 156

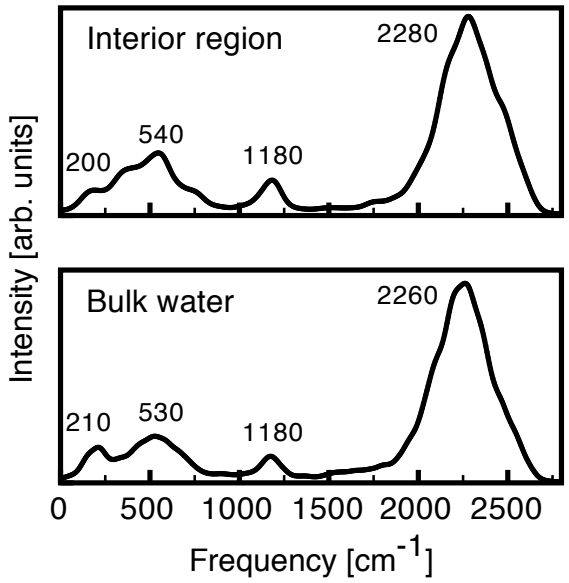

Figure 3.2.5: Computed IR spectrum (top panel) of the interior water region of the alumina/96-water interface model, defined as $z > 5$ Å. The computed bulk liquid D_2O spectrum from Ref. 149 is shown in the bottom panel for comparison.

were performed at $T \sim 100$ K. Note that the computed bulk water OD stretch band at 2260 cm^{-1} is red-shifted from the experimental value of 2498 cm^{-1}. This systematic red shift has been examined in detail in previous work [122] and can be ascribed to the use of the approximate PBE functional; nevertheless, the overall width and shape of the band are in satisfactory agreement with experiment [149].

The comparison of the interior region of the water layer with previously reported bulk simulations shows that both the number density and the IR spectrum are nearly bulk-like (see Fig. 3.2.5). The positions of the low-frequency librational and bending bands are in close agreement with those of the bulk liquid. The shape and width of the OD stretch band in the interior region also resemble the corresponding ones in liquid water, but its position at ~ 2280 cm^{-1} is slightly blue-shifted by about 20 cm^{-1} relative to the bulk. This small blue shift is likely due to confinement effects, possibly similar to the one observed in the case of hydrophobic surfaces. Previous first-principles MD studies of water

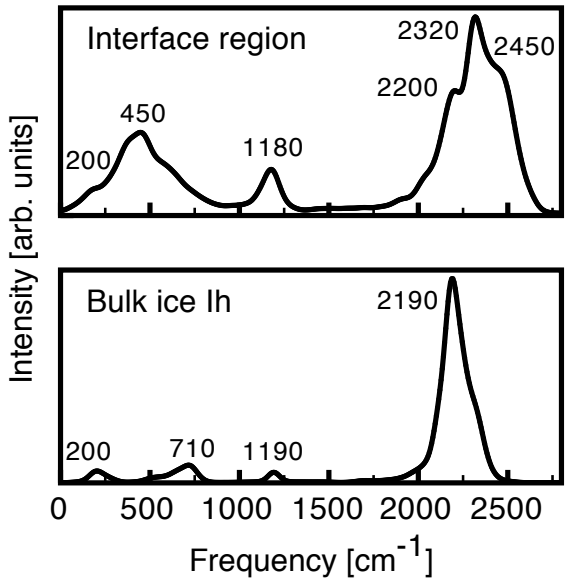

Figure 3.2.6: Computed IR spectrum (top panel) of interfacial water immediately adjacent to the alumina surface, defined as $z < 5$ Å. The computed bulk ice Ih D_2O spectrum from Ref. 156 is shown in the bottom panel for comparison.

confined between graphene sheets found that a separation of 25 Å is necessary to fully recover the IR spectrum of bulk water, and a systematic reduction in this confinement distance resulted in a blue shift in the position of the OD stretching band [157].

The OD stretching band due to interfacial water molecules (Fig. 3.2.6) exhibits additional features not present in bulk water or ice. Specifically, there are three distinct IR peaks arising from the interfacial region: 2200, 2320, and 2450 cm^{-1}. The first two are close to the OD stretch peak in bulk ice (2190 cm^{-1}) and bulk liquid water (2260 cm^{-1}), respectively, and are thus labelled as "ice-like" and "liquid-like". The origins of these features are identified in Fig. 3.2.8, where in panel (a) the interfacial water OD stretch band is shown again, and compared to the spectra arising from interfacial water molecules that donate a hydrogen bond to in-plane surface OD groups (panel (b)). The interfacial water feature at 2200 cm^{-1} is clearly due to these hydrogen-bond-donating interfacial water molecules; in addition, the close correspondence with the 2190 cm^{-1} peak in bulk

ice suggests a strong, ice-like hydrogen bond. Panel (c) shows the spectrum from interfacial molecules accepting a hydrogen bond from the out-of-plane surface OD groups, which give rise to the 2320 cm^{-1} peak in the interfacial water spectrum. Its position is close to that of the OD stretching peak in bulk liquid water, indicating a weaker, "liquid-like" coordination environment. The remaining interfacial water bonds not included in (b) or (c) are classified as broken hydrogen bonds, whose spectrum is shown in panel (d), and provides an explanation for the high-frequency 2450 cm^{-1} peak. A similar peak was also observed in previous simulations of the vibrational spectra of the SiO_2/water interface [158]. Note that we did not observe any dangling OD stretching peak arising from interfacial water molecules, which is indicative of the hydrophilic nature of the alumina surface. In contrast, Car-Parrinello MD simulations of the vibrational properties of the hydrophobic diamond(111)/water interface reported a small but distinct dangling OD feature in the computed IR spectrum at $\sim$ 2500 cm^{-1} [157]. Also note that Car-Parrinello dynamics produces a systematic red shift of $\sim$ 100 cm^{-1} in the infrared OD stretching band, relative to the Born-Oppenheimer dynamics as carried out in this work; had this feature been present in our simulations, we would expect it to appear at $\sim$ 2600 cm^{-1}.

Further characterization of interfacial water molecules is shown in Fig. 3.2.7 which reports the dipole moment distribution in the direction normal to the Al_2O_3(0001) surface. The average dipole moment per molecule shows a sharp increase at the interface due to the strong interaction between interfacial water molecules and the solid. In contrast, in the case of the hydrophobic diamond(111)/water interface, the dipole moment of interfacial water molecules is decreased to a value lower than that in the bulk. This difference in the dipole distribution clearly reflects the different nature of the two surfaces.

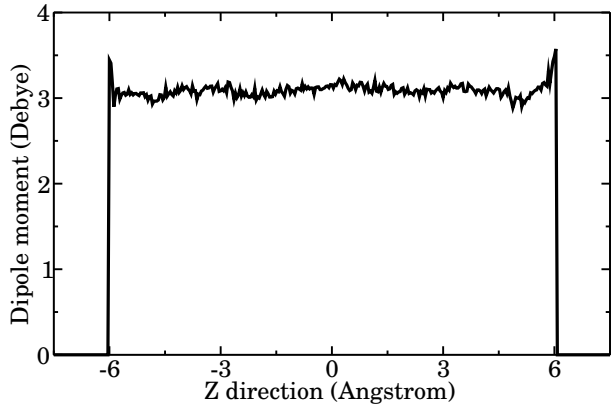

Figure 3.2.7: The dipole distribution of water molecules in the direction normal to the $Al_2O_3(0001)$ surface.

Discussion

At present only measured SFG spectra [134, 135, 136] are available for the single-crystal, water/alumina(0001) interface and no IR and Raman spectra have yet been reported. Unfortunately a direct, quantitative comparison between the experimental SFG spectra and our computed IR spectra is not straightforward. First, the SFG spectrum involves both IR and Raman contributions, and thus the relative intensities of the peaks of the SFG and IR spectra are expected to be different. In addition, as previously mentioned the use of the approximate PBE functional for the exchange-correlation energy leads to a sizable red shift in the computed water OD stretching band compared to experiments [149, 122], and the neglect of quantum effects from light nuclei (D) might further increase this error [159]. Deuterium masses are used here to enable the use of larger MD time steps, however the SFG experiments on alumina(0001)/water were performed on hydrogenated water samples.

However a qualitative comparison between experiments and simulation results can be drawn in the OH stretching region. The SFG spectrum is dominated by a strong peak at ~ 3700 cm^{-1}. Flörsheimer *et al.* [136] decomposed it into four subpeaks arising from

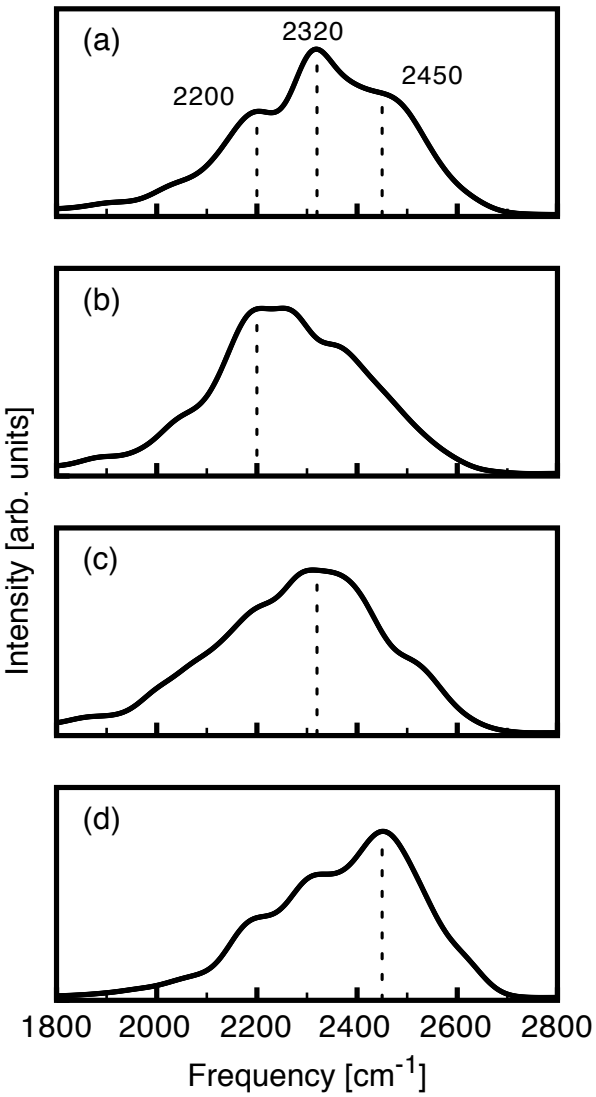

Figure 3.2.8: Panel (a): Computed IR spectrum of the interfacial water region immediately adjacent to the alumina surface, defined as $z < 5$ Å. The remaining panels show spectra arising from (b) water molecules that donate a hydrogen bond to in-plane surface OD groups, (c) water molecules that accept a hydrogen bond from out-of-plane surface OD groups, and (d) broken hydrogen-bonded waters.

alumina surface OH groups of increasing tilt angles: from the highest frequency subpeak at 3690 cm^{-1} assigned to dangling surface OH groups, to the lowest frequency subpeak at 3460 cm^{-1} due to in-plane, hydrogen-bonded surface OH. These deconvoluted peaks correspond to the out-of-plane and in-plane surface OD peaks shown in Fig. 3.2.4. Note that in the calculated IR spectrum, similar to the SFG spectrum, the in-plane surface OD stretch overlaps with a broad band due to vibrations involving out-of-plane surface OD groups that donate a hydrogen bond to interfacial water molecules.

At nearly neutral pH conditions, the SFG experiments also reported a weaker OH stretch band at ~ 3400 cm^{-1}, interpreted as arising from interfacial water molecules with "liquid-like" characteristics, due to its proximity to the bulk water OH stretching band maximum. At low and high pH a third band at ~ 3200 cm^{-1} appears, labelled as "ice-like" by correspondence with the OH stretching mode in bulk ice Ih. These general trends as a function of pH have also been reported for other water interfaces [121]. One proposed explanation in the case of alumina(0001)/water [135] is that the surface acquires a net charge due to protonation (low pH) or deprotonation (high pH) of the surface OH groups, when moving away from the point of zero charge (pzc). This pH-dependent surface charge, in turn, drives the orientation of the interfacial water molecule dipole moments.

Alternatively, Flörsheimer *et al.* [136] argued that in alumina(0001)/water, the orientation of the interfacial water molecules is not driven solely by electrostatics but also by hydrogen bonding with in-plane surface OH groups, whose population relative to the out-of-plane surface ones was observed to vary as a function of pH. Consequently, when in-plane surface OH groups are prevalent, more interfacial water molecules are engaged in strong hydrogen bond donation to these in-plane surface OH, giving rise to an increased intensity in the ice-like OH band. This ice-like band is reduced in intensity when out-of-plane OH groups become prevalent, as more interfacial water molecules experience a comparatively weaker, liquid-like hydrogen bonding.

The analysis of our computed IR spectra appears to support the interpretation of Ref. 136. As discussed in Section 3.2.4, the OD vibrations arising from the interfacial water molecules exhibit features at 2200 cm^{-1} and 2320 cm^{-1} that coincide with the computed OD stretch in bulk ice Ih and bulk water (Fig. 3.2.6 and Fig. 3.2.5). These were identified with a water OD donating a strong hydrogen bond to an in-plane, acceptor surface OD, and a water OD accepting a weak hydrogen bond from an out-of-plane, donor surface OD, respectively. Thus, our simulations support the hypothesis that different types of hydrogen bonding interactions may give rise to distinct "ice-like" and "liquid-like" features in the vibrational spectra. Our analysis is in agreement with that of Ref. [142], which was based on VDOS calculations, and our simulations support the hypothesis that different types of hydrogen bonding interactions may give rise to distinct "ice-like" and "liquid-like" features in the vibrational spectra. We emphasize that the observation of an "ice-like" band in experiments does not necessarily indicate the presence of an ordered ice layer at the interface; as we have shown here, a vibrational signature resembling that of bulk ice may also arise from individual water molecules hydrogen-bonded to the alumina surface, without the presence of any long-range order typical of ice. We also note that the alumina-water hydrogen bonding contributions to the vibrational spectra identified here do not exclude the presence of electrostatic orientational effects; it is possible that both contributions act in concert, or at different pH regimes.

Finally, we note that while an "ice-like" peak is clearly present in the simulated IR spectrum, its intensity is very weak in the experimental SFG spectrum at neutral pH. This discrepancy between theory and experiment has at least two possible explanations. The relative spectral weight of the ice-like 2200 cm^{-1} peak might decrease if Raman contributions were taken into account in our calculations, which instead includes only IR activity. In addition, the relative populations of in- and out-of-plane surface OD groups may not be fully equilibrated in our relatively short simulations ($\sim$10 ps), especially since

the in-plane surface OD groups form strong hydrogen bonds as evidenced by the large red shift relative to the dangling OD signal. As the systematic variation of the pH of the aqueous phase is difficult to carry out in first-principles MD simulations, given the limited supercell sizes that are tractable at present, it may be interesting to vary the population of in-plane surface OD groups in future studies, to understand in detail their contributions to SFG spectra as a function of pH.

3.2.5 Summary

Experiments that employ SFG or X-ray techniques provide an incomplete picture of the structural properties of interfacial water molecules at solid/liquid water interfaces. However the combination of the results of these experiments and those of first-principles simulations provide a powerful strategy to understand the structural and dynamical properties of water at solid interfaces. This is illustrated in the present work, where we investigated the structural properties and IR spectra of the alumina(0001)/water interface using first-principles MD, and compared our results with experiments [134, 135, 136].

We found that the first interfacial water layer was adsorbed at a distance of about 2.5 Å from the alumina(0001) surface, in agreement with synchrotron X-ray experiments. This interfacial water region is localized within 5.0 Å of the solid, beyond which the structural properties of liquid water are essentially recovered. Detailed analyses of the computed IR spectra provided a molecular interpretation of SFG vibrational spectroscopy experiments [134, 135, 136]. In particular, the "ice-like" and "liquid-like" features observed in these spectra can be attributed to different types of interfacial water molecules participating in hydrogen bonds with different strengths. The "ice-like" peak arises from strongly hydrogen-bond-donating water molecules to alumina surface OD, while the "liquid-like" peak is due to water molecules accepting a weak hydrogen bond from the surface OD. No indication of a solid ice layer with long range order was found in our simulations. Our

results support the view that the electrostatic charge of the surface is not solely responsible for the orientation of water molecules at the interface; in addition to electrostatic effects from surface charges, hydrogen bond interactions with the surface play a key role, and should be considered as well in interpretations of experiments.

The available experimental SFG data on interfacial water molecules show common features for a variety of systems [121]. In particular, "ice-like" and "liquid-like" bands have been observed for both hydrophobic and hydrophilic interfaces. However it is yet unclear whether the microscopic origin of these bands is the same in the two cases. Future studies employing the analytical approach taken here may help elucidate this question further.

Chapter 4

Electronic and dielectric properties of nitride surfaces and interfaces

4.1 Dielectric properties of Si_3N_4 thin films

4.1.1 Introduction

The degradation of the dielectric properties of silicon dioxide films as their thickness decreases below few nanometers has led to intense research activities on high dielectric constant materials [160, 161, 162]. Among those, silicon nitride (Si_3N_4) has received much attention lately, since the static and optical dielectric constants of bulk crystalline β-Si_3N_4 are higher than the corresponding ones of SiO_2 [163]: $\epsilon_0 = 8.19$, and $\epsilon_\infty = 4.33$, as compared to $\epsilon_0 = 4.6$, and $\epsilon_\infty = 2.46$, respectively. Therefore silicon nitride might be a good candidate as a gate dielectric to reduce tunneling induced leakage currents as the MOS technology is scaled further [162]. Furthermore, compared to SiO_2, silicon nitride is more stable in strong electric fields and at high temperature, and it is more effective in providing diffusion barriers to impurities such as boron [164].

In order to understand if silicon nitride may be a promising high dielectric constant

material at the nanoscale, it is important to understand how its dielectric properties vary as its size is reduced. In this section, we investigate the dielectric properties of Si_3N_4 thin films with lateral dimensions below 6 nm as a function of size and surface structure. In addition we compare dielectric properties of crystalline and amorphous thin films with the same size, and we discuss the effect of structural disorder on the spatial variation of the film dielectric properties.

4.1.2 Methods

We carried out first-principles calculations using DFT in the LDA approximation. We employed plane wave basis sets with a kinetic energy cutoff of 60 Ry, in conjunction with norm-conserving pseudopotentials. Within this framework we obtain structural properties of bulk crystalline β-Si_3N_4 in good agreement with experiment [165] (lattice constant underestimated by 1 %) and static and optical dielectric constants ($\epsilon_\infty = 4.23$ and $\epsilon_0 = 8.2$) in good accord with previous studies ($\epsilon_\infty = 4.33$ and $\epsilon_0 = 8.19$) [163]. We also compared selected phonon frequencies with available measurements [166] and a previous study [167] and obtained good agreement (average error of phonon frequencies, with respect to experiment, is 3.7%). In addition, we note that the measured band gap of silicon nitride (4.7-4.9 eV) [168] is reproduced fairly well by LDA calculations (the computed value is 4.4 eV).

In order to compute the optical and static dielectric constants of thin films, we carried out calculations in a finite electric field, applied in the form of a saw tooth potential as described in Ref. [169]. Since we performed calculations with periodic boundary conditions, an artificial planar dipole potential was added to our simulation cell, following the prescription proposed by Meyer and Vanderbilt [170]. Similar to Ref. [171], we define the microscopic dielectric constant of a slab as the ratio between the applied electric field E_0

and the screened field $E(z)$:

$$\epsilon(z) = \frac{E_0}{E(z)} \tag{4.1.1}$$

Here z is the direction parallel to the slab growth direction. The screened electric field $E(z)$ can be obtained from the gradient of the change in the planar-averaged potential due to the applied electric field:

$$E(z) = -\frac{\partial \Delta \overline{V}(z)}{\partial z} \tag{4.1.2}$$

where the planar-averaged potential is defined as:

$$\overline{V}(z) = \frac{1}{L_x L_y} \int V(x, y, z) dx dy \tag{4.1.3}$$

Here L_x and L_y are the dimensions of the plane perpendicular to the growth direction of the slab. If the ions are fixed to their equilibrium positions, Eq. 4.1.1 yields the optical dielectric constant ϵ_∞. If the ions are relaxed in response to a finite electric field, an ionic contribution is included in the evaluation of the dielectric constant and Eq. 4.1.1 yields the value of ϵ_0. When ionic relaxations are allowed, $\Delta \overline{V}(z)$ has an oscillatory behavior. Such a behavior can be smoothed out by using, e.g. a Gaussian convolution technique [172], with a standard deviation of about 1 Å. This is equivalent to the assumption that the dielectric response of the system is local over a scale of 1 Å.

The average dielectric constant $\bar{\epsilon}$ of a slab representing a thin film is evaluated by using the relation [172]:

$$\frac{L}{\bar{\epsilon}} = \int \frac{dz}{\epsilon(z)} \tag{4.1.4}$$

where L is the thickness of the slab. When using Eq. 4.1.4 the slab boundaries are taken to coincide with the positions of the outermost atoms.

We considered β-Si_3N_4 crystalline slabs grown in different directions, with varying

number of layers, i.e. with varying thickness. We considered two growth directions, the [0001] direction by using hexagonal unit cells [173], and the [010] direction by using orthorhombic unit cells [174]. We terminated the surface dangling bonds on each side of the slabs by hydrogen atoms, and we optimized the ionic positions of the whole slab using the QUANTUM ESPRESSO code [145].

In order to understand the difference in dielectric properties between crystalline and amorphous Si_3N_4 thin films, we generated several amorphous slab models by combining classical and *ab initio* MD simulations. A Si_3N_4 crystalline, bulk sample was first melted at 3000 K by using classical MD simulations with a modified Tersoff potential [175]. About 10 Å vacuum was then added along the z axis to the bulk liquid in order to generate two free surfaces. Using this structure as the starting configuration, we conducted a classical MD simulation at 3000 K and collected seven snapshots along the MD trajectory every 100 ps. This time interval ensures that the generated snapshots are structurally uncorrelated. Each of the seven snapshots was then equilibrated at 2000 K and annealed to room temperature for about 2 ps using *ab initio* MD employing the Qbox code [89] that is optimized for use on high performance architectures and is highly efficient for *ab initio* MD runs. Finally, all atoms of the annealed configurations were fully relaxed to the positions of the nearest energy minimum. At variance with crystalline slabs, the surfaces of our amorphous thin films were not hydrogen terminated. A detailed description of the amorphous slab structure is discussed in Ref [66].

4.1.3 Results and Discussion

Fig. 4.1.1 and Fig. 4.1.2 show $\epsilon(z)$ (see Eq. 4.1.1) for crystalline slabs with different growth directions. Consistent with previous studies of silicon slabs [171, 172], we find that the dielectric constant decays smoothly at the surface. Its oscillations within the slab originate from the ordered structure of the crystalline thin film. The difference in

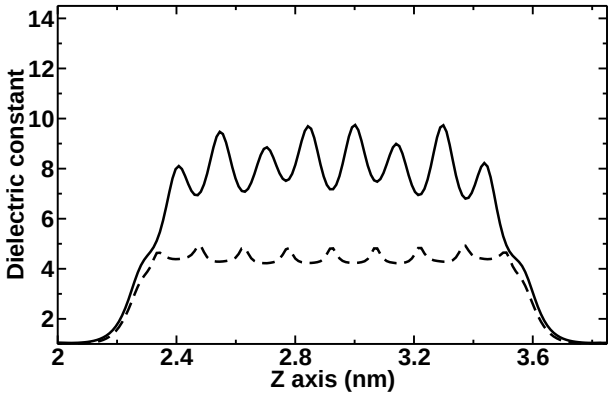

Figure 4.1.1: Spatial variation of the optical (dashed line) and static (solid line) dielectric constants along the z direction (parallel to the growth direction) of a crystalline Si_3N_4 slab grown in the [0001] direction.

the dielectric constant profile in Fig. 4.1.1 and Fig. 4.1.2 is due to the different atomic arrangements in the two films, that are grown in different directions. In Fig. 4.1.3 we show the spatial variation of the dielectric constant along the z axis of one model of amorphous silicon nitride (z is perpendicular to the slab surfaces). As expected, the dielectric constant shows a non-homogenous behavior due to the disordered arrangement of atoms.

In low frequency measurements, one probes the static dielectric constant, inclusive of both electronic and ionic contributions; our results for ϵ_0 of crystalline slabs, as obtained using Eq. 4.1.4, are reported in Fig. 4.1.4. Below 4 nm, crystalline slabs show a substantially reduced screening with respect to the bulk value, indicated by the dashed line in the figure. This reduction is mostly due to the decay of the dielectric constant in proximity of the surface and it depends on the film growth direction. As the thickness of the slab decreases, the ratio of surface to volume increases, inducing a stronger reduction in the dielectric constant. It is interesting to note that, for both growth directions, while the optical dielectric constant reaches values close to that of the bulk already for thicknesses of about 5 nm, ϵ_0 remains smaller than that of the bulk for considerably larger thick-

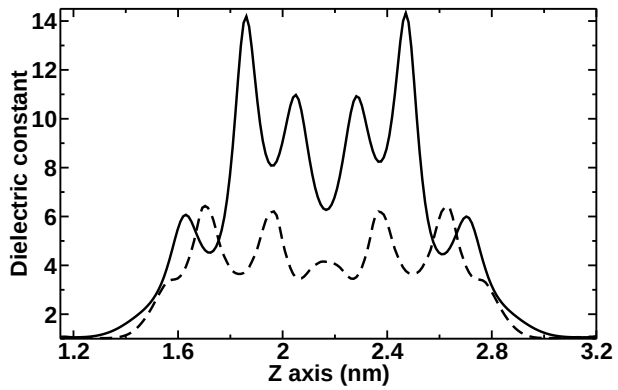

Figure 4.1.2: Spatial variation of the optical (dashed line) and static (solid line) dielectric constant along the z direction (parallel to the growth direction) of a crystalline Si_3N_4 slab grown in the [010] direction. The thickness of the slab (≈ 1.47 nm) is very similar to the one of the slab reported in Fig. 4.1.1 (≈ 1.41 nm).

nesses. In addition, we find that at the nanoscale, the static dielectric constant is much more sensitive to the growth direction than the optical dielectric constant, as indicated in Fig. 4.1.4.

In Fig. 4.1.5 we also show the dielectric constant of amorphous models whose thickness is about the same as that of a nine layer crystalline [0001] slab. We computed the dielectric constants for several amorphous slab models (we chose seven of them), and the mean value and standard deviations were obtained by averaging over the results obtained for seven configurations. We note that both optical and static dielectric constants are very similar for the several amorphous slab models considered here, with standard deviations $\sigma_{\epsilon_\infty} = 0.2$ and $\sigma_{\epsilon_0} = 0.31$, respectively. Both the static and optical dielectric response of the amorphous slabs are smaller than those of the crystalline [0001] slab with the same thickness. However they are very similar to those of the orthorhombic slab (see Fig. 4.1.4). This indicates that amorphization of thin films may be an important factor in decreasing the dielectric constant at the nanoscale. All of our results show that the presence of a surface in the film greatly affects its average dielectric properties and indicates that surface and interface effects play a key role in determining the dielectric response at the nanoscale.

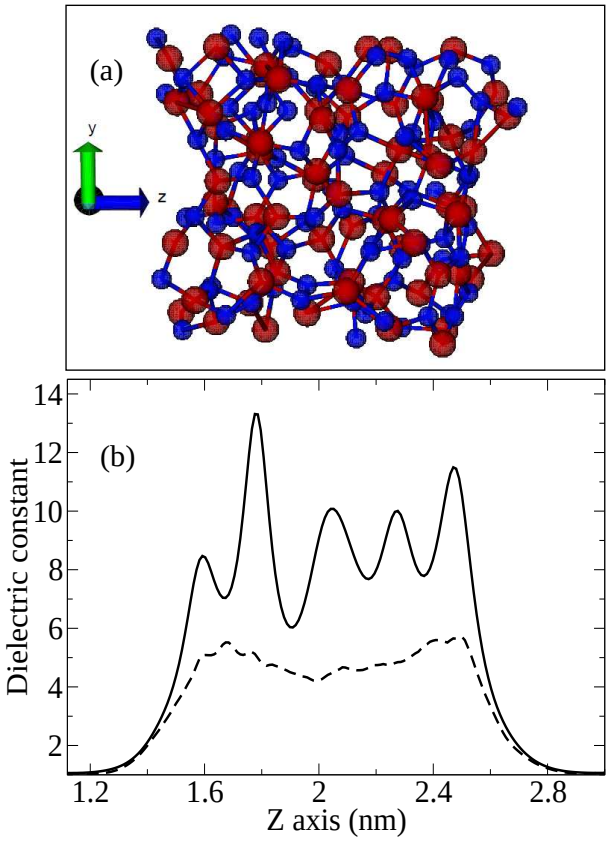

Figure 4.1.3: (a): A snapshot of an amorphous slab model, with 166 atoms used in calculations; Silicon and Nitrogen are represented by red and blue balls, respectively. (b): the spatial variation of the optical (dashed line) and static (solid line) dielectric constant along the z direction (perpendicular to the slab surfaces of the amorphous slab shown in (a)).

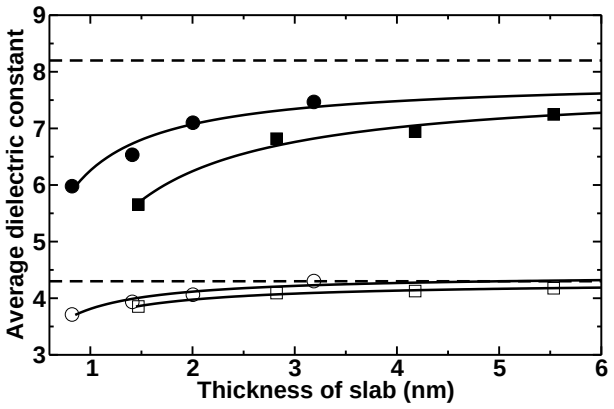

Figure 4.1.4: Dielectric constant of crystalline slabs grown in the [0001] direction (hexagonal cell) compared to that of crystalline slabs grown in the [010] direction (orthorhombic cell). Circles (squares) represent data for hexagonal (orthorhombic) cells, open (filled) symbols represent ϵ_∞ (ϵ_0). Dashed lines indicate values of the dielectric constants for bulk crystalline Si_3N_4.

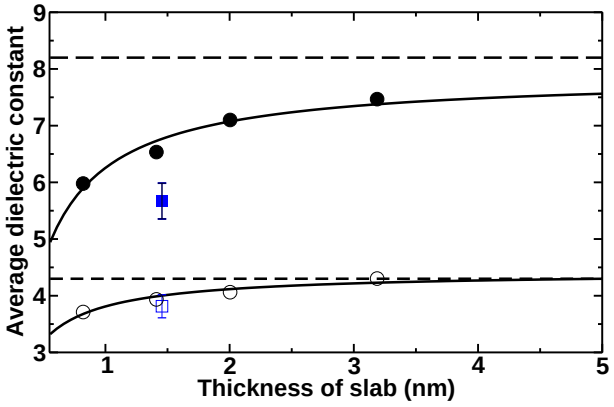

Figure 4.1.5: Dielectric constant of several Si_3N_4 slabs. Circles and squares represent the data for crystalline (in the [0001] direction) and amorphous slabs, respectively. Open (filled) symbols represent ϵ_∞ (ϵ_0). Dashed lines indicate the optical and static dielectric constant values of bulk crystalline Si_3N_4.

We note that for both amorphous and crystalline Si_3N_4 slabs with thicknesses of about 1 nm, the dielectric constants are still larger that those of SiO_2 ($\epsilon_0 = 4.6, \epsilon_\infty = 2.46$).

4.1.4 Summary

In summary, in this section we investigated the behavior of the dielectric properties of silicon nitride films as a function of thickness, using *ab-initio* calculations based on DFT. For crystalline films we find non negligible variations of the static dielectric constants with respect to bulk values for thicknesses up to 6 nm and a reduction by a factor as large as 1.32 and 1.15 for, e.g. 1 and 2 nm slabs grown in the [0001] direction, respectively. Amorphization, that is atomic disorder, may further decrease both the high frequency and the static dielectric constants of the films.

4.2 Band offset and dieletric properties of the amorphous Si_3N_4/Si(100) interface

4.2.1 Introduction

The continuing reduction of the silicon dioxide (SiO_2) gate dielectric thickness in modern metal-oxide-semiconductor (MOS) technology results in high direct tunneling gate leakage currents, leading to excessive standby power consumption. As a result, there is strong interest in alternative gate dielectrics with dielectric constant higher than that of SiO_2 (high-κ materials), which may yield the same gate capacitance as SiO_2 but with a larger physical thickness, thus decreasing gate leakage currents [176]. Silicon nitride (Si_3N_4) has been intensively investigated as a prototype high-κ material to be alloyed with SiO_2 and transition metal oxides, and be part of high-κ stack dielectric films [177, 178]. In particular, the dielectric constant of bulk Si_3N_4 is higher than that of SiO_2 [67, 66], which

may help reduce tunneling induced leakage currents as the MOS technology is scaled further.

Together with the dielectric constant, the performance of Si_3N_4 as a dielectric material critically depends on the electronic structures of the interface with the Si substrate, in particular on the valence band (VBO) and conduction band offsets (CBO), which define the barrier height for injection of electrons or holes [161]. The VBO of amorphous silicon nitride (a-Si_3N_4) thin films on Si(100) has been studied using photoemission spectroscopy, and different results were reported, which appear to depend on the preparation techniques of the nitride film on the Si substrate [179, 180, 181]. Keister [179] deposited a-Si_3N_4 thin films on Si(100) by remote plasma chemical vapor deposition (CVD), using N_2 and SiH_4 sources and reported a VBO of 1.78 eV, while Gritsenko [180], who used the low pressure CVD technique with a SiH_2Cl_2+NH_3 gas mixture found a value of 1.5 eV for the VBO. More recently Higuchi [181] measured a VBO of 1.6 eV after direct nitrition of the Si(100) substrate with NH radicals.

In contrast to the dielectric constant and the VBO, the thickness at which nitride thin films on a Si subtrate recover their bulk properties is less understood. The value of this thickness is important, as it defines the fundamental limit where Si_3N_4 can still operate properly as a gate dielectric. While the SiO_2/Si interface has been widely studied [182, 183, 184, 185], and a value of 7 Å was reported for the thickness of SiO_2 thin films on a Si subtrate by both experiment [186] and *ab-initio* calculations [187, 188], the corresponding value for Si_3N_4 is not yet known.

Most theoretical studies has so far been limited to the β-$Si_3N_4(0001)$/Si(111) interface [189, 190, 191, 192], due to the almost perfect match (lattice constant and symmetry) between the β-$Si_3N_4(0001)$ and Si(111) surfaces, and no *ab-initio* study of a-Si_3N_4/c-Si interfaces is yet available. Omeltchenko [193] reported an investigation of the structural properties of the interface of a-Si_3N_4 with Si(111) using empirical potentials, and band

offsets were investigated by Robertson [161], with a tight binding model that requires experimental data for the band gaps of c-Si and a-Si_3N_4 as input parameters.

In this work we present a microscopic study of the a-Si_3N_4/Si(100) interface. We used a combination of classical and *ab-initio* molecular dynamics (MD) to obtain a realistic model of the interface, and we computed band offsets by means of many-body perturbation theory (MBPT), within the G_0W_0 approximation; we used a GW technique based on a spectral decomposition of the dielectric matrix [72, 194] that allowed us to tackle relatively large samples. We show that density functional theory (DFT) with semi-local functionals is not adequate to obtain the valence band offset, as the cancellation of errors reported [195], e.g. for semiconductor heterojunctions, was not observed in the a-Si_3N_4/c-Si case, similar to other semiconductor/insulator interfaces [196, 197]. MBPT yielded instead a good agreement with experiments. In contrast to the VBO, semi-local DFT calculations yielded dielectric constants of c-Si and a-Si_3N_4 in good agreement with experiments (within 10%); DFT was then used to investigate the local variation of the dielectric constant across the interface. In addition, based on the calculated local band edges and dielectric constants, we estimated the minimum thickness of the nitride required to recover the system bulk properties.

4.2.2 Methods

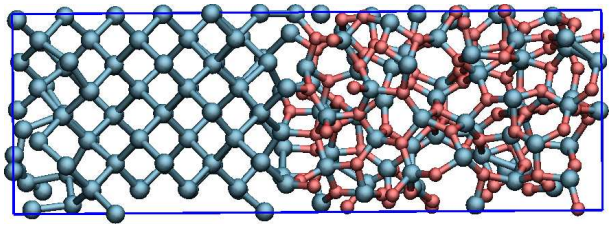

Figure 4.2.1: Atomistic model representing an a-Si_3N_4/Si(100) interface; Si and N atoms are represented by blue and red symbols, respectively.

Our *ab-initio* simulations were carried out with plane-wave basis sets with a kinetic energy cutoff of 70 Ry, in conjunction with norm-conserving pseudo-potentials and the Perdew-Burke-Ernzerhof (PBE) exchange-correlation functional [7]. The a-Si_3N_4/Si(100) interface was modelled by a 264-atom superlattice that contains 96 and 168 atoms in the c-Si and a-Si_3N_4 components, respectively, as illustrated in Fig. 4.2.1. The surface area of the a-Si_3N_4 was chosen to match that of the (2×2) Si(100) surface with a lattice constant of 5.43 Å. The initial interface model was generated by joining the bulk β-Si_3N_4 and c-Si structures along the direction perpendicular to the (100) surface of c-Si. Then the β-Si_3N_4 component was amorphized by melting and quenching from 4000 K to 300 K for about 3 ns, using classical MD with the a modified Tersoff potential [175]. The structural model was further refined using *ab-initio* MD simulations, carried out with the Qbox code [89], by melting and quenching from 2000 K to 300 K for about 14 ps. During the annealing process, the atomic positions of c-Si were kept fixed up to 1000 K, and below this temperature all the atoms were allowed to move. The final structure was further relaxed to the closest local minimum using the Quantum Espresso code [145]. We obtained an average mass density of 3.0 g/cm^3 for the a-Si_3N_4 region of the slab, consistent with the experimental density of 3.1 g/cm^3 [198]. We note that in our interface model, the total thickness of the a-Si_3N_4 component is ∼15 Å, similar to the thickness (15-20 Å) of a-Si_3N_4 films investigated in photoemisson experiments [179, 181].

4.2.3 Results and Discussion

We computed the VBO (CBO) between c-Si and a-Si_3N_4 as the difference between the VBM (CBM) in the bulk-like c-Si and a-Si_3N_4 regions of the interface:

$$\text{VBO} = E_v^{Si} - E_v^{Si_3N_4};\ \text{CBO} = E_c^{Si} - E_c^{Si_3N_4}. \tag{4.2.1}$$

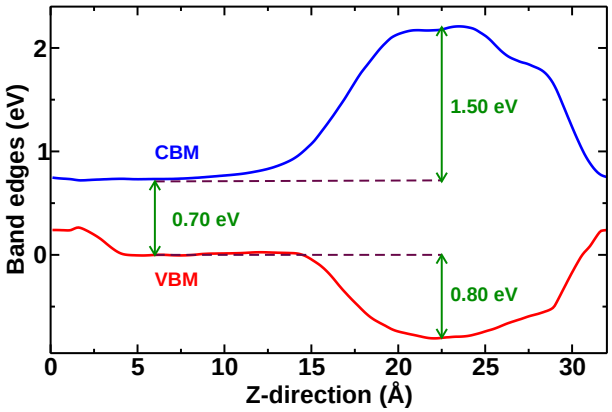

Figure 4.2.2: Variations of the valence band maximum (VBM) and conduction band minimum (CBM) along the direction orthogonal to the interface plane, obtained at the DFT-PBE level of theory. Two interfaces are present in the model, one at ∼0 Å and the other one at ∼14 Å.

To obtain values of the VBM and CBM of silicon and of the nitride we defined the local density of states (LDOS) [199]:

$$D(\epsilon, z) = 2 \sum_{n} |\langle z|\psi_n\rangle|^2 \delta(\epsilon - \epsilon_n), \tag{4.2.2}$$

where z is the direction perpendicular to the interface, $|\langle z|\psi_n\rangle|^2$ is the electron density integrated in the xy plane, the factor of 2 accounts for the spin degeneracy and ψ_n, ϵ_n are the Kohn-Sham eigenstates and eigenvalues, respectively. The variation of the values of the VBM and CBM across the interfaces (shown in Fig. 4.2.2) was determined using the relation [199]:

$$\int_{VBM}^{E_F} D(\epsilon, z)d\epsilon = \int_{E_F}^{CBM} D(\epsilon, z)d\epsilon = \Delta \int_{-\infty}^{E_F} D(\epsilon, z)d\epsilon, \tag{4.2.3}$$

where E_f is the Fermi energy of the interface model and the threshold value Δ is determined by imposing that the band edges of the interface model are recovered. Fig. 4.2.2 shows that the bulk band edges are approximately recovered at about $z \sim 7.5$ Å for c-Si

and at $z \sim 22.5$ Å for a-Si_3N_4, hence the transition region extends into the nitride for about $\approx 6-7$ Å, similar to what found for SiO_2 on a Si subtrate [186, 187, 188]. We note that due to the periodicity of our model, two interfaces are present and are located at $\sim$0 Å and 14 Å, respectively. Of the two interfaces, due to the presence of Si^{3+} defects, the one located at $\sim$0 Å exhibits a slightly thicker transition region, associated with smaller values of the local band gap in the transition region.

At the DFT-PBE level of theory, we obtained a band gap of 3.0 eV in the bulk region of a-Si_3N_4, that underestimates the experimental results of 4.5-5.5 eV [200, 201, 202]. The computed band gap in the bulk region of c-Si is 0.7 eV, in agreement with the corresponding bulk calculation of 0.7 eV; as expected this result underestimates the experiment (1.17 eV) [203]. In addition, we obtained values of 0.8 and 1.5 eV for the VBO and CBO, respectively, at the DFT-PBE level of theory. The value of VBO (see Table. 4.1) underestimates experiment (1.5-1.78 eV [180, 179, 181]) by $\sim$50%, showing that in this case, similar to the SiO_2/Si interface [204], there is no cancellation of errors for the heterojunction.

To improve on the computed band offsets, following Ref. [204] we compute the VBO and the CBO within the G_0W_0 approximation ($VBO^{G_0W_0}$, $CBO^{G_0W_0}$):

$$\begin{aligned} \mathrm{VBO}^{\mathrm{G_0W_0}} &= (E_v^{Si} + \Delta E_v^{Si}) - (E_v^{Si_3N_4} + \Delta E_v^{Si_3N_4}), \\ \mathrm{CBO}^{\mathrm{G_0W_0}} &= (E_c^{Si} + \Delta E_c^{Si}) - (E_c^{Si_3N_4} + \Delta E_c^{Si_3N_4}), \end{aligned} \tag{4.2.4}$$

where $\Delta E_{v,c}^{Si}$ and $\Delta E_{v,c}^{Si_3N_4}$ are G_0W_0 corrections to the band edges of bulk c-Si and a-Si_3N_4, respectively, that were obtained from bulk calculations.

For computational convenience, the G_0W_0 corrections to the VBM and CBM of a-Si_3N_4 were computed for a sample of 56 atoms that showed structural and dielectric properties in good agreement with experiment (we verified that there was only one dan-

gling bond in our sample) [66]. The 56 atom structure was generated from a molten β-Si_3N_4 at 3000 K, using a procedure similar to that described above for the a-Si_3N_4/c-Si interface. At the DFT-PBE level of theory, the computed band gaps of the bulk sample of a-Si_3N_4 is 3.17 eV, close to the value of 3.0 eV calculated in the bulk region of a-Si_3N_4 of the interface model. Such a small difference is attributed to structural disorder. MBPT calculations were carried out using the method proposed in Refs. [72, 194], with 800 dielectric eigenpotentials to represent the dielectric matrix. We emphasize that the approach of Refs. [72, 194] does not require the explicit calculation of empty electronic orbitals and that the convergence of the G_0W_0 calculations is controlled by only one parameter (in addition to the planewave cutoff), i.e. the number of eigenpotentials used for the spectral decomposition of the dielectric matrix.

For bulk a-Si_3N_4, G_0W_0 calculations greatly improve the DFT-PBE values, yielding a band gap of 4.87 eV, that falls within the range of available experimental data (4.5-5.5 eV). Our calculations yielded a large correction to the VBM, e.g., $\Delta E_v^{Si_3N_4} = -1.20$ eV and a smaller one for the CBM: $\Delta E_c^{Si_3N_4} = 0.5$ eV. For bulk c-Si, within the G_0W_0 approximation we obtained $\Delta E_v^{Si} = -0.5$ eV and $\Delta E_c^{Si} = 0.1$ eV, consistent with the values of $\Delta E_v^{Si} = -0.4$ eV and $\Delta E_c^{Si} = 0.2$ eV reported in Ref. [204] for calculations that did not make use of plasmon-pole models. We note that in our GW calculations we did not use any model for the frequency dependent dielectric matrix; frequency integration was explicitly performed. Within the G_0W_0 approximation, the computed VBO, 1.5 eV, is in good agreement with the experimental data ranging from 1.5 and 1.78 eV [179, 180, 181] (see Table I). In addition, we obtained a value of 1.9 eV for the CBO. We note that large uncertainties are associated to the measured band gap of a-Si_3N_4, which may affect the value of CBO extracted from experiment. Using a value of 1.5 eV for the VBO, the experimental values of 1.17 eV and of 4.5-5.5 eV for the band gap of c-Si and a-Si_3N_4, respectively, we estimated the experimental value of the CBO to be within 1.83-2.83 eV.

Table 4.1: Band gaps of bulk c-Si (E_g^{Si}), a-Si_3N_4 ($E_g^{Si_3N_4}$), and band offsets of the a-Si_3N_4/Si(001) interface, calculated at the DFT-PBE and G_0W_0 levels of theory. Photoemission spectroscopy experimental results are also shown for comparison. All values are in eV.

	VBO	CBO	E_g^{Si}	$E_g^{Si_3N_4}$
PBE	0.8	1.5	0.7	3.17
G_0W_0	1.5	1.9	1.3	4.87
Expt.	1.5-1.78 [179, 180, 181]	1.83-2.83	1.17 [203]	4.5-5.5 [200, 201, 202]

We now turn to describe the variation of the dielectric constant at the a-Si_3N_4/Si(100) interface, which was characterized using the local dielectric constant, $\epsilon(z)$, introduced in Ref. [172]; $\epsilon(z)$ is computed as:

$$\epsilon(z) = 1 + 4\pi \frac{p(z)}{E(z)}, \tag{4.2.5}$$

where $p(z)$ and $E(z)$ correspond to the induced microscopic polarization and the change in the total electric field due to an applied external field. These quantities can be obtained from the induced charge density, $\rho_{ind}(z)$, using the relation $\partial p(z)/\partial z = -\rho_{ind}(z)$ and $\partial E(z)/\partial z = 4\pi\rho_{ind}(z)$ [172].

We focus on the optical dielectric constant (ϵ_∞) that represents the electronic response to the electric field. Fig. 4.2.3 shows the variation of the local dielectric constant $\epsilon_\infty(z)$ along the direction orthogonal to the interface plane, obtained at the DFT-PBE level of theory. Similar to the variation of local band edges, we found that the local dielectric constant shows a sharp decrease within $\sim 6-7$ Å from the Si subtrate. In the bulk regions of c-Si and a-Si_3N_4 we found average dielectric constants of 13.0 and 5.2, respectively, in agreement with the value of 12.7 for bulk c-Si and of 4.8 for the bulk a-Si_3N_4 sample. We note that the computed optical dielectric constant in the bulk region of a-Si_3N_4 is about two times larger than the value of 2.46 for SiO_2 [163]. Our computed dielectric

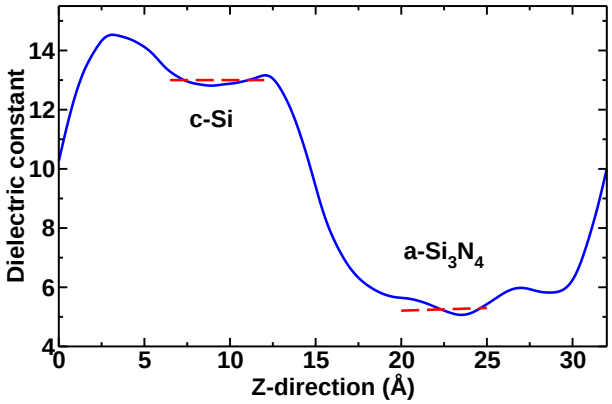

Figure 4.2.3: The variation of the local dielectric constant $\epsilon_\infty(z)$ along the direction orthogonal to the interface plane, obtained at the DFT-PBE level of theory. Red dashed line indicates the average values in the bulk regions of c-Si and a-Si_3N_4.

constants are in reasonably good agreement with the experimental values of 11.7 and 4.5-5.6 for c-Si and a-Si_3N_4 [205, 206], respectively. Hence, even if band gaps and band offsets are not *quantitatively* correct within DFT-PBE, we found that dielectric constants of the interface model are in satisfactory agreement with experiments.

4.2.4 Summary

In summary, we investigated the electronic and dielectric properties of an atomistic model of the amorphous Si_3N_4/Si(100) interface by using density functional theory. We showed that DFT-PBE severely underesimates the experimental value of the VBO. Within the G_0W_0 approximation, we obtained instead a value of 1.5 eV, in good agreement with experiment; our calculations predicted a value of 1.9 eV for the CBO. We found that in the bulk-like region of the interface model, dielectric constants computed within the DFT-PBE level are in reasonably good agreement with experiments. Our calculations also showed that, similar to SiO_2, a minimum thickness of $6-7$ Å is required to recover the bulk properties of a-Si_3N_4; such value thus determines the minimum thickness of an a-Si_3N_4 thin film gate dielectric. Although the value of the minimum thickness is similar

for a-Si_3N_4 and SiO_2, given the higher dielectric constant of the former, using alloys of amorphous nitrides and oxides as gate dielectrics may help suppress direct tunneling leakage currents observed in SiO_2.

Chapter 5

Electronic properties of simple aqueous solutions

5.1 Electronic structure of liquid water

5.1.1 Introduction

Liquid water occupies a central role in many problems concerning energy use and environmental remediation, including its utilization for solar energy capture and conversion in photo-catalytic devices [207]. Yet, surprisingly the electronic properties of water are poorly understood, in spite of their paramount importance in determining the physical and chemical properties of aqueous interfaces, e.g., those with photo-electrodes in water splitting cells. In particular, the determination of basic properties of liquid water, such as the ionization potential and electron affinity remains challenging, from both theoretical and experimental standpoints.

Photoemission spectroscopy (PES) provides information on the valence band structure of liquid water, while inverse photoemission and the use of excess electrons are valuable probes of conduction band energy levels. As discussed, e.g., in Ref. [208], excess electrons

may be "presolvated" in liquid water, prior to reaching a solvated state following solvent reorganization. It has been suggested [209, 210] that thermal fluctuations of molecular dipole moments may give rise to trap states for presolvated electrons, which are localized below the conduction band minimum (CBM) of neutral liquid water. Thus, understanding the electronic structure of neutral liquid water is also crucial to unravel fundamental processes involving excess electrons, which have been the subject of intense experimental and theoretical studies [208].

One of the first extensive PES studies of the electronic structure of liquid water was conducted by Delahay *et al.* [211, 212]. The authors showed that close to the threshold energy E_t, the photocurrent (Y) depends quadratically on the photon energy E, i.e., $Y \sim (E - E_t)^2$, consistent with the theoretical predictions of Brodsky *et al.* [213]. By extrapolating $Y^{0.5}$ to zero, Delahay *et al.* [211, 212] found $E_t = 10.06$ eV, i.e., the valence band maximum (VBM) lying 10.06 eV below the vacuum level.

The development of the vacuum liquid microjet technique led to great advances in liquid PES, and in the understanding of the electronic structure of liquids and solutions [214, 215]. This approach permits the transfer of electrons to the detector essentially free of collisions with gas-phase water molecules, and thus it provides a more precise measurement of the electronic levels of liquid water. Using microjets, Winter *et al.* [215] reported $E_t = 9.9$ eV for water at ambient conditions. However, we note that this value was derived from a linear extrapolation of the photocurrent in the proximity of the threshold energy; by employing the power law used in Refs. [211, 212], one would instead obtain $E_t = 9.3$ eV from the same experimental data. This value is ≈ 0.7 eV smaller than the one previously reported [211, 212], possibly due to a reduction in electron collisions with gas phase water molecules in liquid microjet experiments.

Unlike the VBM position, the CBM of liquid water was inferred only indirectly from solution measurements. Under the assumption that excess electrons are injected into the

water conduction band, photoelectron experiments on electrolytes and photoionization experiments of aqueous solutions led to similar values of $V_0 = -1.2$ eV [216, 217], i.e., the CBM lying 1.2 eV below the vacuum level. However, we note that excess electrons may initially localize in trap states below the water CBM [209, 210], and the value -1.2 eV might not correspond to the position of the CBM of neutral water. An estimate of $V_0 = -0.74$ eV was given based on thermodynamic considerations [210]. By combining the value of $E_t = 9.3 - 10.06$ eV with $V_0 = -(0.74 - 1.2)$ eV one finds a quasiparticle band gap of $E_t - |V_0| = 8.7 \pm 0.6$ eV [210]. Given the large uncertainty in the experimental measurements, theoretical efforts are needed to provide a fundamental understanding of the electronic structure of liquid water.

In this work, we report results for the band gap and band edge positions of liquid water, obtained by combining *ab initio* and classical molecular dynamics (MD) simulations, and electronic structure calculations within many-body perturbation theory (MBPT). We found that the use of MBPT is essential to obtain results in satisfactory agreement with existing measurements. In addition, we analyzed the effect of structural properties on the electronic structure of water, in particular we compared results for the band edge positions obtained with configurations generated by first-principles and classical MD simulations. We note that trap states could be accessed by considering liquid water samples with an explicit excess electron [218]. Since the focus of our work is on the band gap and band edge positions of liquid water, the effect of an excess electron was not considered here.

5.1.2 Methods

An accurate determination of the electronic gap and band edge positions of water encompasses several theoretical and computational challenges, in particular: (i) the generation of well equilibrated MD trajectories to obtain structural models of the liquid; and (ii) the calculation of electronic states of models composed of several hundreds of electrons using

advanced electronic structure methods, such as hybrid functionals or many-body perturbation theory, e.g., within the G_0W_0 approximation [8, 219]. In particular, calculations at the G_0W_0 level of theory are rather demanding from a computational standpoint, and in the case of water they have been mostly limited to 16-molecule samples [220, 221], and only recently 64-molecule cells were used [124]. In addition, MPBT studies were limited to band gaps and absorption spectra, and band edge calculations have not yet been reported for liquid water.

We carried out Born-Oppenheimer *ab initio* MD simulations using the Qbox code [89], with interatomic forces derived from density functional theory (DFT) using the Perdew, Burke, and Ernzerhof (PBE) approximation for the exchange and correlation energy functional [7], The interaction between valence electrons and ionic cores was represented by norm-conserving pseudopotentials [222], and the electronic wavefunctions were expanded in a plane-wave basis set truncated at a cutoff energy of 85 Ry. We used a cubic simulation cell containing 64 water molecules at the experimental density and we sampled the Brillouin zone by the Γ-point only. All hydrogen atoms were replaced with deuterium in order to maximize the allowable time step; we adopted a time step of 10 a.u. in all simulations.The equilibration runs were carried out at constant temperature (NVT conditions) with T= 390 K. The neglect of zero-point motion effects [125] of light nuclei and the use of the PBE approximation are known to yield an overstructured liquid water at ambient conditions, and an elevated simulation temperature of T≈ 400 K may be used to recover the experimental structure and diffusion coefficients at T= 300 K [223, 224]. Snapshots employed in G_0W_0 calculations were extracted from a 20 ps trajectory obtained with *ab initio* runs at constant energy (NVE), following the NVT equilibration. Furthermore, to investigate size effects, smaller snapshots with 32 water molecules were also generated with the same simulation protocol.

In addition to configurations obtained from *ab initio* simulations, we considered snap-

shots of 64 water molecules generated by classical MD with the TIP3P, TIP4P[225] and SPC/E[226] empirical potentials. Classical equilibration runs were carried out at an average temperature of T= 300 K, as empirical potentials were designed to reproduce the properties of water at ambient conditions. Snapshots used in G_0W_0 calculations were collected from 10 ns NVE simulations.

The quasiparticle energies (E_n^{qp}) were calculated at the G_0W_0 level of theory [219], as a first order correction to the Kohn-Sham (KS) energies (ε_n):

$$E_n^{qp} = \varepsilon_n + \langle\psi_n|\Sigma_{G_0W_0}(E_n^{qp})|\psi_n\rangle - \langle\psi_n|V_{xc}|\psi_n\rangle, \tag{5.1.1}$$

where V_{xc} is the exchange-correction potential entering the chosen KS Hamiltonian and $\Sigma_{G_0W_0}$ is the self-energy operator, computed from the one particle Green's function (G_0) and the screened Coulomb interaction (W_0). We used the method proposed in Refs. [72] and [194], which does not require the explicit calculation of empty electronic orbitals nor the use of a plasmon-pole model [219]; the convergence of the G_0W_0 calculation is controlled by only one parameter (in addition to the planewave cutoff), i.e the number of eigenpotentials (N_{eig}) used for the spectral decomposition of the dielectric matrix [32, 34]. We used $N_{eig} = 1000$ and we verified that the quasiparticle energies were converged within 0.04 eV. We averaged our results over 10 configurations chosen to be equally spaced in time over a trajectory of 20 ps and of 10 ns for *ab initio* and classical simulations, respectively. In addition to the G_0W_0 approach, hybrid functionals were also employed to compute the band gap for comparison.

5.1.3 Results and Discussion

Fig. 5.1.3 presents the results for the quasiparticle band gap of 64 and 32-molecule configurations generated by *ab initio* MD, and computed at different levels of theory. The

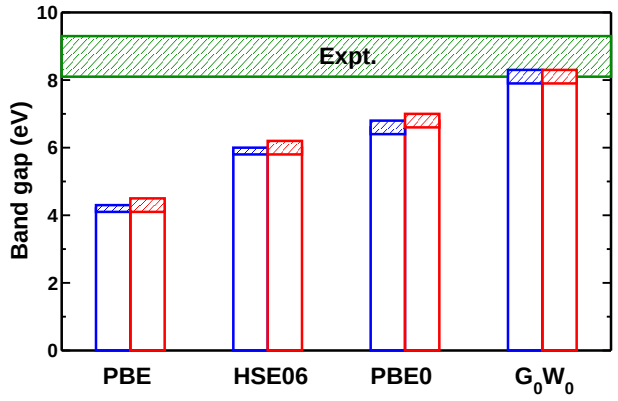

Figure 5.1.1: Quasiparticle band gap (eV) of liquid water models generated by *ab initio* MD simulations with 32 (red rectangle) and 64 (blue rectangle) molecule cells, computed at the DFT level of theory with semi-local (PBE) and hybrid functionals (PBE0 [51, 52] and HSE06 [53]), and at the G_0W_0 level of theory. The experimental band gap of liquid water is 8.7 ± 0.6 eV [210]. Error bars (standard deviation) of theoretical calculations and experimental results (see text) are indicated by shaded areas.

results appear to be insensitive to the cell size, consistent with previous studies [223, 224]. As expected, the PBE functional greatly underestimates the water band gap; although a substantial improvement on the PBE value, the HSE06 and PBE0 band gaps are still lower than the experimental value by ≈2.7 eV and ≈2.0 eV, respectively. These inaccuracies can be qualitatively understood in term of average dielectric screening. In the case of a condensed system, the appropriate fraction of exact exchange (α) entering the definition of a hybrid functional should be approximately proportional to the inverse of the electronic dielectric constant (ϵ_∞) of the material, i.e $\alpha \approx 1/\epsilon_\infty$ [227]. For the PBE0 and HSE06 functionals, $1/\alpha = 4$, a value substantially different from $\epsilon_\infty = 1.7 - 1.8$ [221], hence the gap is underestimated when using these functionals. The best agreement with experiment was obtained using MBPT, yielding a band gap of 8.1 ± 0.2 eV.

In order to evaluate the band edge positions of the liquid with respect to vacuum ($\mathrm{VBM^{vac.}}$ and $\mathrm{CBM^{vac.}}$), we adopted a two-step procedure: (i) we computed the band edge positions ($\mathrm{VBM^{bulk}}$, $\mathrm{CBM^{bulk}}$), including G_0W_0 corrections ($\Delta\mathrm{E^{VBM}}$, $\Delta\mathrm{E^{CBM}}$), with respect to the average electrostatic potential of a bulk model; (ii) we computed the relative

average electrostatic potential (ΔV) of the bulk and vacuum regions, by employing water slabs in contact with a vacuum region. The values $\mathrm{VBM^{vac.}}$ and $\mathrm{CBM^{vac.}}$ are then given by:

$$\begin{aligned} \mathrm{VBM^{vac.}} &= \mathrm{VBM^{bulk}} + \Delta\mathrm{E^{VBM}} + \Delta\mathrm{V} \\ \mathrm{CBM^{vac.}} &= \mathrm{CBM^{bulk}} + \Delta\mathrm{E^{CBM}} + \Delta\mathrm{V}. \end{aligned} \tag{5.1.2}$$

Step (i) was carried out using bulk models generated by *ab initio* simulations and, for comparison, by classical MD simulations. The calculation of ΔV involves the generation of MD trajectories of water slabs in contact with a relatively thick vaccum region, which are computationally expensive to obtain from *ab initio* simulations. To reduce the computational cost, we generated water slabs using classical potentials (TIP3P, TIP4P and SPC/E) and we explored the possibility of computing ΔV for configurations extracted from classical MD trajectories. Our models consisted of 108 water molecules in contact with a vacuum region of $\approx$ 60 Å in a supercell of dimensions of $L_x = L_y = 12.77$ Å and $L_z = 80.0$ Å; the simulations were performed at 300 K for a total simulation time of $\approx$10 ns. The value of ΔV was computed as the difference in the average electrostatic potential $\overline{V}(z)$ between the region where the liquid is present and the vacuum region:

$$\overline{V}(z) = \langle \int\int dxdyV(x,y,z)/A \rangle, \tag{5.1.3}$$

where A is the surface area of the water slab and z is the direction perpendicular to the water/vaccum interface. For each simulation, we used 50 configurations to evaluate $\overline{V}(z)$, chosen to be equally spaced in time over a trajectory of 10 ns.

Fig. 5.1.3 shows the $\overline{V}(z)$ of three different water slabs generated using TIP3P, TIP4P and SPC/E potentials. We found that the value of ΔV is insensitive to the choice of the classical potentials, yielding a value of 3.7 ± 0.05 eV, in good agreement with that

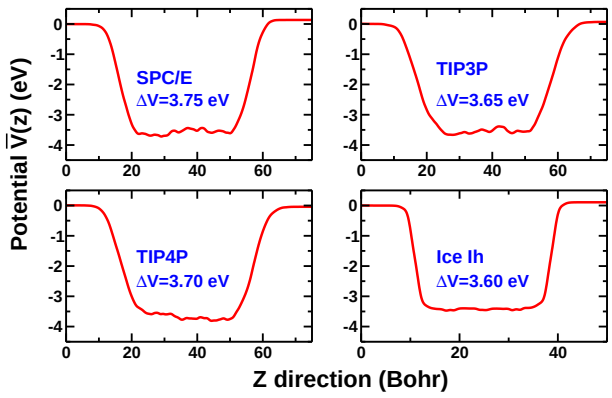

Figure 5.1.2: Planar average of the electrostatic potential ($\overline{V}(z)$) along the z-direction of water slab models generated using different classical potentials (TIP3P, TIP4P and SPC/E). The vacuum level is set at zero. The corresponding average values of the relative average electrostatic potential (ΔV) of the bulk and vacuum regions are presented, with an average fluctuation of $\approx$ 0.1 eV in all cases. The same quantity computed for an ice Ih slab is shown for comparison.

reported in Ref. [228] (ΔV = 3.63 eV) for a SPC/E water slab. We also computed ΔV for an ice Ih slab generated in an *ab initio* simulation at 150 K [156], and we obtained a slightly lower value of ΔV = 3.6 eV. The weak sensitivity of ΔV to the force field is an important result from the computational standpoint, as it avoids the need to carry out full equilibrations of water slabs using *ab initio* simulations. In the following, we used the average value ΔV = 3.7 eV for the evaluation of the positions of the VBM and CBM reported in Table 5.1 and Fig. 5.1.3.

While the average electrostatic potential entering Eq. 5.1.2 is fairly insensitive to the empirical potentials employed, we found that the electronic band gap shows a more complex variability. Table 5.1 shows results for the band gap and band edge positions of liquid water, computed at the PBE and G_0W_0 levels of theory for snapshots generated by classical and *ab initio* simulations. While TIP3P and TIP4P configurations yield band gaps in good agreement with that obtained with *ab initio* configurations, SPC/E water models exhibit a smaller ($0.3 - 0.5$ eV) band gap at both the PBE and G_0W_0 levels of theory. In particular, the G_0W_0 band gap of the SPC/E model shows the largest deviation

Table 5.1: Positions of the valence band maximum (VBM) and conduction band minimum (CBM) of liquid water with respect to the vacuum level, computed at the PBE and G_0W_0 levels of theory for configurations obtained from first-principles (PBE) and classical (SPC/E, TIP3P and TIP4P) MD simulations. Experimental results for the VBM are -10.06 eV [211, 212], -9.9 eV [215], and -9.3 eV as inferred from the experimental results reported in Ref. [215]. The experimental results for the CBM are -1.2 eV[216, 217] and -0.74 eV [210].

	SPC/E	TIP3P	TIP4P	PBE
PBE				
VBM	-5.9 ± 0.1	-6.1 ± 0.1	-6.1 ± 0.2	-6.1 ± 0.1
CBM	-2.2	-2.0	-2.0	-1.9
Band gap	3.7 ± 0.1	4.1 ± 0.1	4.1 ± 0.2	4.2 ± 0.1
G_0W_0				
VBM	-8.7 ± 0.2	-9.0 ± 0.2	-9.0 ± 0.2	-8.8 ± 0.2
CBM	-0.9	-0.7	-0.7	-0.7
Band gap	7.8 ± 0.2	8.3 ± 0.2	8.3 ± 0.2	8.1 ± 0.2

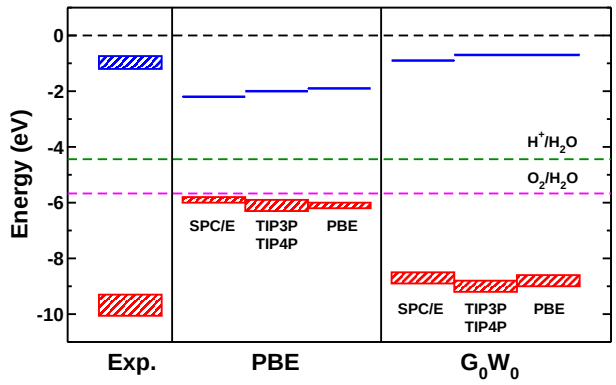

Figure 5.1.3: Positions of the valence band maximum (thick red lines) and conduction band minimum (thick blue lines) of liquid water with respect to the vacuum level computed at the PBE (middle panel) and G_0W_0 (right panel) levels of electronic structure theory (see Table 5.1), compared to experiments [211, 212, 215, 216, 217, 210] (left panel). The vacuum level, reduction and oxidation potentials of water [229] are indicated by the dashed black, green and magenta lines, respectively. We show values obtained using water configurations extracted from first-principles (PBE) and classical (SPC/E, TIP3P and TIP4P) MD simulations.

from experimental values.

To understand why SPC/E water configurations exhibit the smallest band gap of all models, we investigated further how the liquid structure affects the calculated electronic gap. When the OH bond length of the SPC/E model was decreased from the original value of 1.0 Å to 0.957 Å (i.e the OH bond length in the TIP3P and TIP4P water models), the PBE band gap of SPC/E water snapshots increased to 4.0±0.2 eV, in agreement with values obtained for TIP3P and TIP4P configurations. We found similar results for the isolated water molecule, for which the computed HOMO-LUMO gap decreased by ≈0.23 eV when the OH bond length increased from 0.957 Å to 1.0 Å. These results indicate that the effect of the intra-molecular structure on the liquid band gap is not negligible (of the order of 0.3−0.5 eV).

The temperature and water density used to generate configurations in MD simulations also affects the calculated gap. For example, the PBE band gap of configurations generated by *ab initio* MD simulations slightly increases with decreasing temperature (4.2±0.1 eV at 390 K and 4.5±0.1 eV at 317 K for configurations at the experimental density of 1.1 g/cm^3) and with increasing density (3.9±0.1 eV at 0.85 g/cm^3 and 4.2±0.1 eV at 1.1 g/cm^3 for configurations generated at T=390 K). Our calculations thus indicate that the overall variation in the band gap of the liquid may be as large as 0.7 eV, depending on the level of theory, and on the specific thermodynamic conditions used to generate liquid water configurations.

We now turn to the comparison of the computed band edge positions with experimental results. Similar to the band gap, the band edge positions obtained with TIP3P, TIP4P and *ab initio* configurations are in agreement within 0.2 eV. In the following we focus on the results computed for configurations obtained from *ab initio* simulations.

As shown in Table 5.1, the positions of the VBM and CBM predicted by the PBE functional severely deviate from experiment [230], in particular the VBM position is

qualitatively incorrect, as it lies only ≈ 0.5 eV below the experimental water oxidation potential. Instead, a satisfactory agreement is obtained when the G_0W_0 approach is employed. When compared to the value reported by Winter *et al.* [215], the position of the VBM computed by MBPT for *ab initio* configurations is still ≈ 1.1 eV too high. As discussed in the introduction, if a power law of the type $Y \sim (E - E_t)^2$ were used to describe the photocurrent in the region near the threshold energy, a value of $E_t = 9.3$ eV would be obtained. Using this value for the threshold energy, the agreement between our theoretical results and experiment is within 0.5 eV for the position of the VBM (see Table 5.1). Since G_0W_0 calculations probe the photocurrent onset, i.e the very top of the valence band, the value $E_t = 9.3$ eV appears to be appropriate for a comparison with our results. The remaining discrepancy with experiments may stem from experimental conditions, e.g., possible electrokinetic charging effects in the experiments [231, 232] that are not present in the calculations. Discrepancies could also be due to the neglect of off-diagonal matrix elements in the self-energy and the use of approximate PBE wavefunctions as input in G_0W_0 calculations. In particular, it was shown that by taking into account off-diagonal matrix elements of the self-energy [233] or by using PBE0 wavefunctions as input [56], the value of the HOMO of an isolated water molecule decreases by $\approx 0.3 - 0.4$ eV compared to the one obtained with PBE wavefunctions, resulting in a better agreement with experiment.

While the use of PBE0 wavefunctions may affect the valence band position obtained with G_0W_0, we expect that the CBM position would be weakly modified by the use of PBE0, as the LUMO of water is a delocalized state. In particular, the IPR (inverse participation ratio) ratio between HOMO and LUMO wavefunctions of *ab initio* water configurations is $\approx$210. The IPR of a wavefunction ψ is defined as:

$$IPR = \frac{\frac{1}{N}\sum_{i=1}^{N} |\psi(r_i)|^4}{[\frac{1}{N}\sum_{i=1}^{N} |\psi(r_i)|^2]^2}, \tag{5.1.4}$$

where N is the number of points in the real-space grid used to represent the wavefunction. An IPR value of 1 indicates that the wavefunction is completely delocalized and the value increases from 1 with the localization of the wavefunction. Therefore we consider our result for the position of the CBM, -0.7 eV, to be in reasonable agreement with the -0.74 eV estimate proposed in Ref. [210]. We suggest that experimental results putting the CBM at -1.2 eV may possibly be probing presolvated electron states and not the CBM of the neutral liquid.

5.1.4 Summary

In summary, we reported the first theoretical determination of the positions of the VBM and CBM of liquid water using many-body perturbation theory coupled with *ab initio* MD simulations, and we presented a comprehensive analysis of the effect of structural models on the electronic properties of the liquid. We showed that the positions of the VBM and CBM predicted by DFT with a semi-local density functional are qualitatively incorrect, and the use of MBPT is crucial to obtain (semi)-quantitative agreement with experiments. We found that intramolecular structural effects and variation of temperature and pressure close to ambient conditions may lead to variations of the water band gap of $0.3-0.7$ eV. We also showed that the structure of water slabs, used to compute the position of the vacuum level, may be conveniently generated using classical force fields, in particular the TIP3P and TIP4P empirical potentials, thus leading to substantial computational savings. Our interpretation of experiments, especially the position of the CBM, and our assessment of the level of theory necessary to describe the electronic properties of water, represent an important step towards building a robust computational scheme to investigate the electronic structure of aqueous interfaces, and the position of the water redox potential relative to the band edges of semiconductors and insulators.

5.2 Electronic structure the solvated chloride anion

5.2.1 Introduction

Aqueous solutions of simple salts such as alkali halides and common acids such as sulfates, are of great interest in electrochemistry [234, 235] and atmospheric chemistry [236]. Much progress was made in recent years in the microscopic description of simple aqueous solutions, including the determination of the structure of the solvation shells of cations and anions [237, 238, 236, 239, 240, 234, 235, 241] and of their vibrational properties [242, 243, 244, 245, 246, 247].

On the other hand, electronic structure studies are still limited. Experimentally, recent developments in liquid microjet techniques [248, 249] enabled the measurement of photoelectrons emitted directly from the liquid solutions and constituted a fundamental step forward in understanding the electronic properties of aqueous solutions. From a theoretical standpoint, only few interpretations or predictions [248, 250, 251, 252, 253] of photoelectron spectra appeared in the literature; indeed the microscopic description of aqueous solutions from first principles remains a challenging task, both from a theoretical and computational standpoint, and especially so in the case of anions, whose electronic structure is in general more complex than that of cations [254, 240, 255].

In this work, we consider a chloride anion in aqueous solution and report an analysis of its electronic properties based on first principles molecular dynamics (MD), using several levels of theory; these include semi-local (PBE [256, 257]) and hybrid (PBE0 [258]) functionals and many body perturbation theory (MBPT) [10] within the G_0W_0 approximation [8, 72]. To the best of our knowledge, these are the first calculations of the electronic properties of a solvated ion in water using ab initio MD with hybrid functionals and MBPT. By comparing trajectories obtained with the PBE and the PBE0 functionals, we first identified differences between the structural properties provided by the two

levels of theory, which mainly stem from the average orientation of water molecules in the first anion solvation shell. We then studied how these differences impact the electronic properties of the solution, especially the energy difference between the highest occupied molecular orbital (HOMO) of the anion and the valence band maximum (VBM) of water. We first briefly describe the theoretical and computational methods used in our study, and then discuss our results and compare them with recent photoelectron spectra [248].

5.2.2 Methods

We performed first principles MD simulations of a chloride anion in water using the gradient-corrected PBE functional [256, 257] and the PBE0 hybrid functional [258]. We used supercells containing one chloride ion and 63 D_2O molecules, corresponding to a 0.87 M concentration. The density was fixed at the experimental value of D_2O at ambient conditions, and we then removed one water molecule and added one chloride ion. Our calculations were carried out with the *Qbox* code [89]. We adopted plane wave basis sets and norm-conserving pseudopotentials [259, 260] with a kinetic energy cutoff of 85 Ry. Simulations were carried out in the NVE ensemble with a time step of 10 a.u., within a Born-Oppenheimer framework. After the solution was equilibrated for $\sim$5 ps with either the PBE or the PBE0 functional, the PBE trajectories were collected for 20 ps at an average temperature of 388$\pm$19 K, and the PBE0 ones were collected for $\sim$6 ps with an average temperature of 379$\pm$18 K. We chose a temperature higher than 300 K due to the tendency of liquid water to be overstructured at ambient conditions, when using semi-local [223] and some hybrid functionals [261]. The electronic contributions to the molecular dipole moment were computed using maximally localized Wannier Functions (MLWFs), evaluated at each MD step with the algorithm proposed in Ref. [262]. Electronic structure calculations were carried out at three different levels of theory, using the PBE and PBE0 functionals and within the G_0W_0 approximation [8] with the method

recently developed in Ref. [72, 194], which allowed us to perform calculations for relatively large supercells for several snapshots. In the following, we denote the results of electronic structure calculations using the PBE (PBE0) functionals and PBE0 (PBE) trajectories, as PBE/PBE0 (PBE0/PBE) calculations.

5.2.3 Results and Discussion

We discuss the electronic structure of the solvated chloride anion, and we show how the subtle structural differences between the PBE and PBE0 results observed in our simulations influence the electronic properties of the system. The computed band gap of the solution is 6.55±0.18 eV and 4.18±0.18 eV, in our PBE0/PBE0 and PBE/PBE calculations, respectively. The values were obtained as averages over eight and fifteen configurations extracted from the respective trajectories, and are consistent with previous calculations using smaller cells with 32 water molecules [261]. (As we will discuss below, in our calculations the difference between the band gap of the Cl^- solution and of pure water is at most 0.3 eV). Although a substantial improvement on the PBE result, the PBE0 gap underestimates the value of the quasi particle gap obtained by photoemission experiments (8.7±0.5 eV [263]), and by *GW* calculations (8.1±0.2 eV [264] and 8.7 eV [220, 221]). We note that within statistical errors, we obtained the same value of the gap (6.58±0.21 eV) in PBE0/PBE calculations. This result indicates that the value of the electronic gap of the pristine liquid is not sensitive to the details of the atomic structure and the differences in hydrogen bonding configurations found between PBE and PBE0 samples [261].

On the contrary, we found that the electronic structure of the solvated anion is sensitive to the structural properties of the solution: in particular, the position of the HOMO of the anion with respect to the VBM of water is qualitatively different when using PBE and PBE0 trajectories. It is only when the PBE0 functional is adopted to describe both the trajectories and the electronic structure of the solvated Cl^- that we obtain a HOMO

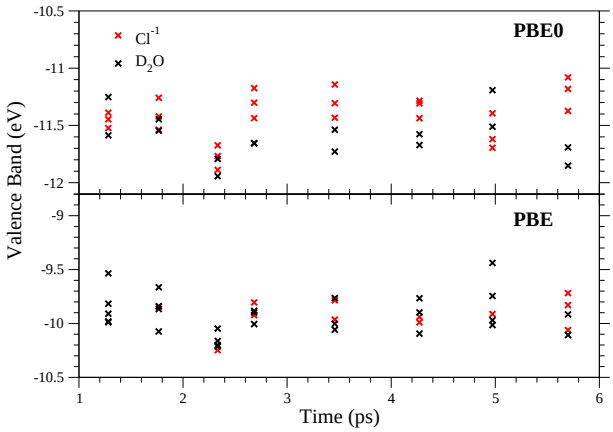

Figure 5.2.1: Top five valence bands of a 0.87 M solution of Cl^- in water calculated with the PBE0 and PBE functionals on a ∼6 ps PBE0 trajectory. The assignment was made based on the localization of the corresponding maximally localized Wannier Functions (MLWFs).

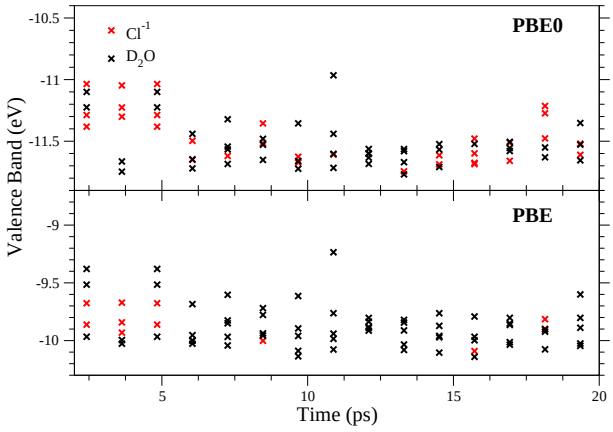

Figure 5.2.2: Top five valence bands computed with the PBE0 and PBE functionals on a 20 ps PBE trajectory. Red squares and black dots represent states assigned to the anion and to water, respectively. The assignment was made based on the localization of the corresponding maximally localized Wannier Functions (MLWFs).

level of the anion on average above the VBM of the liquid. At the PBE level of theory, irrespective of the trajectories used, the HOMO of the anion is below the VBM of water. In PBE0/PBE simulations, the Cl^- HOMO moves slightly upward but not enough to be above the water VBM. These findings are illustrated in Fig. 5.2.1 and Fig. 5.2.2, where we report the five topmost energy levels of the solutions, obtained from PBE0/PBE0 and PBE/PBE0 simulations (left panels) and PBE0/PBE and PBE/PBE simulations (right panels). The energy levels were classified as belonging to water (black circles) or the anion (red squares) based on the analysis of the localization of the square moduli of MLWFs. We note that, in general the single particle states are more delocalized at the PBE level than at the PBE0 level of theory, as illustrated in Fig. S1. This result is consistent with other reports appeared in the literature on Cl^-/H_2O clusters [254] and shows that the PBE0 functional at least partially corrects the delocalization error of the semi-local PBE description.

We also carried out MBPT calculations, within the G_0W_0 approximation, for snapshots extracted from PBE0 trajectories, starting from PBE wavefunctions. We used the techniques proposed in Ref. [72, 194] with 1000 eigenpotentials for the spectral decomposition of the dielectric matrix and the same kinetic energy cutoff as that adopted for the solution of the Kohn-Sham equations (85 Ry). Table SI reports the energy difference between the Cl^- HOMO and the VBM of water for eight snapshots as obtained with PBE, PBE0 and G_0W_0 calculations. Fig. 5.2.3 reports the top five uppermost valence band eigenvalues as computed within G_0W_0. Although the use of MBPT significantly improves on the PBE results, it does not suffice to bring the Cl^- HOMO significantly above the water valence band.

The results on the electronic structure of the anion obtained at the PBE0 and G_0W_0 levels of theory are in qualitative agreement with experiment, however, the energy difference between the Cl^- HOMO and the VBM of water ($\sim$0.23 eV within PBE0 and $\sim$0.08

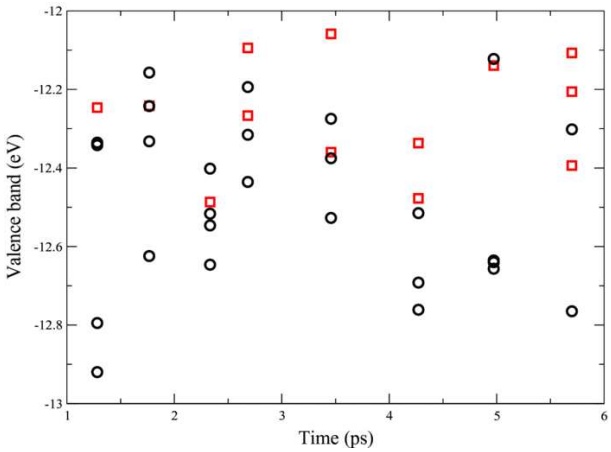

Figure 5.2.3: Top five valence bands of a 0.87 M solution of Cl^- in water calculated using G_0W_0 corrections of PBE eigenvalues for snapshots extracted from a ∼6 ps PBE0 trajectory. Red squares and black dots represent states assigned to the anion and to water, respectively. The assignment was made based on the localization of the corresponding maximally localized Wannier Functions (MLWFs).

eV within G_0W_0) is an underestimate of the experimental value. The measured electron binding energy of a 3 M chloride solution and water are 9.6±0.07 [248] and 11.16±0.04 eV [215], respectively, and the experimental values of the threshold energy of photoelectron emission spectra are 8.81 and 10.06 eV [212] for a 2 M chloride solution and liquid water, respectively. This indicates that the HOMO of Cl^- is located about 1.25−1.56 eV above the VBM of water. We note that the comparison between our calculations and the experiments of Ref. [248] is not straightforward. The photoemission spectra of water and those of NaCl in water reported by Winter et al. [248] show a considerable overlap of the bands of the solution and of pristine water below 10 eV and the onset of the NaCl aqueous solution spectrum is not easy to detect. A weak shoulder is present between 8.8 and 9.6 eV (see Fig. 1 of Ref. [248]). Most importantly, as pointed out in Ref. [248], the spectra of the solution may have contributions from both bulk and surface anions. The structural and electronic properties of the latter differ from those of the bulk anions, and Cl^- anions are known to preferably reside at the surface [265, 266, 267, 268]. Hence the signal detected experimentally may have an overall significant component originating

from surface anions. Finally, we note that the concentration of the solution simulated here (∼0.87 M) is lower than those studied experimentally (2 M [269, 212] and 3M [248]) and the concentration difference may also be partially responsible for the discrepancy with the photoelectron detachment measurements.

5.2.4 Summary

In summary, we presented a study of the electronic properties of the solvated chloride anion using first principles MD and both semi-local (PBE) and hybrid (PBE0) functionals. We showed that it is only when using both PBE0 trajectories and the PBE0 functional to compute electronic properties that the HOMO of the anion is found above the VBM of water. This finding is consistent with photoelectron detachment measurements, although our results underestimate the energy difference between the HOMO of the chloride ion and the VBM of water found experimentally. Work is in progress to investigate the contribution of surface anions to the signals detected in photodetachment experiments, and to explore possible ways to improve on MBPT calculations of aqueous solutions.

Chapter 6

Electronic properties of semiconductors interfaced with liquid water

6.1 Introduction

The photocatalysis of water splitting is a promising way to capture and store solar energy, and is an active research field. In photoelectrochemical cells (PCE), one directly harvests photons to create electron-hole pairs in semiconductor materials and uses these charge carriers in photoelectrochemical reactions to turn water into hydrogen and oxygen, which later may be used as chemical fuels.

The simplest way to construct a PEC cell is to use a single semiconductor material with appropriate band gap and band edge positions [207]. In this case, the optimum value of the solid band gap should be larger than 1.9 eV, a value determined by the energy necessary to split water plus the thermodynamic losses and the overpotential, and smaller than 3.1 eV to fall within the visible range of the solar spectrum [10]. In addition, the semiconductor's

valence band maximum (VBM) and conduction band minimum (CBM) must straddle the water redox potentials to ensure the reactions will be thermodynamically accessible upon photon absorption.

Alternatively, a potentially more efficient design of PEC is based on two semiconductor-liquid junctions: a n-type semiconductor for the photoanode where water oxidation occurs, and a p-type one for the photocathode where water reduction takes place [207, 270]. Each semiconductor provides part of the water splitting potential, and thus semiconductors with smaller band gaps which absorb visible light more efficiently can be utilized to improve the overall conversion efficiency of the device. In this scheme, the thermodynamic constraints are that the CBM of the photocathode must be higher than the water reduction potential H^+/H_2O, and the VBM of the photoanode lower than the water oxidation potential O_2/H_2O, on an absolute scale. Therefore, irrespective of the chosen scheme to build PEC cell, one of the most important information that makes a semiconductor a possible candidate for water splitting is the alignment between its band edge positions and water redox potentials.

From a theoretical point of view, an accurate determination of the alignment between photoeletrodes band edges and water redox potentials requires explicit calculations of the electronic properties of semiconductor-liquid interfaces. This remains a challenging task, as it requires: (i) the generation of molecular dynamics trajectories to obtain structural models of the interfaces; and (ii) the calculation of electronic states of models composed of several hundreds of electrons using advanced electronic structure methods, such as density functional theory (DFT) with hybrid functionals [271] or many-body perturbation theory, e.g., within the *GW* approximation [11]. In particular, *GW* calculations are rather demanding from a computational standpoint, although they are now feasible for systems containing several hundreds of electrons [72, 194]. Due to these difficulties, calculations of the alignment between photoelectrode band edges and water redox potentials usually

involve several approximations as we discuss below.

Most computational approaches to align semiconductor band edges and water redox potentials neglect the effects of the liquid [272, 273, 274, 275]. Band edge positions of a semiconductor in vacuum are computed and aligned to the experimental values [276] of water redox potentials obtained on the absolute scale. The neglect of solid-liquid interfaces allows one to use relatively small supercells, for which a high level of theory, e.g. the *GW* approximation, may be employed to overcome DFT errors on band gaps and band edge positions. This computational procedure has been adopted to investigate numerous candidate materials for water splitting, including transition metal oxides (MnO, FeO, Fe_2O_3, NiO, Cu_2O) [272] and single layer transition metal dichalcogenides MX_2 [273, 274]. An additional assumption which is usually made is that the band gap centers (BGC) computed at the DFT and *GW* levels of theory are the same, thus further reducing the computational cost of *GW* calculations, as the absolute values of quasiparticle energies converge much slower than the value of the quasiparticle band gap [272, 273, 274].

An approach similar to that of Ref. [272, 273, 274] was recently employed to compute band edge positions of a set of 17 materials (both n-type and p-type) of interest for water splitting [275]. To address the effect of liquid water, the authors compared the band edges computed for the materials in vacuum with those measured electrochemically at the point of zero charge (PZC), and suggested that taking into account the interaction with water molecules at interfaces would shift the computed band edges closer to the vacuum level by approximately 0.5 eV, on average. Interestingly, the authors did not resort to the BGC approximation and in fact concluded that such an approximation does not hold for either semiconductors or insulating transition metal oxides.

Calculations of semiconductors band edge alignment with water redox potentials for explicit semiconductor-liquid interfaces have so far been rather sparse. Band edge positions of TiO_2 relative to the normal hydrogen electrode potential were computed in

Ref. [277] using the solvation energy of the H^+ ion as reference. Although the semiconductor-liquid interface was treated explicitly, the electronic properties were obtained using DFT with a semilocal density functional, which yielded substantial errors on the band positions, when compared to experimental values (errors of 0.4 eV and 1.6 eV for the CBM and VBM, respectively). Similar to Ref. [275], the authors also showed that the interaction with water brings the TiO_2 band edge positions closer to the vacuum level, but with a significantly larger shift ($\approx$2.0 eV).

Explicit calculations for semiconductor-liquid interfaces were also reported in Ref. [278] for six photocatalyst materials (TiO_2, WO_3, CdS, ZnSe, GaAs and GaP). The authors computed the band alignment of these systems with the LUMO level of the H_3O^+ ion in a sample of liquid water in contact with the semiconductor, at the DFT level, assuming cancellation of errors would occur between the values obtained for the band positions and the hydronium ion level. Computational procedures to generate interfacial models were not discussed in Ref. [278]. As pointed out later in our our, details of surface structures may significantly affect the semiconductor band edge positions and their alignment with water redox potentials.

The purpose of this work is to present a general computational framework for the calculation of band edge positions of a photoelectrode in liquid water, and their alignment with water redox potentials. Our computational framework includes the effects of the liquid and solid surface structures and hence of the interface, as our samples are generated using *ab initio* molecular dynamics; in addition we employ many-body *GW* calculations for the electronic properties of semiconductor-liquid interfaces. We applied our approach to the calculations of the band edges of functionalized Si(111) surfaces, which have been used as a p-type photoelectrode material for water splitting [207]. In particular we considered three hydrophobic (H−, CH_3−, CF_3−terminated) and one hydrophilic (COOH−terminated) Si(111) substrates.

In agreement with Ref. [279], our results, obtained with a slightly different methodology, showed that surface terminations strongly influence the electronic structure of the solid, i.e., the band edge positions of the Si(111) surface may vary as much as 1.5 eV as a function of the surface termination. Our findings thus confirm those of Ref. [279] indicating that the band edges alignment with water redox potential may be achieved by engineering the surface termination of the semiconductors.

For the three hydrophobic H−, CH_3− and CF_3−terminated Si(111) surfaces, we found that the effect of liquid water on the calculated energy alignment is similar and within 0.5 eV; thus in materials screening studies, trends between different systems with hydrophobic terminations may be determined without explicit consideration of liquid water, if an accuracy less than 0.5 eV is acceptable. In contrast, the effect of water presence is significant for the case of hydrophilic COOH−terminated Si(111) surfaces, and the use of a simple computational scheme neglecting water may lead to substantial errors.

The rest of the chapter is organized as follows. In section 6.2 we discuss a general approach to evaluate band edge positions of semiconductor-liquid water interfaces and computational details. In section 6.3 we present results for Si(111)-water interfaces and we discuss the effect of water and surface structures on the solid band edges. Conclusions and outlook are presented in section 6.4.

6.2 Calculations of Semiconductor Band Edges

In this section, we first briefly summarize the approach to compute band edge positions of a semiconductor in vacuum, widely employed in, e. g., Ref. [272, 273, 274, 275]. We then propose a scheme to compute these quantities for semiconductors in the presence of liquid water. Results obtained with the former approach may be compared with photoemission measurements, while those obtained with the latter are to be compared to electrochemical measurements. We also give computational details.

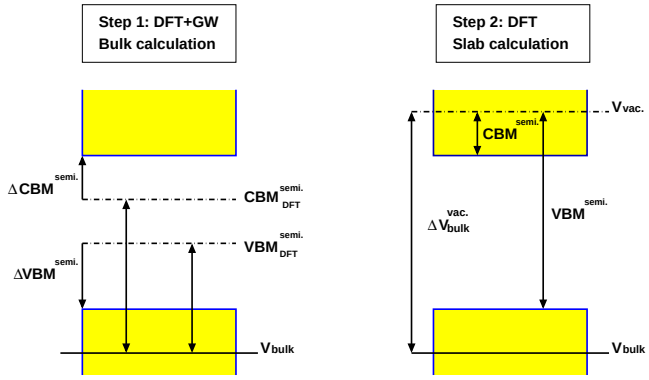

Figure 6.2.1: A computational approach (scheme I) to compute band edge positions (V(C)BM$^{\text{semi.}}$) of a semiconductor in vacuum: V(C)BM$^{\text{semi.}}_{\text{DFT}}$ are band edge positions computed at the DFT level of theory with respect to the average electrostatic potential of the bulk semiconductor in a periodic supercell; ΔV(C)BM$^{\text{semi.}}$ are many-body corrections to the bulk semiconductor band edges, obtained at the G_0W_0 level of theory; and $\Delta \text{V}^{\text{vac.}}_{\text{bulk}} = \text{V}_{\text{vac.}} - \text{V}_{\text{bulk}}$ is the electrostatic potential of the semiconductor relative to vacuum. Theoretical results correspond to photoemission measurements.

Calculations of the VBM and CBM positions of a semiconductor in vacuum are relatively straight forward to carry out, following two separate steps as shown in Fig. 6.2.1. In the first step, band edge positions with many-body corrections are computed with respect to the average electrostatic potentials of the corresponding bulk system. The second step involves the evaluation of the average electrostatic potential of the bulk semiconductor relative to vacuum by employing a slab model with a specific surface structure in contact with an appropriately thick vacuum region. The band edge positions relative to the vacuum level, i.e., band edge positions on an absolute scale, are then computed as:

$$\begin{aligned} \text{VBM} = &\quad \text{VBM}^{\text{semi.}}_{\text{DFT}} + \Delta\text{VBM}^{\text{semi.}} + \Delta\text{V}^{\text{vac.}}_{\text{bulk}} \\ \text{CBM} = &\quad \text{CBM}^{\text{semi.}}_{\text{DFT}} + \Delta\text{CBM}^{\text{semi.}} + \Delta\text{V}^{\text{vac.}}_{\text{bulk}}, \end{aligned} \tag{6.2.1}$$

where VBM$^{\text{semi.}}_{\text{DFT}}$ and CBM$^{\text{semi.}}_{\text{DFT}}$ are band edge positions computed at the DFT level of theory with respect to the average electrostatic potential of the bulk semiconductor in a periodic supercell; ΔV(C)BM$^{\text{semi.}}$ are many-body corrections to the bulk semiconductor

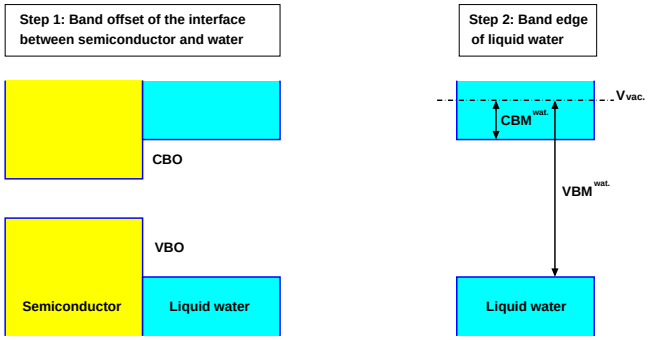

Figure 6.2.2: A computational approach (scheme II) to compute band edge positions of a semiconductor in the presence of liquid water: VBO and CBO are valence and conduction band offsets between liquid water and a semiconductor, respectively; $\mathrm{VBM}^{\mathrm{wat.}}$ and $\mathrm{CBM}^{\mathrm{wat.}}$ are band edges of liquid water relative to vacuum. Theoretical results correspond to electrochemical measurements.

band edges, obtained at the G_0W_0 level of theory. The relative electrostatic potential of the semiconductor relative to vacuum, $\Delta \mathrm{V}^{\mathrm{vac.}}_{\mathrm{bulk}} = \mathrm{V}_{\mathrm{vac.}} - \mathrm{V}_{\mathrm{bulk}}$, is computed within DFT, since in our GW calculations only eigenvalue corrections are computed, without updating the wavefunctions self-consistently (G_0W_0 approximation).

The VBM and CBM obtained from Eq. 6.2.1 correspond to the ionization potential (IP) and electron affinity (EA) of the semiconductor measured in photoemission experiments, in the absence of water.

In constrast, band edge positions measured in electrochemical experiments on an absolute scale provide energies relative to a point in the vacuum just outside the solution [276, 280, 230]. A computational framework to compute these band edge positions is outlined in Fig. 6.2.2 and it involves two steps: (i) calculation of the band offset between liquid water and the semiconductor with a chosen specific surface structure; and (ii) calculation of band edge positions of liquid water relative to vacuum.

The latter (step 2 of Fig. 6.2.2) were computed in Ref. [264] using many-body perturbation theory within the G_0W_0 approximation. Calculated positions of the VBM and CBM of liquid water are 8.8 eV and 0.7 eV below the vacuum level, in qualitative good

agreement with the experimental values of 9.3-9.9 eV and 0.74-1.2 eV [210], respectively. Thus, we are left with the task to carry out band offset calculations at the semiconductor-liquid interface (step 1 of Fig. 6.2.2), which we performed using a bulk-plus-line up procedure [281] that contains two terms, i.e., the bulk band positions, $\mathrm{V(C)BM}^{\mathrm{semi.(wat.)}}$, and the line up term ΔV:

$$\begin{aligned} \mathrm{VBO} &= \mathrm{VBM}^{\mathrm{semi.}} - \mathrm{VBM}^{\mathrm{wat.}} - \Delta\mathrm{V}, \\ \mathrm{CBO} &= \mathrm{CBM}^{\mathrm{semi.}} - \mathrm{CBM}^{\mathrm{wat.}} - \Delta\mathrm{V}. \end{aligned} \tag{6.2.2}$$

Here $\mathrm{V(C)BM}^{\mathrm{semi.(wat.)}}$ are band edge positions of semiconductor and liquid water with respect to the average electrostatic potentials of the corresponding bulk systems, while ΔV is the discontinuity of the average electrostatic potential across the semiconductor-liquid interfaces. The line up term ΔV is computed within DFT for the whole interface containing the semiconductor and water; to correct the DFT errors on band offsets, we included G_0W_0 corrections obtained from bulk calculations on VBM and CBM eigenvalues, following the scheme described in Ref. [282].

6.3 Application to Si(111) surfaces

6.3.1 Computational Details

Calculations of IPs and EAs of the functionalized Si(111) surfaces in vacuum were performed using scheme I (Fig. 6.2.1). We considered the H−terminated Si(111) surface, which has been extensively studied by experiments [283, 284], and additionally, the CH_3-, CF_3- and COOH−terminated Si(111) surfaces, as chemically stable Si-C bonds were shown to passivate Si surfaces both electrically and chemically while allowing for secondary functionalization [285, 286, 287, 288]. For all surfaces we used full adsor-

bate coverage; we note that experimentally, the maximum coverage obtained for the COOH−terminated Si(111) surface is ≈ 0.5 [289], with the rest of the surface Si bonds terminated by H atoms; hence the results presented here for the relative position of band edges and water redox potentials represent a model hydrophilic surface. For Si we used a lattice constant of $a = 5.48$ Å, obtained at the DFT/PBE level of theory, to model the bulk and slabs. All DFT calculations were performed with the PBE approximation for the exchange and correlation functional [7], and the interaction between valence electrons and ionic cores was represented by norm-conserving pseudopotentials. We used a kinetic energy cutoff of 85 Ry to represent wavefunctions.

Band edge positions of functionalized Si(111) surfaces in water were computed using scheme II (Fig. 6.2.2). For each solid/liquid interface, we carried out Car-Parrinello *ab initio* MD simulations with interatomic forces derived from DFT/PBE and we obtained trajectories over ≈ 30 ps. We used ultrasoft pseudopotentials and the electronic wavefunctions and charge densities were expanded in a plane-wave basis set truncated at a cutoff energy of 25 Ry and 180 Ry, respectively. We used an effective electronic mass of $\mu = 340$ a.u. which was shown to be an appropriate value when investigating the structural properties of the liquid at ambient conditions [223, 224]. In all simulations we used 108 heavy water molecules in order to maximize the allowable time step, which was chosen equal to 3 a.u. A six-layer slab of 72 Si atoms at the equilibrium volume was used to model the surface with periodic boundary conditions. All MD simulations were performed with the Quantum-ESPRESSO code [145].

First-principles simulations were carried out at constant temperature (NVT conditions) with T= 375 K. The neglect of quantum zero-point effects [125] of light nuclei and the use of the PBE approximation were shown to yield an overstructured liquid in *ab initio* simulations of water, and an elevated simulation temperature of T$\approx$ 400 K is necessary to recover the experimental structure and diffusion coefficients at T= 300 K [223, 224].

Snapshots used in the band offset calculations were extracted every $\approx$ 0.5 ps from the NVT simulations; we averaged our results over $\approx$30 snapshots. Calculations of the electronic properties were then carried out using the same norm-conserving pseudopotentials as in the calculations for Si(111) surfaces in vacuum. Further information on the structural and dynamical properties of liquid water at the surfaces studied here were dicussed in Ref. [290].

Calculations of many-body corrections were carried out within the G_0W_0 approximation, using a spectral decompostions of the dielectric matrix [72, 194]. Corrections to the band edges of bulk Si and liquid water were obtained from Ref. [291] and Ref. [264], respectively.

6.3.2 Results and Discussions

We first present our results for IPs and EAs of different Si(111) surfaces in vacuum, and compare them with photoemission experiments. We then present our calculations of the band edge positions of Si(111) surfaces in liquid water and we discuss the effect of surface structure and the presence of the liquid on the computed band edge positions.

Ionization Potentials and Electron Affinities of Functionalized Si(111) surfaces in Vacuum

Table 6.1 shows the computed IPs and EAs of three Si(111) surfaces (H, CH_3, CF_3) obtained at the DFT and G_0W_0 levels of theory. For H$-$ and CH_3-terminated Si(111) surfaces, for which experimental results are available, we found that IPs computed within DFT severely underestimated experimental data. The use of G_0W_0 many-body corrections significantly improved the calculated values of the IPs, i.e., the discrepancy with experiments was reduced within 0.3 eV.

Our results for H$-$ and CH_3-terminated Si(111) surfaces are in good agreement with

Table 6.1: Ionization potential (IP) and electron affinity (EA) of H−, CH_3−, and CF_3−terminated Si(111) surfaces computed using Density Functional Theory (DFT) with the PBE functional and the G_0W_0 approximation. Results from photoelectron spectroscopy (Expt.) are also included when available [285, 279]. Theoretical results obtained with the LDA approximation are given in the parentheses [279].

	DFT	G_0W_0	Expt.[285, 279]
H−			
IP	4.76 (4.83)	5.26 (5.46)	5.29-5.31
EA	4.02 (4.10)	3.92 (3.93)	
CH_3−			
IP	3.96 (4.06)	4.46 (4.71)	4.76-4.80
EA	3.22 (3.36)	3.12 (3.21)	
CF_3−			
IP	6.32	6.82	
EA	5.58	5.48	

those of a previous theoretical study [279]. A small difference ($\approx 0.20 - 0.25$ eV) in the G_0W_0 result for the IP of the H− and CH_3−terminated Si(111) surfaces may originate from the use of different exchange-correlation functionals, as the DFT wavefunctions and eigenvalues were obtained using the local density approximation (LDA) in Ref. [279]. In addition, G_0W_0 calculations were directly obtained from slab models in Ref. [279], while here, to avoid carrying out G_0W_0 of thick slabs in large supercells, we estimated G_0W_0 corrections to the DFT band edges of Si(111) surfaces from bulk calculations [291]. In agreement with Ref. [292], we found that VBM and CBM states of the Si(111) surfaces considered here are bulklike, thus validating a posteriori our choice of bulk G_0W_0 corrections.

In addition, we found that the G_0W_0 corrections to the CBM and VBM are opposite in sign and of different magnitude. Similar conclusions applied to liquid water where the G_0W_0 correction to the VBM (≈ 2.7 eV) is significantly larger than that to the CBM (≈ 1.2 eV). This indicates that the BGC approximation does not hold, neither for Si nor for liquid water, consistent with the results of Ref. [275].

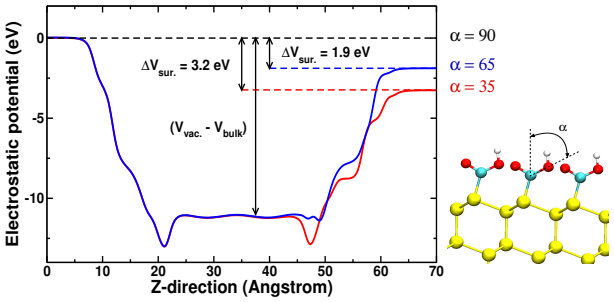

Figure 6.3.1: Electrostatic potential along the z-direction of Si(111) surfaces terminated by different COOH orientations. Orientation of the COOH group is defined in term of the angle (α) that the single C−O bond forms with the direction normal to the surface. The vacuum level of the surface corresponing to $\alpha = 90.0°$ is set to zero. The difference in the electrostatic potential at a point in vacuum and at the center of the slab defines the relative electrostatic potential $\Delta V^{vac.}_{bulk} = V_{vac.} - V_{bulk}$ in Eq. 6.2.1.

As expected, the absolute band edge positions of Si(111) surfaces strongly depend on their termination. In particular, the IP of the H−terminated Si(111) surface is 0.8 eV and 1.3 eV lower than those of the CH_3− and CF_3−terminated ones, respectively. As discussed in Ref. [279] the shift of the band edges is due to the change in the surface potential, which is related to the surface dipole moment (μ_z):

$$\Delta V_{sur.} = -4\pi\mu_z/A, \tag{6.3.1}$$

where A is the surface area. In particular, the electron-donating group $-CH_3$ leads to a positive value of μ_z, and thus decreases the IP and EA with respect to those computed for the H−terminated surface. In contrast, the electron-withdrawing group $-CF_3$ results in a negative value of μ_z and in an increase of the IP and EA values. The COOH−terminated Si(111) surface exhibits a complex structure. We considered different surface arrangements of the −COOH group, defined in terms of the angle (α) that the single C−O bond forms with the direction normal to the surface.

Fig. 6.3.1 shows that changing α has a profound effect on the relative electrostatic

Table 6.2: Band edge positions of H−, CH_3−, CF_3− and COOH−terminated Si(111) surfaces in the presence of liquid water, as computed at the DFT/PBE and G_0W_0 levels of theory. Results from electrochemical measurements are also included when available [283].

	DFT	G_0W_0	Expt.[283]
H−			
VBM	4.49	4.99	5.10
CBM	3.75	3.65	3.98
CH_3−			
VBM	3.62	4.12	
CBM	2.88	2.78	
CF_3−			
VBM	5.88	6.38	
CBM	5.14	5.04	
COOH−			
VBM	4.23	4.73	
CBM	3.49	3.39	

potential $\Delta V^{vac.}_{bulk}$ (see Eq. 6.2.1), and thus on the absolute position of the solid band edges. In particular, $\Delta V^{vac.}_{bulk}$ is shifted by as much as 3.2 eV when the angle α is varied from 90.0° to 35.0°. This large variation originates from the modification in the value of the $z-$component of the surface dipole moment μ_z.

These results indicate that in order to accurately determine the absolute band edge positions of the COOH−terminated Si(111) surface in solution, it is crucial to explicitly perform simulations of the COOH−Si(111)-water interface and determine the surface structure in the presence of the liquid.

Band Edge Positions of Functionalized Si(111) Surfaces in Liquid Water

The band edge positions of different functionalized Si(111) surfaces in liquid water, computed at the PBE and G_0W_0 levels of theory with Eq. 6.2.2 are shown in Tab. 6.2 and as green lines in Fig. 6.3.2. The band offset was computed as the difference between the VBM (CBM) in the bulk-like Si and water regions of the interface.

We verified that band offsets obtained from Eq. 6.2.2 agree with those obtained from the computed Local Density of States (LDOS) within 0.3 eV. We obtained LDOS as [291]:

$$D(\epsilon, z) = 2 \sum_{n} |\langle z|\psi_n\rangle|^2 \delta(\epsilon - \epsilon_n), \tag{6.3.2}$$

where z is the direction perpendicular to the interface, the factor of 2 accounts for the spin degeneracy and ψ_n, ϵ_n are the Kohn-Sham eigenstates and eigenvalues, respectively. The band offset was computed as the difference between the VBM (CBM) in the bulk-like Si and water regions of the interface. We found that the ≈0.3 eV difference comes almost entirely from an overestimate of the Si bulk band gap obtained with the LDOS procedure, due to the use of a relatively thin Si slab composed of six layes. Indeed, it was shown in Ref. [293] that a thickness of at least 30 monolayers is required to recover the bulk band gap. In addition, we found that the electronic properties, e.g., the electronic gap, of bulk liquid water, is recovered within $\approx 5 - 7$ Å from the Si(111) surfaces. Hence, from an electronic standpoint, our LDOS calculations indicated an interface thickness of $\approx 5 - 7$ Å.

Our results are in satisfactory agreement (within $\approx 0.1 - 0.3$ eV) with experimental results [283] for the H−terminated Si(111) surface for which electrochemical measurements at the point of zero charge are available. As expected, similar to calculations for Si surfaces in vacuum, the surface termination has significant effects on the band edge positions of the solid in liquid water. In particular, we found that the CBM of H− and CH_3−terminated Si(111) surfaces are higher than the standard hydrogen electrode (SHE), while that of the CF_3−terminated Si(111) surface is lower. The hydrophilic COOH− terminated surface exhibit a slight shift (≈0.25 eV) of the band edges compared to those of the H− one, and its CBM is higher than the SHE.

We note that the VBM and CBM of water computed with the G_0W_0 approximation and PBE wavefunctions are 8.8 eV and 0.7 eV below the vacuum. Using, e.g., PBE0

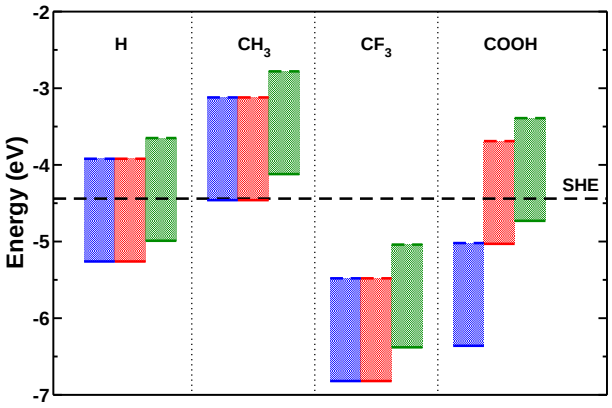

Figure 6.3.2: Valence band maxima (solid lines) and conduction band minima (dotted lines) of the Si(111) surface functionalized with various groups shown on the x−axis, as computed with the G_0W_0 approximation. Blue and red lines indicate results obtained with scheme I and surface geometries optimized in the absence of liquid water, and for geometries extracted from molecular dynamics simulations performed in the presence of liquid water. For the COOH−terminated surface, we used the configuration with $\alpha = 90°$ (see Fig. 6.3.1). The green lines indicate results obtained with scheme II; the dashed line indicates the experimental value of the standard hydrogen potential (SHE).

wavefunctions as input for G_0W_0 calculations may improve the agreement with experimental results of 9.3−9.9 eV for the VBM, while the CBM position would be weakly modified as the LUMO of water is a delocalized state. However, we expect that the results presented in Tab. 6.2 are not affected by the change in G_0W_0 corrections due to wavefunction effects, as these corrections appear in both two steps: calculation of the band offset between water and the semiconductor; and calculation of band edge positions of liquid water relative to vacuum.

Aligning the band edge positions reported in Tab. 6.2 with water redox potentials requires knowdlege of the PZC as water redox potentials at room temperature vary as a function of the pH, following the Nernst equation [294]:

$$\mathrm{E}_{\mathrm{red}}^{\mathrm{pH}} = \mathrm{E}_{\mathrm{red}}^{\mathrm{pH=0}} - 0.059 \times \mathrm{pH}, \tag{6.3.3}$$

where $\mathrm{E}_{\mathrm{red}}^{\mathrm{pH}}$ is the redox potential measured on the SHE scale. However, qualitative con-

clusions may be drawn. The CBM of the H−terminated Si(111) surface is higher than the water reduction potential at the PZC (≈2.2), and thus it is suitable for hydrogen evolution. The same conclusions hold for the CH_3− and COOH−terminated surfaces, as their CBM positions are higher than the water reduction potential, even at a very high value of pH, e.g. pH= 14. In contrast, the CF_3−terminated surface is not suitable for the hydrogen evolution reaction, as its CBM position is always lower than the water reduction potential. In addition, only the CF_3−terminated surface is suitable for oxygen evolution reaction as its VBM position is lower than the water oxidation potential at any pH value ranging from 0 to 14.

The presence of liquid water alters the position of the solid band edges computed in vacuum in several ways: (i) the interaction with liquid water may modify the surface structure and hence the surface dipole; and (ii) interfacial water molecules may form an additional dipole layer, counteracting or enhancing the solid surface dipole. To identify the magnitude of effect (i), we extracted Si(111) geometries from our MD simulations carried out in the presence of water, and we computed band edge positions using scheme I, and we obtained the results shown by the red lines in Fig. 6.3.2. Therefore, comparison between red and blue lines highlights the effect, on the electronic states of the surface in vaccum, of the structural modifications caused by the presence of the liquid; the difference between red and green lines reveals instead the effect of the change in the magnitude and direction of the interfacial dipole due to solid-liquid interaction.

For the three hydrophobic surfaces, we found negligible differences between results obtained with different geometries, indicating that hydrophobic Si(111) surface structures are weakly affected by the presence of liquid water, and the difference between the results computed with schemes I and II stems from the interfacial dipole only. Taking into account the interaction with liquid water shifts the band edge positions of the three hydrophobic surfaces closer to the vacuum level by $\approx 0.3 - 0.5$ eV (see blue and green

lines), and such a small difference implies the presence of a rather weak interfacial dipole moment. Overall, our analysis indicates that if the accuracy required in screening surfaces for optimal photoelectrodes is within 0.5 eV, effect of the solution may be neglected in the case of hydrophobic surfaces.

In contrast, both surface modification and the interface dipole play an important role in the case of the hydrophilic COOH-terminated Si(111) surface. In particular, band edges computed with the surface geometry extracted from MD trajectory are significantly shifted closer to vacuum (≈1.3 eV) when compared to those determined with the one obtained without liquid water and $\alpha = 90°$[1] (see Fig. 6.3.1). This analysis clearly indicates that the electronic structure of hydrophilic surfaces may be remarkably modified in the presence of liquid water.

In addition to the effect of surface modification, there is an additional effect of liquid water, which further shifts the band edge of the COOH-terminated Si(111) surface closer to vacuum by ≈0.3 eV. Interestingly, the magnitude and direction of band edge shifts would suggest the presence of similar interfacial dipoles in the case of the hydrophobic and hydrophilic surfaces considered here.

In order to investigate the intrinsic dipole of water molecules at the interfaces, we computed the OH bond tilt angle (θ), defined as the angle that a vector along an OH bond forms with the direction perpendicular to the surface (see Fig. 6.3.3). The probability distribution of θ as a function of the angle and the distance from the outermost surface Si atom is shown in Fig. 6.3.3.

We found that the three hydrophobic surfaces exhibit similar angular probability distributions. In particular, at about ≈2.1 Å from the surfaces, we observed maxima at 60 and 160 degrees, orientations corresponding to OH bonds pointing slightly outwards and towards the surface, respectively. This indicates that the molecular dipoles tend to point

[1]The COOH−terminated surface with $\alpha = 90°$ was chosen as it is more energetically favorable than the configurations with $\alpha = 35°$ and $\alpha = 65°$.

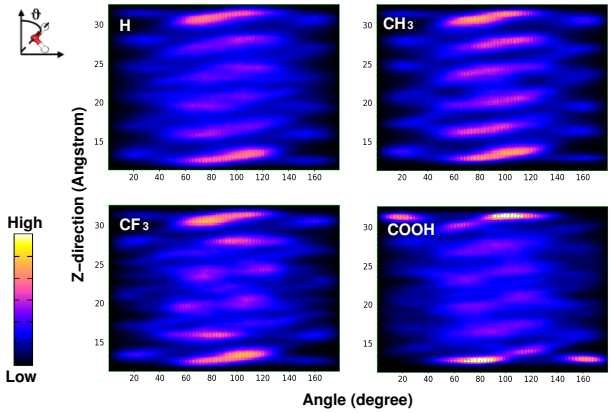

Figure 6.3.3: Probability of finding different orientations of water OH bond vector as a function of the distance from the outermost surface Si atom, computed for H−, CH_3−, CF_3−, and COOH−terminated surfaces.

away from the hydrophobic surfaces. Our argument is further validated by examining the probability of specific molecular orientations at the interface (see Fig. 6.3.4). The formation of an interfacial water layer with a positive dipole thus explains why the band edges of hydrophobic surfaces shift toward the vacuum level, in the presence of water.

In contrast, the water orientation at the interface with the hydrophilic COOH− Si(111) surface exhibits a noticeable different character. Fig. 6.3.3 indicates that in addition to a peak at 60°, the OH bond distribution shows a strong peak at 160°. Furthermore, we found that water molecules exhibit a large probability of having dipole moment pointing toward the surface (maroon rectangles in Fig. 6.3.4), i.e. negative dipole moments, instead of positive ones, as in the case of hydrophobic surfaces.

Thus one would expect a shift of the solid band edge positions in the opposite direction with respect to that of hydrophobic surfaces, which is not what we found in Fig. 6.3.2. We further examined the electronic density at the interface, in particular we computed:

$$\Delta\rho_e = \rho_e(solid/liquid) - \rho_e(solid) - \rho_e(liquid). \tag{6.3.4}$$

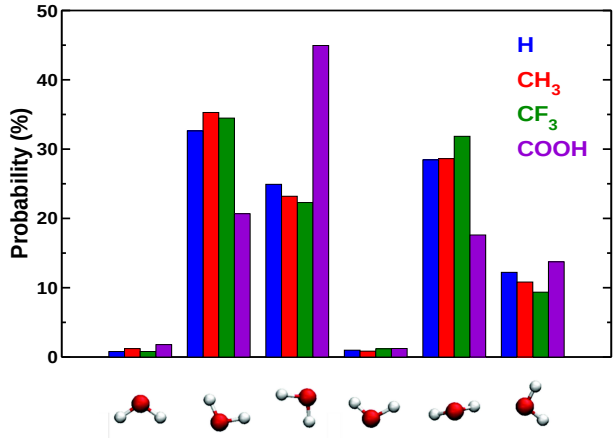

Figure 6.3.4: Probability of specific orientation of water molecules in the interfacial region. Molecular orientation of individual waters is classified in six different categories, depending on the orientation of two OH vectors with respect to the Si(111) surface.

The quantity $\Delta\rho_e$ is positive when there is an addition of charge to the system with respect to the isolated fragments, and negative in the opposite case.

In Fig. 6.3.5, we plot the planar average electronic density difference $\Delta\rho_e$ for a particular snapshot of the COOH−terminated Si(111) surface-water interface. The main changes in the charge density are localized directly at the interface, and there is a charge transfer from liquid water towards the semiconductor. This charge transfer leads to the creation of an additional dipole layer with a positive dipole moment that shifts the solid band edges closer to vacuum. In the case of the COOH−terminated surface, the charge transfer is clearly more significant than for the H−terminated one (inset of Fig. 6.3.5) and overcompensates the effect of the intrinsic dipole moment of the water layer, leading to a net shift closer to vacuum. A similar behavior has been observed for the work function of metals, e.g., Pt and Ru, interfaced with water [295].

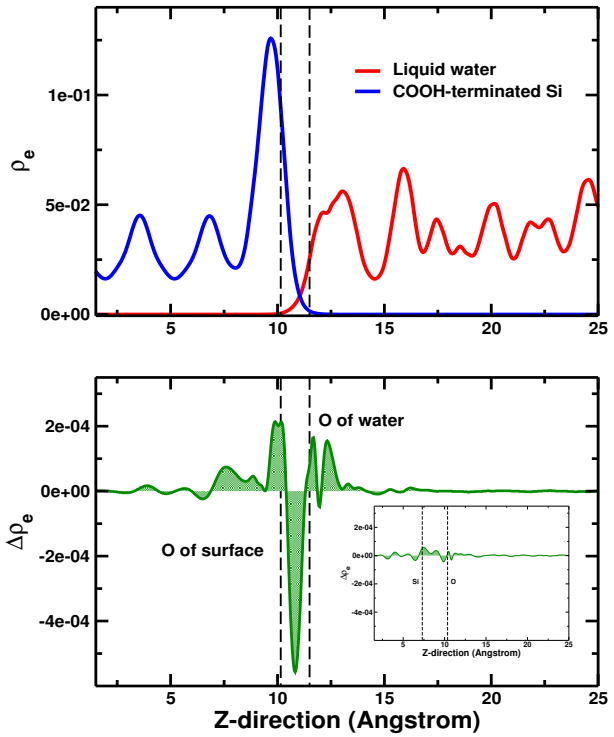

Figure 6.3.5: Charge transfer at the COOH-terminated Si(111)-water interface. The inset shows the same quantity computed for the H-terminated Si(111)-water interface.

6.4 Summary

In summary, we present a general computational framework for the calculation of band edge positions of a photoelectrode in liquid water, and their alignment with water redox potentials. Our computational framework includes the effects of the liquid and solid surface structures, and we employ accurate many-body *GW* calculations for the electronic properties of semiconductor-liquid interfaces.

Application of our approach to the calculations of the band edges of functionalized Si(111) surfaces shows that: (i) in agreement with the previous study [279], the absolute band edge positions of the surfaces strongly depends on their termination; (ii) the effect of liquid water on band edges of hydrophobic surfaces is similar and within 0.5 eV, thus indicates that in materials screening studies, trends between different systems with hy-

drophobic terminations may be determined without explicit consideration of liquid water, if an accuracy less than 0.5 eV is acceptable; (iii) the effect of water may be significant for hydrophilic surfaces, and the use of a simple computational scheme neglecting water may lead to substantial errors; (iv) band edge positions of the surfaces, in particular of the hydrophilic one, are not only affected by intrinsic dipole of interfacial water molecules, but also by the charge rearrangement at the interfaces; (v) many body GW corrections are essential to obtain band edge positions in good agreement with experiment; (vi) the band gap center approximation does not hold neither for Si or liquid water, e.g., it fails drastically for liquid water. We note that results presented here correspond to electrochemical measurements at the PZC, work is in progress to simulate photoelectrodes in contact with water under different pH conditions, and to compute the corresponding band edge alignment.

Conclusions

With the aim of devising new theoretical and computational techniques to describe heterogeneous interfaces, the work presented in this thesis represent progress in: (i) theoretical and computational developments in electronic structure; and (ii) understanding of the physical and chemical properties of complex materials interfaces.

Theoretical and computational developments

In this thesis I presented an efficient and accurate way to carry out GW calculations for realistic systems, based on the Projective Dielectric EigenPotential method for Green's function formulation of the many-body problem. The algorithm devised in the thesis offers several advantages over those currently employed in the literature: i) explicit summations over unoccupied single particle states are not required to obtain the Green's function or dielectric matrix; ii) the storage and inversion of large dielectric matrices are avoided; iii) the use of plasmon-pole models is not necessary; and iv) the numerical accuracy of the calculation of single particle energies controlled by a single parameter, i.e., the number of dielectric eigenvalues and eigenvectors in the spectral decomposition of the dielectric matrix. These advantages open the possibility to compute the electronic properties of realistic systems with the GW approach.

In addition to the theoretical formulation summarized above, I have been one of the main developers of the first version of a GW code (WEST). The WEST code was written in

Fortran90 and interfaced with the open-source Quantum-ESPRESSO package. The code was used by several members of the Galli group at the University of California Davis (now at the University of Chicago), and by several of our collaborators over the world. Further development of the WEST code is being continued, with the goal of making the final product an open source code.

Physical and chemical properties of complex material interfaces

In this thesis I used a combination of theoretical and computational techniques to provide a microscopic description of heterogeneous interfaces, which allowed for a direct comparison with experiments.

I used DFT and DFT-based first-principles molecular dynamics to provide insights into experimental measurements, e.g., to interpret the vibrational spectra of the ice Ih surface and Al_2O_3/water interface, and to explain the variation of the dielectric constant of nitride thin films as a function of thickness. Ground state properties were found to be in good agreement with experiments.

The development of the *GW* approach based on the spectral decomposition of the dielectric matrix allowed us to *quantitatively* describe excited state properties of hetegeneous interfaces. For examples, this thesis showed that the use of the *GW* method improves the comparison with photoemission experiments for, e.g., the band offsets of the Si_3N_4/Si(001) and Si/water interfaces, and of the band gaps and band edge positions of liquid water and aqueous solutions.

The *GW* approach developed here allowed us, for the first time, to investigate the electronic properties of realistic heterogeneous interfaces and to establish relationships between local interface structures, electronic properties of semiconductors and their reactivity in aqueous solutions. An example presented in this thesis is that of functionalized Si surfaces interfaced with water, in which we studied the effect of surface functionalization

on the alignment between Si band edges and water redox potentials, and we suggested a possible approach to engineer and design semiconductor surfaces for photoelectrochemical water splitting.

Future work

The work presented in this thesis represents the very first step towards the development of a general computational framework for describing heterogeneous interfaces. There are a number of outstanding challenges and issues raised by the present study, which require additional theoretical and computational developments.

First, although the GW approach developed in this work is applicable to systems with hundreds of electrons, further developments to optimize the GW code developed in this thesis are necessary so as to apply the approach to interfacial systems with thousands of atoms.

Secondly, all GW calculations presented in the thesis were performed with the G_0W_0 approximation, and it would be interesting to go beyond this approximation, and evaluate the effect of the choice of the starting wavefunctions on the energy alignment at hetegeneous interfaces.

Thirdly, *ab-initio* simulations to obtain models of heterogeneous interfaces are still challenging, as the level of theory necessary to describe, e.g., chemical bonds and adsorption events at the interfaces is yet unclear. Levels of theory more sophisticated than semi-local density functionals, such as DFT with hybrid functionals or with many-body dispersion interaction corrections may be necessary to investigate structural properties of complex materials interfaces.

As one could see, there are a lot of additional developments and calculations required to understand the physics and chemistry occuring at heterogeneous interfaces. However, any thesis should have its last sentence. I stop my writing here and I am already starting

developments in future directions.

Livermore, California, July 2014

List of publications

Publications related to the thesis

1. **T. Anh Pham**, D. Lee, E. Schwegler and G. Galli, "Interfaces of Semiconductor Photoelectrodes with Liquid Water: Electronic Properties from First-Principles Calculations", (Submitted).

2. **T. Anh Pham**, C. Zhang, E. Schwegler and G. Galli, "Probing the electronic structure of liquid water with many-body perturbation theory", Phys. Rev. B **89**, 060202(R) (2014).

3. P. Huang, **T. Anh Pham**, G. Galli and E. Schwegler, "Alumina(0001)/Water interface: structural properties and infrared spectra from first-principles molecular dynamics simulations", J. Chem. Phys. C **108**, 8944 (2014).

4. **T. Anh Pham**, T. Li, Huy-Viet Nguyen, S. Shankar, F. Gygi and G. Galli,"Band offset and dielectric properties of the amorphous Si_3N_4/Si(001) interface: a first-principles study", Appl. Phys. Lett. **102**, 241603 (2013).

5. C. Zhang, **T. Anh Pham**, F. Gygi and G. Galli, "Electronic structure of solvated chloride anion solution from first-principles calculations", J. Chem. Phys. **138**, 181102 (2013).

6. **T. Anh Pham**, Huy-Viet Nguyen, D. Rocca and G. Galli, "*GW* calculations using

the spectral decomposition of the dielectric matrix: Verification, validation, and comparison of methods", Phys. Rev. B **87**, 155148 (2013).

7. **T. Anh Pham**, P. Huang, E. Schwegler and G. Galli, "First-principles study of the infrared spectra of the ice Ih(0001) surface", J. Phys. Chem. A **116**, 9255 (2012).

8. Huy-Viet Nguyen, **T. Anh Pham**, D. Rocca and G. Galli, "Improving accuracy and efficiency of calculations of photoemission spectra within many-body perturbation theory", Phys. Rev. B **85**, 081101(R) (2012).

9. **T. Anh Pham**, T. Li, S. Shankar, F. Gygi and G. Galli, "Microscopic modeling of the dielectric properties of silicon nitride", Phys. Rev. B **84**, 045308 (2011).

10. **T. Anh Pham**, T. Li, S. Shankar, F. Gygi and G. Galli, "First principles investigations of the dielectric properties of crystalline and amorphous silicon nitride thin films", Appl. Phys. Lett. **96**, 062902 (2010).

Publications not included in the thesis

1. D. Opalka, **T. Anh Pham**, M. Sprik and G. Galli, "The ionization potential of aqueous hydroxide computed using many-body perturbation theory", J. Chem. Phys. **141**, 034501 (2014).

2. A. Kaur, E. R. Ylvisaker, D. Lu, **T. Anh Pham**, G. Galli and W. E. Pickett, "Spectral representation analysis of dielectric screening in solids and molecules", Phys. Rev. B **87**, 155144 (2013).

3. **T. Anh Pham**, R. Gebauer and S. Scandolo, "Magnetism and vibrations in the phase epsilon of oxygen", Solid State Commun. **149**, 160 (2009).

4. **T. Anh Pham**, V. Hung Nguyen and V. Lien Nguyen, "Phonon-assisted tunneling and shot noise in double barrier structures in a longitudinal magnetic field", Phys.

Lett. A **372**, 4947 (2008).

5. V. Hung Nguyen, V. Lien Nguyen and **T. Anh Pham**, "Current and shot noise in double barrier resonant tunneling structures in a longitudinal magnetic field", Phys. Rev. B **76**, 235326 (2007).

References

[1] G. W. Crabtree and N. S. Lewis. *Physics Today* **60**, 309 (2008).

[2] N. S. Lewis and D. G. Nocera. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 15729 (2006).

[3] Y. Wu, P. Lazic, G. Hautier, K. Persson, and G. Ceder. *Energ. Environ.* **6**, 157 (2013).

[4] P. Liao and E. A. Carter. *Chem. Soc. Rev.* **42**, 2401 (2013).

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

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

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

[8] L. Hedin. *Phys. Rev.* **139**, A796 (1965).

[9] M. R. Martin. Cambridge University Press, Cambridge, 2004.

[10] Y. Ping, D. Rocca, and G. Galli. *Chem. Soc. Rev.* **42**, 2437 (2013).

[11] G. Onida, L. Reining, and A. Rubio. *Rev. Mod. Phys.* **74**, 601 (2002).

[12] M. S. Hybertsen and S. G. Louie. *Phys. Rev. B* **34**, 5390 (1986).

[13] M. Jain, J. R. Chelikowsky, and S. G. Louie. *Phys. Rev. Lett.* **107**, 216806 (2011).

[14] R. Shaltaf, G. M. Rignanese, X. Gonze, F. Giustino, and A. Pasquarello. *Phys. Rev. Lett.* **100**, 186401 (2008).

[15] K. S. Thygesen and A. Rubio. *Phys. Rev. B* **77**, 115333 (2008).

[16] P. Umari, G. Stenuit, and S. Baroni. *Phys. Rev. B* **79**, 201104 (2009).

[17] B.-C. Shih, Y. Xue, P. Zhang, M. L. Cohen, and S. G. Louie. *Phys. Rev. Lett.* **105**, 146401 (2010).

[18] G. Samsonidze, M. Jain, J. Deslippe, M. L. Cohen, and S. G. Louie. *Phys. Rev. Lett.* **107**, 186404 (2011).

[19] I. Tamblyn, P. Darancet, S. Y. Quek, S. A. Bonev, and J. B. Neaton. *Phys. Rev. B* **84**, 201402 (2011).

[20] P. Umari, G. Stenuit, and S. Baroni. *Phys. Rev. B* **81**, 115104 (2010).

[21] F. Giustino, M. L. Cohen, and S. G.. Louie, S. G. *Phys. Rev. B* **81**, 115105 (2010).

[22] J. A. Berger, L. Reining, and F. Sottile. *Phys. Rev. B* **82**, 041103 (2010).

[23] C. Friedrich, M. C. Müller, and S. Blügel. *Phys. Rev. B* **83**, 081101 (2011).

[24] W. von der Linden and P. Horsch. *Phys. Rev. B* **37**, 8351 (1988).

[25] R. W. Godby and R. J. Needs. *Phys. Rev. Lett.* **62**, 1169 (1989).

[26] G. E. Engel and B. Farid. *Phys. Rev. B* **47**, 15931 (1993).

[27] M. Stankovski, G. Antonius, D. Waroquiers, A. Miglio, H. Dixit, K. Sankaran, M. Giantomassi, X. Gonze, M. Côté, and G.-M. Rignanese. *Phys. Rev. B* **84**, 241201 (2011).

[28] A. Baldereschi and E. Tosatti. *Solid State Comm.* **29**, 131 (1979).

[29] R. Car, E. Tosatti, S. Baroni, and S. Leelaprute. *Phys. Rev. B* **24**, 985 (1981).

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

[31] S. Galamic-Mulaomerovic, C. D. Hogan, and C. H. Patterson. *Phys. Stat. Sol.* **188**, 1291 (2001).

[32] H. F. Wilson, F. Gygi, and G. Galli. *Phys. Rev. B* **78**, 113303 (2008).

[33] H.-V. Nguyen and S. de Gironcoli. *Phys. Rev. B* **79**, 205114 (2009).

[34] H. F. Wilson, D. Lu, F. Gygi, and G. Galli. *Phys. Rev. B* **79**, 245106 (2009).

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

[36] F. Gygi and A. Baldereschi. *Phys. Rev. B* **34**, 4405 (1986).

[37] D. Lu, F. Gygi, and G. Galli. *Phys. Rev. Lett.* **100**, 147601 (2008).

[38] Y. Saad. *Iterative Methods for Sparse Linear Systems*. SIAM, Philadelphia, second edition, 2003.

[39] D. W. Bullet, R Haydock, V. Heine, and M.J. Kelly. volume 35 of *Solid State Physics*. Academic Press, New York, 1980.

[40] D. Rocca, R. Gebauer, Y. Saad, and S. Baroni. *J. Chem. Phys.* **128**, 154105 (2008).

[41] J. B. Neaton, M. S. Hybertsen, and S. G. Louie. *Phys. Rev. Lett.* **97**, 216405 (2006).

[42] M. L. Tiago and J. R. Chelikowsky. *Phys. Rev. B* **73**, 205334 (2006).

[43] G. Onida, L. Reining, R. W. Godby, R. Del Sole, and W. Andreoni. *Phys. Rev. Lett.* **75**, 818 (1995).

[44] H. N. Rojas, R. W. Godby, and R. J. Needs. *Phys. Rev. Lett.* **74**, 1827 (1995).

[45] C. Fridh, L. Åsbrink, and E. Lindholm. *Chem. Phys. Lett.* **15**, 282 (1972).

[46] P. D. Burrow, J. A. Michejda, and K. D. Jordan. *J. Chem. Phys.* **86**, 9 (1987).

[47] National Institute of Standard, NIST Chemistry webbook, http://webbook.nist.gov/chemistry/.

[48] X. Blase, C. Attaccalite, and V. Olevano. *Phys. Rev. B* **83**, 115103 (2011).

[49] L. A. Curtiss, K. Raghavachari, P. C. Redfern, and J. A. Pople. *J. Chem. Phys.* **106**, 1063 (1997).

[50] X. Gonze, B. Amadon, P.-M. Anglade, J.-M. Beuken, F. Bottin, P. Boulanger, F. Bruneval, D. Caliste, R. Caracas, M. Ct, T. Deutsch, L. Genovese, Ph. Ghosez, M. Giantomassi, S. Goedecker, D.R. Hamann, P. Hermet, F. Jollet, G. Jomard, S. Leroux, M. Mancini, S. Mazevet, M.J.T. Oliveira, G. Onida, Y. Pouillon, T. Rangel, G.-M. Rignanese, D. Sangalli, R. Shaltaf, M. Torrent, M.J. Verstraete, G. Zerah, and J.W. Zwanziger. *Comput. Phys. Commun.* **180**, 2582 (2009).

[51] J. P. Perdew, M. Ernzerhof, and K. Burke. *J. Chem. Phys.* **105**, 9982 (1996).

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

[53] J. Heyd, G. E. Scuseria, and M. Ernzerhof. *J. Chem. Phys.* **118**, 8207 (2003).

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

[55] C. Rostgaard, K. W. Jacobsen, and K. S. Thygesen. *Phys. Rev. B* **81**, 085103 (2010).

[56] F. Bruneval and M. A. L. Marques. *J. Chem. Theory Comput.* **9**, 324 (2013).

[57] M. K. Y. Chan and G. Ceder. *Phys. Rev. Lett.* **105**, 196403 (2010).

[58] S. Baroni and R. Resta. *Phys. Rev. B* **33**, 7017 (1986).

[59] A. Baldereschi and E. Tosatti. *Phys. Rev. B* **17**, 4710 (1978).

[60] M. R. Martin, L. Steinbeck, I. D. White, H. N. Rojas, and R. W. Godby. *Comput. Phys. Commun.* **117**, 211 (1999).

[61] A. Fleszar and W. Hanke. *Phys. Rev. B* **56**, 10228 (1997).

[62] W. G. Aulbur, M. Städele, and A. Görling. *Phys. Rev. B* **62**, 7121 (2000).

[63] M. Rohlfing, P. Krüger, and J. Pollmann. *Phys. Rev. B* **48**, 17791 (1993).

[64] Y. Li and G. Galli. *Phys. Rev. B* **82**, 045321 (2010).

[65] S. Lebègue, B. Arnaud, M. Alouani, and P. E. Bloechl. *Phys. Rev. B* **67**, 155208 (2003).

[66] T. A. Pham, T. Li, S. Shankar, F. Gygi, and G. Galli. *Phys. Rev. B* **84**, 045308 (2011).

[67] T. A. Pham, T. Li, S. Shankar, F. Gygi, and G. Galli. *Appl. Phys. Lett.* **96**, 062902 (2010).

[68] J. Bauer. *Phys. Status Solidi A* **39**, 411 (1977).

[69] K. Lenzke, L. Landt, M. Hoener, H. Thomas, J. E. Dahl, S. G. Liu, R. M. K. Carlson, T. Möller, and C. Bostedt. *J. Chem. Phys.* **127**, 084320 (2007).

[70] N. D. Drummond, A. J. Williamson, R. J. Needs, and G. Galli. *Phys. Rev. Lett.* **95**, 096801 (2005).

[71] T. M. Willey, C. Bostedt, T. van Buuren, J. E. Dahl, S. G. Liu, R. M. K. Carlson, R. W. Meulenberg, E. J. Nelson, and L. J. Terminello. *Phys. Rev. B* **74**, 205432 (2006).

[72] H.-V. Nguyen, T. A. Pham, D. Rocca, and G. Galli. *Phys. Rev. B* **85**, 081101 (2012).

[73] M. J. Molina, T. Tso, L. T. Molina, and F. C.-Y. Wang. *Science* **238**, 1253 (1987).

[74] L. Pauling. *J. Am. Chem. Soc.* **57**, 2680 (1935).

[75] N. Materer, U. Starke, A. Barbieri, M.A. Van Hove, G.A. Somorjai, G.-J. Kroes, and C. Minot. *Sur. Sci.* **381**, 190 (1997).

[76] N. Materer, U. Starke, A. Barbieri, M van Hove, G. A. Somorjai, G J. Kroes, and C. Minot. *J. Phys. Chem.* **99**, 6267 (1995).

[77] J. Braun, A. Glebov, A. P. Graham, A. Menzel, and J. P. Toennies. *Phys. Rev. Lett.* **80**, 2638 (1998).

[78] A. Glebov, A. P. Graham, A. Menzel, J. P. Toennies, and P. Senet. *J. Chem. Phys.* **112**, 11011 (2000).

[79] A. P. Graham, A. Menzel, and J. P. Toennies. *J. Chem. Phys.* **111**, 1169 (1999).

[80] A. P. Graham and J. P. Toennies. *J. Chem. Phys.* **118**, 2879 (2003).

[81] M. T. Suter, P. U. Andersson, and J. B. C. Pettersson. *J. Chem. Phys.* **125**, 174704 (2006).

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

[83] N. H. Fletcher. *Philosophical Magazine Part B* **66**, 109 (1992).

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

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

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

[87] C. L. Bishop, D. Pan, L. M. Liu, G. A. Tribello, A. Michaelides, E. G. Wang, and B. Slater. *Faraday Discuss.* **141**, 277 (2009).

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

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

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

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

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

[93] R. Ramirez, T. Lopez-Ciudad, P. Kumar, and D. Marx. *J. Chem. Phys.* **121**, 3973 (2004).

[94] R. Iftimie and M. E. Tuckerman. *J. Chem. Phys.* **122**, 214508 (2005).

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

[96] W. Eisenberg, D. S.; Kauzmann. Clarendon, Oxford, 1969.

[97] D. Donadio, G. Cicero, E. Schwegler, M. Sharma, and G. Galli. *J. Phys. Chem. B* **113**, 4170 (2009).

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

[99] M. Watkins, D. Pan, E. Wang, A. Michaelides, J. VandeVondele, and B. Slater. *Nat. Mat.* **10**, 794 (2011).

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

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

[102] V. Buch, T. Tarbuck, G. L. Richmond, H. Groenzin, I. Li, and M. J. Shultz. *J. Chem. Phys.* **127**, 204710 (2007).

[103] B. Rowland, N. S. Kadagathur, J. P. Devlin, V. Buch, T. Feldman, and M. J. Wojcik. *J. Chem. Phys.* **102**, 8328 (1995).

[104] T. Hamashima, K. Mizuse, and A. Fujii. *J. Phys. Chem. A* **115**, 620 (2011).

[105] V. Buch, B. Sigurd, J. Paul Devlin, U. Buck, and J. K. Kazimirski. *Int. Rev. Phys. Chem.* **23**, 375 (2004).

[106] C. Steinbach, P. Andersson, J. K. Kazimirski, U. Buck, V. Buch, and T. A. Beu. *J. Phys. Chem. A* **108**, 6165 (2004).

[107] A. Lenz and L. Ojamae. *Phys. Chem. Chem. Phys.* **7**, 1905 (2005).

[108] J. E. Schaff and J. T. Roberts. *Langmuir* **14**, 1478 (1998).

[109] J. E. Schaff and J. T. Roberts. *J. Phys. Chem.* **100**, 14151 (1996).

[110] J. E. Schaff and J. T. Roberts. *J. Phys. Chem.* **98**, 6900 (1994).

[111] J. T. Roberts. *Acc. Chem. Res.* **31**, 415 (1998).

[112] J. P. Devlin, C. Joyce, and V. Buch. *J. Phys. Chem. A* **104**, 1974 (2000).

[113] J. P. Devlin, J. Sadlej, and V. Buch. *J. Phys. Chem. A* **105**, 974 (2001).

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

[115] G.-J. Kroes. *Sur. Sci.* **275**, 365 (1992).

[116] G. E. Brown, V. E. Henrich, W. H. Casey, D. L. Clark, C. Eggleston, A. Felmy, D. W. Goodman, M. Grätzel, G. Maciel, M. I. McCarthy, et al. *Chem. Rev.* **99**, 77 (1999).

[117] A. Verdaguer, G. M. Sacha, H. Bluhm, and M. Salmeron. *Chem. Rev.* **106**, 1478 (2006).

[118] G. Rubasinghege and V. H. Grassian. *Chem. Commun.* **49**, 3071 (2013).

[119] F. Zaera. *Chem. Rev.* **112**, 2920 (2012).

[120] P. Fenter and N. C. Sturchio. *Prog. Surf. Sci.* **77**, 171 (2004).

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

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

[123] J. VandeVondele, P. Troster, P. Tavan, and G. Mathias. *J. Phys. Chem. A* **116**, 2466 (2012).

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

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

[126] J. C. Grossman, E. Schwegler, E. W. Draeger, F. Gygi, and G. Galli. *J. Chem. Phys.* **120**, 300 (2004).

[127] E. Schwegler, J. C. Grossman, F. Gygi, and G. Galli. *J. Chem. Phys.* **121**, 5400 (2004).

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

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

[130] P. J. Eng, T. P. Trainor, G. E. Brown, Jr., G. A. Waychunas, M. Newville, S. R. Sutton, and M. L. Rivers. *Science* **288**, 1029 (2000).

[131] J. G. Catalano. *Geochim. Cosmochim. Acta* **75**, 2062 (2011).

[132] J. G. Catalano, C. Park, Z. Zhang, and P. Fenter. *Langmuir* **22**, 4668 (2006).

[133] J. G. Catalano. *J. Phys. Chem. C* **114**, 6624 (2010).

[134] B. Braunschweig, S. Eissner, and W. Daum. *J. Phys. Chem. Lett.* **112**, 1751 (2008).

[135] L. Zhang, C. Tian, G. A. Waychunas, and Y. R. Shen. *J. Am. Chem. Soc.* **130**, 7686 (2008).

[136] M. Flörsheimer, K. Kruse, R. Polly, A. Abdelmonem, B. Schimmelpfennig, R. Klenze, and T. Fanghänel. *Langmuir* **24**, 13434 (2008).

[137] J. Sung, L. Zhang, C. Tian, Y. R. Shen, and G. A. Waychunas. *J. Phys. Chem. C* **115**, 13887 (2011).

[138] J. Sung, Y. R. Shen, and G. A. Waychunas. *J. Phys.: Condens. Matt.* **24**, 124101 (2012).

[139] J. A. Kelber. *Surf. Sci. Rep.* **62**, 271 (2007).

[140] K. C. Hass, W. F. Schneider, A. Curioni, and W. Andreoni. *Science* **282**, 265 (1998).

[141] K. C. Hass, W. F Schneider, A. Curioni, and W. Andreoni. *J. Phys. Chem. B* **104**, 5527 (2000).

[142] M.-P. Gaigeot, M. Sprik, and M. Sulpizi. *J. Phys.: Condens. Matter* **24**, 124106 (2012).

[143] M. A. Nygren, D. H. Gay, C. Richard, and A. Catlow. *Surf. Sci.* **380**, 113 (1997).

[144] D. Argyris, T. Ho, D. R. Cole, and A. Striolo. *J. Phys. Chem. C* **115**, 2038 (2011).

[145] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. C., 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. Matt.* **39**, 395502 (2009).

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

[147] R. T. Cygan, J.-J. Liang, and A. G. Kalinichev. *J. Phys. Chem. B* **108**, 1255 (2004).

[148] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak. *J. Chem. Phys.* **81**, 3684 (1984).

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

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

[151] R. W. G. Wyckoff. *Crystal Structures*. Wiley-Interscience, New York, 2nd ed. edition, 1964.

[152] G. Cicero, J. C. Grossman, E. Schwegler, F. Gygi, and G. Galli. *J. Am. Chem. Soc.* **130**, 1871 (2008).

[153] S. H. Lee and P. J. Rossky. *J. Chem. Phys.* **100**, 3334 (1994).

[154] G. Cicero, J. C. Grossman, A. Catellani, and G. Galli. *J. Am. Chem. Soc.* **127**, 6830 (2005).

[155] J.-J. Max and C. Chapados. *J. Chem. Phys.* **131**, 184505 (2009).

[156] T. A. Pham, P. Huang, E. Schwegler, and G. Galli. *J. Phys. Chem. A* **116**, 9255 (2012).

[157] M. Sharma, D. Donadio, E. Schwegler, and G. Galli. *Nano Lett.* **8**, 2959 (2008).

[158] M. Sulpizi, M.-P. Gaigeot, and M. Sprik. *J. Chem. Theory Comput.* **8**, 1037 (2012).

[159] F. Paesani and G. A. Voth. *J. Chem. Phys.* **132**, 014105 (2010).

[160] E. Vogel. *Nat. Nanotechnol.* **2**, 25 (2007).

[161] J. Robertson. *Rep. Prog. Phys.* **69**, 327 (2006).

[162] G. D. Wilk, R. M. Wallace, and J. M. Anthony. *J. Appl. Phys.* **89**, 5243 (2001).

[163] D. Fischer, A. Curioni, S. Billeter, and W. Andreoni. *Phys. Rev. Lett.* **92**, 236405 (2004).

[164] T. P. Ma. *IEEE Trans. Electron Devices* **45**, 680 (1998).

[165] R. W. G. Wyckoff. *Crystal structures*, volume 2. Interscience publishers New York, 1963.

[166] N. Wada, S.A. Solin, J. Wong, and S. Prochazka. *J. Non-Cryst. Solids* **43**, 7 (1981).

[167] Y. Cai, L. Zhang, Q. Zeng, L. Cheng, and Y. Xu. *Phys. Rev. B* **74**, 174301 (2006).

[168] R. D. Carson and S. E. Schnatterly. *Phys. Rev. B* **33**, 2432 (1986).

[169] R. Resta and K. Kunc. *Phys. Rev. B* **34**, 7146 (1986).

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

[171] S. Hamel, A. J. Williamson, H. F. Wilson, F. Gygi, G. Galli, E. Ratner, and D. Wack. *Appl. Phys. Lett.* **92**, 3115 (2008).

[172] F. Giustino and A. Pasquarello. *Phys. Rev. B* **71**, 144104 (2005).

[173] V. M. Bermudez. *Sur. Sci.* **579**, 11 (2005).

[174] S. Ogata, N. Hirosaki, C. Kocer, and H. Kitagawa. *Phys. Rev. B* **64**, 172102 (2001).

[175] F. de Brito Mota, J. F. Justo, and A. Fazzio. *Phys. Rev. B* **58**, 8323 (1998).

[176] R. P. Ortiz, A. Facchetti, and T. J. Marks. *Chem. Rev.* **110**, 205 (2010).

[177] Y. C. Yeo, T. J. King, and C. Hu. *Appl. Phys. Lett.* **81**, 2091 (2002).

[178] A. Biberman and K. Bergman. *Rep. Prog. Phys.* **75**, 046402 (2012).

[179] J. W. Keister, J. E. Rowe, J. J. Kolodziej, H. Niimi, T. E. Madey, and G. Lucovsky. *J. Vac. Sci. Technol. B* **17**, 1831 (1999).

[180] V. A. Gritsenko, A. V. Shaposhnikov, W. M. Kwok, H. Wong, and G. M. Jidomirov. *Thin Solid Films* **437**, 135 (2003).

[181] M. Higuchi, S. Sugawa, E. Ikenaga, J. Ushio, H. Nohira, T. Maruizumi, A. Teramoto, T. Ohmi, and T. Hattori. *Appl. Phys. Lett.* **90**, 123114 (2007).

[182] A. Pasquarello, M. S. Hybertsen, and R. Car. *Nature* **396**, 58 (1998).

[183] A. Bongiorno and A. Pasquarello. *Appl. Phys. Lett.* **83**, 1417 (2003).

[184] A. Bongiorno, A. Pasquarello, M. S. Hybertsen, and L. C. Feldman. *Phys. Rev. Lett.* **90**, 186101 (2003).

[185] F. Giustino, A. Bongiorno, and A. Pasquarello. *Appl. Phys. Lett.* **86**, 192901 (2005).

[186] D. A. Muller, T. Sorsch, S. Moccio, F. H. Baumann, K. Evans-Lutterodt, and G. Timp. *Nature* **399**, 758 (1999).

[187] S. Tang, R. M. Wallace, A. Seabaugh, and D. King-Smith. *Appl. Surf. Sci.* **135**, 137 (1998).

[188] F. Giustino, A. Bongiorno, and A. Pasquarello. *J. Phys.: Condens. Matt.* **17**, S2065 (2005).

[189] G. L. Zhao and M. E. Bachlechner. *Phys. Rev. B* **58**, 1887 (1998).

[190] M. E. Bachlechner, J. Zhang, Y. Wang, J. Schiffbauer, S. R. Knudsen, and D. Korakakis. *Phys. Rev. B* **72**, 094115 (2005).

[191] M. E. Bachlechner, D. Srivastava, E. T. Owens, J. Schiffbauer, J. T. Anderson, M. R. Burky, S. C. Ducatman, A. M. Gripper, E. J. Guffey, and F. S. Ramos. *Phys. Rev. B* **74**, 075327 (2006).

[192] M. Yang, R. Q. Wu, W. S. Deng, L. Shen, Z. D. Sha, Y. Q. Cai, Y. P. Feng, and S. J. Wang. *J. Appl. Phys.* **105**, 024108 (2009).

[193] A. Omeltchenko, M. E. Bachlechner, A. Nakano, R. K. Kalia, P. Vashishta, I. Ebbsjö, A. Madhukar, and P. Messina. *Phys. Rev. Lett.* **84**, 318 (2000).

[194] T. A. Pham, H.-V. Nguyen, D. Rocca, and G. Galli. *Phys. Rev. B* **87**, 155148 (2013).

[195] X. Zhu and S. G. Louie. *Phys. Rev. B* **43**, 14142 (1991).

[196] F. Devynck, F. Giustino, P. Broqvist, and A. Pasquarello. *Phys. Rev. B* **76**, 075351 (2007).

[197] P. Broqvist, J. F. Binder, and A. Pasquarello. *Appl. Phys. Lett.* **94**, 141911 (2009).

[198] S. M. Sze. *Physics of Semiconductor Devices*. John Wiley & Sons, Inc., New York, 1981.

[199] T. Yamasaki, C. Kaneta, T. Uchiyama, T. Uda, and K. Terakura. *Phys. Rev. B* **63**, 115314 (2001).

[200] A. M. Goodman. *Appl. Phys. Lett.* **13**, 275 (1968).

[201] J. Bauer. *Phys. Status Solidi A* **39**, 411 (1977).

[202] S. V. Deshpande, E. Gulari, S. W. Brown, and S. C. Rand. *J. Appl. Phys.* **77**, 6534 (1995).

[203] C. Kittel. *Introduction to Solid State Physics*. John Wiley & Sons, Inc., New York, 1996.

[204] R. Shaltaf, G.-M. Rignanese, X. Gonze, F. Giustino, and A. Pasquarello. *Phys. Rev. Lett.* **100**, 186401 (2008).

[205] J. D. Walton. *Amer. Ceram. Soc. Bull.* **53**, 255 (1974).

[206] J. S. Thorp and R. I. Sharif. *J. Mater. Sci.* **12**, 2274 (1977).

[207] M. G. Walter, E. L. Warren, J. R. McKone, S. W. Boettcher, Q. Mi, E. A. Santori, and N. S. Lewis. *Chem. Rev.* **110**, 6446 (2010).

[208] L. Turi and P. J. Rossky. *Chem. Rev.* **112**, 5641 (2012).

[209] C. E. Krohn, P. R. Antoniewicz, and J. C. Thompson. *Surf. Sci.* **101**, 241 (1980).

[210] A. Bernas, C. Ferradini, and J.-P. Jay-Gerin. *Chem. Phys.* **222**, 151 (1997).

[211] P. Delahay and K. Von Burg. *Chem. Phys. Lett.* **83**, 250 (1981).

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

[213] A. M. Brodsky and A. V. Tsarevsky. *J. Chem. Soc., Faraday Trans. 2* **72**, 1781 (1976).

[214] B. Winter and M. Faubel. *Chem. Rev.* **106**, 1176 (2006).

[215] B. Winter, R. Weber, W. Widdra, M. Dittmar, M. Faubel, and I. V. Hertel. *J. Phys. Chem. A* **108**, 2625 (2004).

[216] D. Grand, A. Bernas, and E. Amouyal. *Chem. Phys.* **44**, 73 (1979).

[217] A. Bernas, D. Grand, and E. Amouyal. *J. Phys. Chem.* **84**, 1259 (1980).

[218] J. Schnitker, P. J. Rossky, and G. A. Kenney-Wallace. *J. Chem. Phys.* **85**, 2986 (1986).

[219] M. S. Hybertsen and S. G. Louie. *Phys. Rev. Lett.* **55**, 1418 (1985).

[220] V. Garbuio, M. Cascella, L. Reining, R. Del Sole, and O. Pulci. *Phys. Rev. Lett.* **97**, 137402 (2006).

[221] D. Lu, F. Gygi, and G. Galli. *Phys. Rev. Lett.* **100**, 147601 (2008).

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

[223] J. C. Grossman, E. Schwegler, E. W. Draeger, F. Gygi, and G. Galli. *J. Chem. Phys.* **120**, 300 (2004).

[224] E. Schwegler, J. C. Grossman, F. Gygi, and G. Galli. *J. Chem. Phys.* **121**, 5400 (2004).

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

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

[227] M. A. L. Marques, J. Vidal, M. J. T. Oliveira, L. Reining, and S. Botti. *Phys. Rev. B* **83**, 035119 (2011).

[228] K. Leung. *J. Phys. Chem. Lett.* **1**, 496 (2010).

[229] S. Trasatti. *Pure Appl. Chem.* **58**, 955 (1986).

[230] J. Cheng and M. Sprik. *Phys. Chem. Chem. Phys.* **14**, 11245 (2012).

[231] A. M. Duffin and R. J. Saykally. *J. Phys. Chem. C* **111**, 12031 (2007).

[232] N. Preissler, F. Buchner, T. Schultz, and A. Lbcke. *J. Phys. Chem. B* **117**, 2422 (2013).

[233] C. Rostgaard, K. W. Jacobsen, and K. S. Thygesen. *Phys. Rev. B* **81**, 085103 (2010).

[234] Q. Mi, A. Zhanaidarova, B. S. Brunschwig, H. B. Gray, and N. S. Lewis. *Energ. Environ.* **5**, 5694 (2012).

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

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

[237] A. W. Omta, M. F. Kropman, S. Woutersen, and H. J. Bakker. *Science* **301**, 347 (2003).

[238] L. X. Dang, G. K. Schenter, V.-A. Glezakou, and J. L. Fulton. *J. Phys. Chem. B* **110**, 23644 (2006).

[239] R. Mancinelli, A. Botti, F. Bruni, M. A. Ricci, and A. K. Soper. *J. Phys. Chem. B* **111**, 13570 (2007).

[240] H. J. Kulik, N. Marzari, A. A. Correa, D. Prendergast, E. Schwegler, and G. Galli. *J. Phys. Chem. B* **114**, 9594 (2010).

[241] A. Bankura, V. Carnevale, and M. L. Klein. *J. Chem. Phys.* **138**, 014501 (2013).

[242] P. B. Petersen and R. J. Saykally. *J. Phys. Chem. B* **110**, 14060 (2006).

[243] S. Park and M. D. Fayer. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 16731 (2007).

[244] J. D. Smith, R. J. Saykally, and P. L. Geissler. *J. Am. Chem. Soc.* **129**, 13847 (2007).

[245] B. S. Mallik, A. Semparithi, and A. Chandra. *J. Chem. Phys.* **129**, 194512 (2008).

[246] A. M. Jubb and H. C. Allen. *J. Phys. Chem. C* **116**, 13161 (2012).

[247] J. L. Skinner, P. A. Pieniazek, and S. M. Gruenbaum. *Acc. Chem. Res.* **45**, 93 (2011).

[248] 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).

[249] R. Seidel, S. Thurmer, and B. Winter. *J. Phys. Chem. Lett.* **2**, 633 (2011).

[250] C. Adriaanse, M. Sulpizi, J. VandeVondele, and M. Sprik. *J. Am. Chem. Soc.* **131**, 6046 (2009).

[251] C. Adriaanse, J. Cheng, V. Chau, M. Sulpizi, J. VandeVondele, and M. Sprik. *J. Phys. Chem. Lett.* **3**, 3411 (2012).

[252] D. Ghosh, A. Roy, R. Seidel, B. Winter, S. Bradforth, and A. I. Krylov. *J. Phys. Chem. B* **116**, 7269 (2012).

[253] E. Pluharova, M. Oncak, R. Seidel, C. Schroeder, W. Schroeder, B. Winter, S. E. Bradforth, P. Jungwirth, and P. Slavicek. *J. Phys. Chem. B* **116**, 13254 (2012).

[254] A. J. Cohen, P. Mori-Sanchez, and W. T. Yang. *Science* **321**, 792 (2009).

[255] A. J. Cohen, P. Mori-Sánchez, and W. Yang. *Chem. Rev.* **112**, 289 (2011).

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

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

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

[259] D. R. Hamann. *Phys. Rev. B* **40**, 2980 (1989).

[260] L. Kleinman and D. M. Bylander. *Phys. Rev. Lett.* **48**, 1425 (1982).

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

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

[263] A. Bernas, C. Ferradini, and J.-P. Jay-Gerin. *Chem. Phys.* **222**, 151 (1997).

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

[265] S. Ghosal, J. C. Hemminger, H. Bluhm, B. S. Mun, E. L. D. Hebenstreit, G. Ketteler, D. F. Ogletree, F. G. Requejo, and M. Salmeron. *Science* **307**, 563 (2005).

[266] D. J. Tobias and J. C. Hemminger. *Science* **319**, 1197 (2008).

[267] C. Caleman, J. S. Hub, P. J. van Maaren, and D. van der Spoel. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 6838 (2011).

[268] M. D. Baer and C. J. Mundy. *J. Phys. Chem. Lett.* **2**, 1088 (2011).

[269] I. Watanabe, J. B. Flanagan, and P. Delahay. *J. Chem. Phys.* **73**, 2057 (1980).

[270] M. Grätzel. *Nature* **414**, 338 (2001).

[271] M. Marsman, J. Paier, A. Stroppa, and G. Kresse. *J. Phys.: Condens. Matt.* **20**, 064201 (2008).

[272] M. C. Toroker, D. K. Kanan, N. Alidoust, L. Y. Isseroff, P. Liao, and E. A. Carter. *Phys. Chem. Chem. Phys.* **13**, 16644 (2011).

[273] H. Jiang. *J. Phys. Chem. C* **116**, 7664 (2012).

[274] H. L. Zhuang and R. G. Hennig. *J. Phys. Chem. C* **117**, 20440 (2013).

[275] V. Stevanović, S. Lany, D. S. Ginley, W. Tumas, and A. Zunger. *Phys. Chem. Chem. Phys.* **16**, 3706 (2014).

[276] S. Trasatti. *Pure Appl. Chem.* **58**, 955 (1986).

[277] J. Cheng and M. Sprik. *Phys. Rev. B* **82**, 081406 (2010).

[278] Y. Wu, M. K. Y. Chan, and G. Ceder. *Phys. Rev. B* **83**, 235301 (2011).

[279] Y. Li, L. E. O'Leary, N. S. Lewis, and G. Galli. *J. Phys. Chem. C* **117**, 5188 (2013).

[280] V. Tripkovic, M. E. Björketun, E. Skúlason, and J. Rossmeisl. *Phys. Rev. B* **84**, 115452 (2011).

[281] C. G. Van de Walle and R. M. Martin. *Phys. Rev. B* **35**, 8154 (1987).

[282] S. B. Zhang, D. Tomanek, S. G. Louie, M. L. Cohen, and M. S. Hybertsen. *Solid State Commun.* **66**, 585 (1988).

[283] M. J. Madou, B. H. Loo, K. W. Frese, and S. R. Morrison. *Sur. Sci.* **108**, 135 (1981).

[284] G. S. Higashi, Y. J. Chabal, G. W. Trucks, and K. Raghavachari. *Appl. Phys. Lett.* **56**, 656 (1990).

[285] R. Hunger, R. Fritsche, B. Jaeckel, W. Jaegermann, L. J. Webb, and N. S. Lewis. *Phys. Rev. B* **72**, 045317 (2005).

[286] W. J. Royea, A. Juang, and N. S. Lewis. *Appl. Phys. Lett.* **77**, 1988 (2000).

[287] K. E. Plass, X. Liu, B. S. Brunschwig, and N. S. Lewis. *Chem. Mat.* **20**, 2228 (2008).

[288] L. E. O'Leary, E. Johansson, B. S. Brunschwig, and N. S. Lewis. *J. Phys. Chem. B* **114**, 14298 (2010).

[289] A. Faucheux, A. C. Gouget-Laemmel, C. Henry de Villeneuve, R. Boukherroub, F. Ozanam, P. Allongue, and J. Chazalviel. *Langmuir* **22**, 153 (2006).

[290] D. Lee, E. Schwegler, and Y. Kanai. *J. Phys. Chem. C* **118**, 8508 (2014).

[291] T. A. Pham, T. Li, H.-V. Nguyen, S. Shankar, F. Gygi, and G. Galli. *Appl. Phys. Lett.* **102**, 241603 (2013).

[292] A. Aliano, Y. Li, G. Cicero, and G. Galli. *J. Phys. Chem. C* **114**, 11898 (2010).

[293] Y. Li and G. Galli. *Phys. Rev. B* **82**, 045321 (2010).

[294] J. M. Bolts and M. S. Wrighton. *J. Phys. Chem.* **80**, 2641 (1976).

[295] S. Schnur and A. Groß. *New J. Phys.* **11**, 125003 (2009).