Insights into reactivity and defect properties of semiconductor surfaces from first-principles computations

INSIGHTS INTO REACTIVITY AND DEFECT PROPERTIES OF SEMICONDUCTOR SURFACES FROM FIRST-PRINCIPLES COMPUTATIONS

A dissertation presented by

DMITRY VINICHENKO

to

The Department of Chemistry and Chemical Biology

in partial fulfillment of the requirements

for the degree of

Doctor of Philoshophy

in the subject of

Chemical Physics

HARVARD UNIVERSITY

Cambridge, Massachusetts

June 2016

Dissertation advisors:
Professor Cynthia M. Friend
Professor Efthimios Kaxiras

Dmitry Vinichenko

INSIGHTS INTO REACTIVITY AND DEFECT PROPERTIES OF SEMICONDUCTOR SURFACES FROM FIRST-PRINCIPLES COMPUTATIONS

ABSTRACT

We use first-principles computations based on Ehrenfest dynamics and density functional theory to study water and methanol photo-oxidation on a model photo-catalytic material – the (110) surface of rutile titanium dioxide. We simulate photo-excitation in titania and the subsequent excited-state reaction trajectories. Analysis of the coupled dynamics of the electronic and ionic subsystems allows us to establish a novel reaction mechanism, for which we propose the name "photo-induced C-H acidity." We provide a detailed and intuitive interpretation of the mechanism in terms of Lewis structures, identify the driving forces of the process, and propose general design principles for efficient photo-catalytic systems.

Another important factor in the reactivity of semiconductor catalysts is the presence of defects in surface and subsurface regions. Knowledge of the formation energies of defects and impurities in different charge states is required in order to obtain insight into their concentration and stability. We develop an internally consistent method for calculating formation energies of charged defects based on density functional theory, which is applicable to both surface regions of semiconducting materials and two-dimensional materials. We discuss the implementation and usage details of the method and provide an example of its usage for studying sulfur vacancy formation in MoS_2 monolayer.

CONTENTS

ACKNOWLEDGEMENTS

The rate of my professional growth has undergone an unexpectedly rapid increase when I started my graduate studies in the Chemical Physics program at Harvard; and while I'll have plenty of time to reflect on the beneficial effects of such non-adiabatic transition, it is pretty obvious even now that it would be impossible without the mentorship and guidance of my advisors, Professor Cynthia M. Friend and Professor Efthimios Kaxiras. I am very grateful to them for creating an intellectually rich and stimulating environment in their research groups, which I had an honor to be a part of, for providing the numerous opportunities to hone my critical thinking skills while working on exciting and challenging research projects, as well as for all the advice throughout my time in graduate school. I would also like to thank Professor Alán Aspuru-Guzik for the insightful conversations, which are a very memorable part of my graduate school experience.

This work was made possible with the financial support from NSF CHE program and Harvard Consortium for Energy and Environment (HUCE). The actual calculations were carried out on the machines of the National Energy Research Supercomputing Center (NERSC), Extreme Science and Engineering Discovery Environment (XSEDE), and Odyssey cluster maintained by Harvard FAS Research Computing division.

I have been very lucky to work together with brilliant researchers in Kaxiras and Friend groups, in particular: Robert Hoyt, Fanny Hiebel, Branko Zugic, Georgios Tritsaris, and Brad Malone. Oscar Grånäs has proven to be a great colleague and a very knowledgeable resource for all the deep and rarely-discussed questions about the various aspects of first-principle computations. It was my great pleasure to work with Mehmet Gokhan Sensoy on the computation of defect formation energies and to have the insightful discussions with him on that

topic and about research in general. I would like to especially acknowledge Grigory Kolesov, his indomitable and contagious enthusiasm for science, unmatched expertise in scientific computing, and all the help and advice I've received from him throughout these years.

I would also like to acknowledge my mentors at Moscow State University: Professor Vladimir Zlomanov for introducing me to the world of academic research, and Dr. Andrey Scherbinin for his support and encouragement during my transition from experimental chemistry to computational material science, and all the great advice regarding my diploma thesis.

Finally, I'd like to mention that completion of this work would be not possible without the support of my family and close friends – and to express my sincere gratitude to them.

CHAPTER 1. INTRODUCTION

The unprecedented increase of available computing power has led to equally rapid progress in computational science and has enabled researchers to explore new exciting topics and tackle associated scientific problems. In this work, we focus on methods for the computational simulation of the properties of semiconductor materials, which have great prospects for various practical applications, particularly for photo-catalysis. In order to obtain a fundamental understanding of materials performance for this application, microscopic insight into dynamics of the electronic and ionic components of the system in question is necessary. Part I of this thesis describes the modern computational techniques used for simulating the real-time dynamics of chemical systems on a microscopic level. In particular, we discuss the localized basis set implementation of Ehrenfest dynamics and density functional theory – TDAP-2.0, developed by the Kaxiras research group.

We then discuss the results of applying that method to the simulation of model systems: water and methanol adsorbed on a rutile titania (110) surface. Titania has been extensively studied as a model photo-catalytic system and has numerous proposed applications, from solar cells to photo-catalysis for solar hydrogen production and environmental pollutant cleanup. For water photo-oxidation we establish a dependence of the reaction path on the external conditions, like temperature, and the presence of defects in the subsurface layer of the photo-catalyst. This partly motivates the need for developing advanced methods for computation of defect formation energies, which we discuss in Part II. Another example we have considered is the oxidation of methanol to formaldehyde – it is both an important industrial process, and an interesting model reaction, which has been a subject of many experimental studies. We simulate both ground- and excited-state reaction trajectories for this process and identify the key driving forces and

intermediates. Establishing a connection between the results of first-principle simulation based on advanced electronic structure theory and traditional chemical notation, Lewis structures and electron pushing schemes, allows us to put forward a novel interpretation of the reaction mechanism, "photo-induced C−H acidity". We then generalize the obtained insights into guiding principles for the design of efficient photo-catalytic systems.

In Part II we discuss problems associated with bridging the gap between the theoretical models and the properties and performance of real materials, the most important of them being the presence of defects and impurities in seminconductors. Impurities can very strongly affect the host material – the understanding of this fact also happened to give a giant boost to humanity's technological progress through the invention of transistors. In the case of catalysts, however, the results of the presence of impurities can be more mixed. First-principle simulations allow to obtain unique insight into the microscopic structure of the defect local environment, as well as provide information about defect formation energy. A number of issues, however, exist with simulation of charged defects, especially in the surface regions of semiconducting materials, which are necessary to obtain the full picture of material behavior. The available methods based on the density functional theory with periodic boundary conditions and the supercell approach tend to use different names for the same quantities, have variability in signs of the terms involved in the expression for defect formation energy, use unnecessary and sometimes even unjustified approximations, and tend overemphasize some parts of the problem without paying attention to the rest of it. We propose a consistent method for the computation of defect formation energies which can be used for any charge states of defects on surfaces of bulk semiconductors and in monolayers of two-dimensional materials, such as graphene or molybdenum disulfide. We describe a computationally efficient parallel implementation of the

method, highlight several non-trivial aspects of its practical usage, and conclude with discussing the potential limitations and the scope of applicability of such approach. Then we demonstrate an example of application of this technique by studying the behavior of sulfur vacancy in monolayer MoS_2 – the most abundant intrinsic defect in this material. Our method allows us to address the relative stability of different charge states of this defect, for both isolated vacancies and pairs of vacancies. We discuss the defect-induced changes in the geometric and electronic structure of the material in the context of the formation of recently proposed "ripplocation" structure in monolayer MoS_2 and associated strain. Finally, we demonstrate another real-world application of semiconductor materials by considering the recently developed solar cells based on hybrid organic-inorganic absorber with perovskite structure. We propose an integrated model for a solar cell and demonstrate how the chemical composition of the material translates into the key performance index – power conversion efficiency. Our model allows us to infer the most important properties of the absorber and carrier transporting materials from the viewpoint of cell efficiency, and to establish guidelines for the experimental research efforts.

PART I

MODELING HETEROGENEOUS PHOTO-CHEMICAL PROCESSES: METHODS AND APPLICATIONS

In this part of the thesis we discuss in detail the modern methods for modeling the dynamics of coupled electron-ion systems on atomic scale, and their applications to the systems of practical relevance. In Chapter 2 we describe the relevant approximations invoked for translating the quantum-mechanical equations into readily applicable computational schemes for studying both ground-state and excited-state properties of the system of interest. In particular, we discuss the Ehrenfest dynamics methodology and the aspects of its efficient numerical implementation.

In Chapter 3 we demonstrate an example application – water photo-oxidation on the rutile TiO_2 (110) surface, which is important in the context of utilization of solar energy for hydrogen production. The results in this chapter demonstrate an intricate interplay of external factors like temperature and the properties of material and adsorbate, in particular the presence of defects and the electronic structure of the catalyst, in influencing the reaction pathway. In Chapter 4 we use the Ehrenfest dynamics technique to simulate the photo-oxidation of methanol to formaldehyde on rutile titania, and to establish a novel mechanism of chemical reaction – "photo-induced C–H acidity". We then generalize our results and propose a set of design rules for the design of efficient photo-catalytic systems.

CHAPTER 2. FIRST-PRINCIPLES METHODS FOR MODELING ELECTRONIC STRUCTURE AND IONIC DYNAMICS

While the quantum-mechanical description of many-body systems is non-trivial by itself, another layer of complexity is introduced by the computational schemes which translate the theoretical equations into computer code. In this chapter we review the most commonly used approximations to the simulation of electronic structure and dynamics, and consider methods for molecular dynamics. We conclude with reviewing the localized basis set implementation of the Ehrenfest dynamics method – code named TDAP-2.0 described in the paper: G. Kolesov, O. Granas, R. Hoyt, D. Vinichenko, E. Kaxiras. Real-time TD-DFT with Classical Ion Dynamics: Methodology and Applications. Journal of Chemical Theory and Computation, 2016, V. 12, pp. 466-476.

2.1. Separation of electronic and nuclear sub-problems

In the non-relativistic approximation, all properties of a system of n electrons and N nuclei can be extracted from the many-body wavefunction $\Phi\left(\{\vec{r}_i\},\{\vec{R}_I\},t\right)$, where $\{\vec{r}_i\}$ are electron coordinates, $i = 1 \ldots n$, and $\{\vec{R}_I\}$ are ion positions, $I = 1 \ldots N$. The time evolution of this wavefunction is governed by the *time-dependent Schrodinger equation* (TDSE), which in the atomic system of units ($\hbar = 1, e = 1, m_e = 1, \epsilon_0 = 1/4\pi$) can be written as:[1-3]

$$i\frac{\partial}{\partial t}\Phi\left(\{\vec{r}_i\},\{\vec{R}_I\},t\right) = \mathcal{H}\Phi\left(\{\vec{r}_i\},\{\vec{R}_I\},t\right) \tag{2.1.1}$$

where $\mathcal{H}$ denotes the system Hamiltonian:

$$\begin{aligned}\mathcal{H} &= -\sum_I \frac{1}{2M_I}\nabla_I^2 - \sum_i \frac{1}{2}\nabla_i^2 + \sum_{i<j}\frac{1}{|\vec{r}_i - \vec{r}_j|} - \sum_{i,I}\frac{Z_I}{|\vec{r}_i - \vec{R}_I|} \\ &\quad + \sum_{I<J}\frac{Z_J Z_I}{|\vec{R}_J - \vec{R}_I|} \\ &= -\sum_I \frac{1}{2M_I}\nabla_I^2 - \sum_i \frac{1}{2}\nabla_i^2 + V_{n-e}\left(\{\vec{r}_i\},\{\vec{R}_I\}\right) \\ &= -\sum_I \frac{1}{2M_I}\nabla_I^2 + \mathcal{H}_e\left(\{\vec{r}_i\},\{\vec{R}_I\}\right)\end{aligned} \tag{2.1.2}$$

where M_I is the mass of nucleus I, Z_I is the corresponding charge, ∇_I^2 is the sum of second derivatives of the wavefunction wrt nucleus I positions, proportional to the operator of kinetic energy of nucleus I, $V_{n-e}\left(\{\vec{r}_i\},\{\vec{R}_I\}\right)$ is the operator of potential energy of electron-electron, electron-nuclei, and nuclei-nuclei Coulomb interaction, and $\mathcal{H}_e\left(\{\vec{r}_i\},\{\vec{R}_I\}\right)$ is the electronic Hamiltonian.

If the Hamiltonian $\mathcal{H}$ of the system does not depend on time explicitly, then separation of variables can be carried out in the TDSE (2.1.1), and the wavefunction can be expanded into a series of the eigenstates of the Hamiltonian $\widetilde{\Phi}_k$:

$$\Phi\left(\{\vec{r}_i\},\{\vec{R}_I\},t\right) = \sum_k c_k(t)\widetilde{\Phi}_k\left(\{\vec{r}_i\},\{\vec{R}_I\}\right) \tag{2.1.3}$$

where $\tilde{E}_k$ and $\widetilde{\Phi}_k$ are eigenvalues and eigenfunctions of $\mathcal{H}$, defined through the *time-independent Schrodinger equation* (TISE):

$$\mathcal{H}\widetilde{\Phi}_k\left(\{\vec{r}_i\},\{\vec{R}_I\}\right) = \tilde{E}_k\widetilde{\Phi}_k\left(\{\vec{r}_i\},\{\vec{R}_I\}\right) \tag{2.1.4}$$

An approximation made to simplify the problem is based on the fact that the mass of the electrons is at least 1800 times smaller than the mass of nuclei; therefore, the velocities of the nuclei are much smaller than the velocities of electrons, and the electrons are instantaneously

adapting to the motion of the nuclei. Mathematically, it can be expressed in the following way: first, for a given set of ionic positions $\{\vec{R}_I\}$ the eigenstates $\Psi_k(\{\vec{r}_i\};\{\vec{R}_I\})$ of the electronic Hamiltonian are found from the *electronic* Schrodinger equation:

$$\mathcal{H}_e(\{\vec{r}_i\},\{\vec{R}_I\})\Psi_k(\{\vec{r}_i\};\{\vec{R}_I\}) = E_k(\{\vec{R}_I\})\Psi_k(\{\vec{r}_i\};\{\vec{R}_I\}) \tag{2.1.5}$$

It is important to note that these electronic wavefunctions $\Psi_k(\{\vec{r}_i\};\{\vec{R}_I\})$ have only parametric dependence on the positions of the nuclei, and they comprise a complete and orthonormal set with respect to electronic coordinates. The corresponding eigenvalues $E_k(\{\vec{R}_I\})$ define the electronic *potential energy surfaces* (PES), which also are parametrically dependent on nuclei positions.

Then the total wavefunction can be expanded in the basis of electronic eigenstates:

$$\Phi(\{\vec{r}_i\},\{\vec{R}_I\},t) = \sum_{k=1}^{\infty} \chi_k(\{\vec{R}_I\},t)\Psi_k(\{\vec{r}_i\};\{\vec{R}_I\}) \tag{2.1.6}$$

where $\chi_k(\{\vec{R}_I\},t)$ are the nuclear wavefunctions and can be treated as time-dependent coefficients. The equations of motion for those coefficients can be derived by inserting this ansatz in the Schrodinger equation for the overall system (2.1.1), which results in the equations of motion for the nuclear wavefunctions

$$i\frac{\partial \chi_l}{\partial t} = -\sum_I \frac{1}{2M_I}\nabla_I^2\chi_l + E_l(\{\vec{R}_I\})\chi_l + \sum_k C_{lk}\chi_k \tag{2.1.7}$$

where C_{lk} are the non-adiabatic coupling coefficients, given by an operator

$$C_{lk} = -\sum_I \frac{1}{2M_I}\int \Psi_k^*\nabla_I^2\Psi_k d\vec{r} - \frac{1}{M_I}\sum_I \Psi_k^*\nabla_I\Psi_l d\vec{r}\,\nabla_I \tag{2.1.8}$$

In the "adiabatic approximation" to the problem only the diagonal coupling coefficients are retained, which leads to an energy correction to the decoupled adiabatic levels. The

implication of such approximation is that the motion of the nuclei proceeds along a given adiabatic potential energy surface, and the expansion of the wavefunction (2.1.6) can be reduced to only one term:

$$\Phi\left(\{\vec{r}_i\},\{\vec{R}_I\},t\right) = \chi_l\left(\{\vec{R}_I\},t\right)\Psi_l\left(\{\vec{r}_i\};\{\vec{R}_I\}\right) \tag{2.1.9}$$

Even further simplification can be achieved by neglecting diagonal terms, which leads to the "Born-Oppenheimer" approximation –the simplest form of the equation of motion for the nuclei

$$i\frac{\partial\chi_l}{\partial t} = \left(-\sum_I \frac{1}{2M_I}\nabla_I^2 + E_l\left(\{\vec{R}_I\}\right)\right)\chi_l \tag{2.1.10}$$

In such manner, the problem can be separated into two components: ionic dynamics and electronic structure; in the following discussion, we will consider those parts in turn.

2.2. Ionic dynamics

2.2.1. Born-Oppenheimer molecular dynamics

The most important approximation for ionic dynamics is the model of classical nuclei represented by point charges. This limit is equivalent to setting Planck's constant to 0 in the equation 2.1.10. Such transition can be done correctly by representing the nuclear wavefunction χ_l in terms of an amplitude factor A_l and phase S_l (the Planck's constant is written explicitly):

$$\chi_l\left(\{\vec{R}_I\};t\right) = A_l\left(\{\vec{R}_I\};t\right)\exp\left(\frac{iS_l\left(\{\vec{R}_I\};t\right)}{\hbar}\right) \tag{2.2.1}$$

which upon plugging into the Born-Oppenheimer dynamics equation lead to the following equations of motion for the amplitude and phase:

$$\frac{\partial S_l}{\partial t} + \sum_I \frac{1}{2M_I}(\nabla_I S_l)^2 + E_l = \hbar^2 \sum_I \frac{1}{2M_I}\frac{\nabla_I^2 A_l}{A_l} \tag{2.2.2}$$

$$\frac{\partial A_l}{\partial t} + \sum_I \frac{1}{M_I}(\nabla_I A_l)(\nabla_I S_l) + \sum_I \frac{1}{2M_I} A_l(\nabla_I^2 S_l) = 0 \tag{2.2.3}$$

In the classical limit, $\hbar \to 0$, the right-hand side of the equation 2.2.2 vanishes and it becomes isomorphic to the classical-mechanics Hamilton-Jacobi equations of motion. Therefore, defining the connecting transformation $\vec{P}_I \equiv \nabla_I S_l$, the Newtonian equations of motion can be inferred:

$$\frac{\partial \vec{P}_I}{\partial t} = -\nabla_I E_k, \text{or } M_I \ddot{\vec{R}}_I = -\nabla_I V_l^{BO}\left(\{\vec{R}_I(t)\}\right) \tag{2.2.4}$$

Overall, in the classical nuclei approximation, the potential energy surface defining the nuclear motion is obtained by solving time-independent electronic Schrodinger equation for l^{th} state at a given nuclei configuration $\{\vec{R}_I(t)\}$. Then, the forces acting on the atoms are computed according to equation 2.2.4, the nuclei are shifted to the new positions, the electronic Hamiltonian $\mathcal{H}_e$ is rediagonalized, and the forces corresponding to the l^{th} electronic eigenstate are computed. This mechanism leads to the independence of the electronic structure of the system on the previous history; effectively, there is no electronic dynamics in such picture. This has two main consequences: first, the time steps in Born-Oppenheimer molecular dynamics can be set large enough to resolve atomic motion – that is, on the order of femtoseconds, whereas resolving electronic dynamics explicitly would require time steps on the order of tens of attoseconds. Second corollary is that the dynamics becomes ill-defined when the qualitative character of the l^{th} electronic eigenstate changes; this happens in the regions where the potential energy surfaces cross. Failure to identify the presence of such regions can make the simulation results meaningless – however, such crossings are widespread, especially when considering

trajectories corresponding to the chemical reactions. Therefore, explicit coupling of electronic and ionic dynamics is required to treat chemical reactions; one of the methods for doing that is known as Ehrenfest dynamics, which we will consider in the following section.

2.2.2. Ehrenfest dynamics

The derivation of the Ehrenfest dynamics starts by assuming the mean-field approximation to the total wavefunction:

$$\Phi\left(\{\vec{r}_i\},\{\vec{R}_I\},t\right) \approx \Psi(\{\vec{r}_i\};\, t)\chi\left(\{\vec{R}_I\};\, t\right)\exp\left(i\int_{t_0}^{t}\tilde{E}_e(t')dt'\right) \tag{2.2.5}$$

with the phase factor $\tilde{E}_e$ involving the electronic Hamiltonian $\mathcal{H}_e$: $\tilde{E}_e = \int \Psi^*(\{\vec{r}_i\};\, t)\chi^*\left(\{\vec{R}_I\};\, t\right)H_e\Psi(\{\vec{r}_i\};\, t)\chi\left(\{\vec{R}_I\};\, t\right)d\vec{R}d\vec{r}$, and normalization of both electronic and nuclear wavefunctions at all moments of time: $\int \Psi^*(\{\vec{r}_i\};\, t)\Psi(\{\vec{r}_i\};\, t)d\vec{r} = 1, \int \chi^*\left(\{\vec{R}_I\};\, t\right)\chi\left(\{\vec{R}_I\};\, t\right)d\vec{R} = 1$. It is important to note that "mean-field" here refers to the absence of explicit correlation between electron and ion wavefunction, unlike in equation 2.1.6; the electronic part of the wavefunction can take electron-electron static and dynamic correlation into account explicitly. The resulting equations of motion for electronic Ψ and nuclear χ parts, under the condition of conservation of total energy, are:

$$\begin{aligned} i\frac{\partial\Psi}{\partial t} &= -\sum_i \frac{1}{2}\nabla_i^2\Psi \\ &\quad + \left(\int \chi^*\left(\{\vec{R}_I\};\, t\right)V_{n-e}\left(\{\vec{r}_i\},\{\vec{R}_I\}\right)\chi\left(\{\vec{R}_I\};\, t\right)d\vec{R}\right)\Psi \end{aligned} \tag{2.2.6}$$

$$i\frac{\partial\chi}{\partial t} = -\sum_I \frac{\hbar^2}{2M_I}\nabla_I^2\chi + \left(\int \Psi^*(\{\vec{r}_i\};\, t)\mathcal{H}_e\left(\{\vec{r}_i\},\{\vec{R}_I\}\right)\Psi(\{\vec{r}_i\};\, t)d\vec{r}\right)\chi \tag{2.2.7}$$

Those equations comprise the time-dependent self-consistent field method[4] – in essence, the ions and electrons move in the potentials obtained by averaging over the other degrees of

freedom. Accordingly, in order to get to the classical limit of point charge nuclei, one can carry out the transformation to phase and amplitude for the nuclear wavefunction, and arrive at a Jacobi-like equation similar to the above (equation 2.2.2):

$$\frac{\partial S}{\partial t} + \sum_I \frac{1}{2M_I}(\nabla_I S)^2 + \int \Psi^* \mathcal{H}_e \Psi d\vec{r} = 0 \tag{2.2.8}$$

and the corresponding Newton-type equation

$$M_I \ddot{\vec{R}}_I = -\nabla_I \int \Psi^* \mathcal{H}_e \Psi d\vec{r} = -\nabla_I V_e^E\left(\{\vec{R}_I(t)\}\right) \tag{2.2.9}$$

Overall, the ions move in the classical potential induced by the electronic wavefunction which now originates from the time-dependent single particle equation. The last substitute to make is to replace the nuclear density with delta-functions in the classical limit, and with that the equation for the electronic dynamics becomes

$$\begin{aligned} i\hbar \frac{\partial \Psi}{\partial t} &= -\sum_i \frac{\hbar^2}{2m_e}\nabla_i^2 \Psi + V_{n-e}\left(\{\vec{r}_i\}, \{\vec{R}_I\}\right)\Psi \\ &= \mathcal{H}_e\left(\{\vec{r}_i\}, \{\vec{R}_I(t)\}\right)\Psi\left(\{\vec{r}_i\}, \{\vec{R}_I\}; t\right) \end{aligned} \tag{2.2.10}$$

It is important to note that this approximation can involve multi-determinant electronic wavefunction Ψ, which can be, for example, expanded in terms of a complete basis of the eigenstates of the electronic Hamiltonian $\mathcal{H}_e$: $\mathcal{H}_e\left(\{\vec{r}_i\}; \{\vec{R}_I\}\right)\Psi_l(\{\vec{r}_i\}) = E_k\left(\{\vec{R}_I\}\right)\Psi_l(\{\vec{r}_i\})$:

$$\Psi\left(\{\vec{r}_i\}, \{\vec{R}_I\}; t\right) = \sum_{l=1}^{\infty} c_l(t)\Psi_l\left(\{\vec{r}_i\}; \{\vec{R}_I\}\right) \tag{2.2.11}$$

With such approximation, the time-dependent coefficients will represent the populations of the adiabatic Hamiltonian eigenstates. Considering the simplest case of electronic system residing in the ground state, one obtains the expression for forces corresponding to the Born-

Oppenheimer dynamics (equation 2.2.4, $l = 0$). Upon plugging the expansion 2.2.11 in the equations for electronic (2.2.10) and ionic (2.2.9) dynamics, one obtains:

$$i\hbar \frac{dc_l(t)}{dt} = c_l(t)E_l - i\hbar \sum_k c_k(t) D^{kl} \tag{2.2.12}$$

$$\begin{aligned} M_I \ddot{\vec{R}}_I = -\nabla_I \int \Psi^* \mathcal{H}_e \Psi d\vec{r} = M_I \ddot{\vec{R}}_I = -\nabla_I \sum_k |c_k(t)|^2 E_k \\ = -\sum_k |c_k(t)|^2 \nabla_I E_k + \sum_{k,l} c_k^* c_l (E_k - E_l) \vec{d}_I^{kl} \end{aligned} \tag{2.2.13}$$

where D^{kl} are the non-adiabatic coupling elements given by

$$D^{kl} = \int \Psi_k^* \frac{\partial}{\partial t} \Psi_l d\vec{r} = \sum_I \dot{\vec{R}}_I \cdot \int \Psi_k^* \nabla_I \Psi_l d\vec{r} = \sum_I \dot{\vec{R}}_I \cdot \vec{d}_I^{kl} \tag{2.2.14}$$

after utilizing the chain rule. From the above expressions one can see that the mean-field approach to the coupled electron-ion dynamics, along with the classical nuclei approximation, allows to capture the transitions between adiabatic ground states. Moreover, the forces acting on atoms are averaged over several adiabatic states, with additional terms due to non-adiabaticity of the wavefunction. This puts limits on the applicability of the method compared to other approaches, for example surface hopping;[5-8] we will discuss those limitations in the next chapter.

Overall, the key difference between the Ehrenfest and Born-Oppenheimer dynamics is the explicit presence of both electronic and ionic dynamics in the former. Due to that, the time steps in the Ehrenfest dynamics have to be small enough to resolve the electronic dynamics – that is, order of attoseconds. On the other hand, the diagonalization of electronic Hamiltonian at each time step is avoided – and since it can be computationally demanding step, scaling as 3rd power of system size for simpler electronic structure methods to above 5th power for more advanced multi-reference approaches, for large systems Ehrenfest dynamics ends up being more

computationally efficient. Moreover, due to inclusion of non-adiabatic effects, this method is the preferred one for simulation of chemical processes. We will consider the methods for dealing with electronic structure in the following discussion.

2.3. Electronic structure methods

As discussed above, after decoupling of the time evolution of nuclei and electrons by assuming that the electrons react instantaneously to any change in the position of nuclei, the electronic part of the problem may be expressed in terms of time-independent Schrodinger equation (TISE),

$$\mathcal{H}_e\left(\{\vec{r}_i\},\{\vec{R}_I\}\right)\Psi(\{\vec{r}_i\}) = E\Psi(\{\vec{r}_i\}), \tag{2.3.1}$$

where $\Psi(\{\vec{r}_i\})$ is the *electronic* many-body wavefunction, and the many-body Hamiltonian $\mathcal{H}_e\left(\{\vec{r}_i\},\{\vec{R}_I\}\right)$ includes terms corresponding to kinetic energy of electrons, electron-electron interaction, electron-nuclei interaction, and the nuclear repulsion:

$$\begin{aligned}\mathcal{H}_e\left(\{\vec{r}_i\},\{\vec{R}_I\}\right) = &-\sum_i \frac{1}{2}\nabla_i^2 + \sum_{i<j}\frac{1}{|\vec{r}_i - \vec{r}_j|} - \sum_{i,I}\frac{Z_I}{|\vec{r}_i - \vec{R}_I|} \\ &+ \sum_{I<J}\frac{Z_J Z_I}{|\vec{R}_J - \vec{R}_I|}\end{aligned} \tag{2.3.2}$$

The main problem with this expression is that the wavefunction $\Psi(\{\vec{r}_i\})$ depends on the positions of all electrons in the system – for n electrons the number of degrees of freedom is $3n$, and the TISE is an eigenvalue problem for a differential operator in $3n$ coordinates. This makes analytical solution impossible for more than one electron (even for helium atom the closed-form solution is not known), and numerical solution challenging as well for many-body systems due to the necessity of dealing with high-dimensional discrete grids.

Several approaches exist to tackle this problem. The first group of methods, so-called mean-field methods, assumes that the many-body wavefunction can be factorized into one-body contributions – orbitals. This implies that the electron-electron two-body interaction term is replaced by an effective potential, which describes the electrostatic interaction of electrons in a mean-field manner: each electron interacts with the average electrostatic potential created by all other electrons. This approximation obviously neglects the instantaneous correlation of electrons movement, which is induced by two-body Coulomb term; however, various approaches to recovering these so-called dynamic correlations exist. The main drawback of these methods is extremely unfavorable scaling with the number of electrons in the system – for those including dynamic correlation computational complexity scales as at least 5^{th} power of the number of electrons in the system, and in the worst cases the scaling is factorial.[9,10]

Another approach is based on the notion that the wavefunction in only necessary for computation of observable quantities, and by itself is only an auxiliary tool. Therefore, one can attempt to infer the properties of the system in question from an observable quantity – for example, density of electrons. This group of methods is known as density functional theory (DFT), which will discuss in greater detail below.

2.3.1. Origins of DFT

The main idea of DFT is that all the observable quantities can be expressed in terms of the electron density of the system; this approach dates back to early work of Thomas and Fermi,[11,12] however it was not until the work of Hohenberg and Kohn[13] that this connection was proven formally. More specifically, it was shown that there is unique correspondence between the electron density of the system and the external potential (that is, Coulomb potential of the nuclei and electric and magnetic fields, if present); moreover, the true ground state electron

density minimizes the total energy of the system. However, the exact expression for total energy as a function of electron density was not provided.

Out of the three components of the total energy of electronic system – kinetic energy, electron-electron interaction, and electron-ion interaction, the kinetic energy term was especially challenging to express in terms of density – that leads to, for example, the failure to predict chemical bonding.[14] A workaround for this problem was proposed in the work of Kohn and Sham[15] by introducing a reference system of non-interacting negatively charged fermions. Due to absence of direct many-body interactions the behavior of such reference was governed by single-particle equations, which made the solutions tractable. The density of particles in such system is supposed to reproduce the density of electrons in actual system under consideration accurately, so the total energy $E[n(\vec{r})]$ can be calculated directly from the quasi-particle density $n(\vec{r})$:

$$E[n(\vec{r})] = T[n(\vec{r})] + \int n(\vec{r}) v_{ext}(\vec{r}) d\vec{r} + \frac{1}{2}\int\int \frac{n(\vec{r})n(\vec{r}')}{|\vec{r}-\vec{r}'|} d\vec{r} d\vec{r}' + E_{XC}[n(\vec{r})] \tag{2.3.3}$$

where $T[n(\vec{r})]$ is the kinetic energy of the system of non-interacting fermions, $v_{ext}(\vec{r})$ is the external potential denoting the interaction with nuclei, and $E_{XC}[n(\vec{r})]$ is the exchange-correlation energy, meaning of which will be explained below. A substantial simplification is achieved in the kinetic energy term due to the fact that for a system of N non-interacting particles described by the single-body wavefunctions $\varphi_i(\vec{r})$ the kinetic energy $T[n(\vec{r})]$ can be expressed as

$$T[n(\vec{r})] = \frac{1}{2}\sum_i \int \varphi_i(\vec{r}) \nabla^2 \varphi_i(\vec{r}) d\vec{r} = \frac{1}{2}\sum_i \langle \varphi_i(\vec{r}) | \nabla^2 | \varphi_i(\vec{r}) \rangle \tag{2.3.4}$$

The exchange-correlation energy has three components in it. First is the difference between the kinetic energies of the reference system of non-interacting fermions, equation 2.3.4,

and the actual kinetic energy of electrons. Second is the correction for dynamic correlation of motion of electrons and associated lowering of the energy (since electrons tend to avoid each other due to Coulomb repulsion). Third term is the so-called "exchange" energy which is supposed to mitigate two effects: first, electrons are fermions and due to the associated many-body wavefunction being antisymmetric the probability of finding two electrons of the same spin in a given point in space is exactly zero (this effect is called exchange hole; for the case of bosons the probability, on the contrary, is enhanced); second, in the expression for the Coulomb interaction of reference fermions the contribution from any given particle is counted twice – that is, there is extra so-called "self-interaction" energy.

Upon variation of the energy functional $E[n(\vec{r})]$ with respect to the density of fermions, under the constraint of conservation of the number of particles in the system, the single-particle eigenvalue equations for the wavefunctions $\varphi_i(\vec{r})$ can be obtained, known as Kohn-Sham equations:

$$\begin{aligned} \mathcal{H}_{KS}\varphi_i(\vec{r}) &= \left(-\frac{1}{2}\nabla^2 + v_{ext}(\vec{r}) + \int \frac{n(\vec{r}')}{|\vec{r}-\vec{r}'|}d\vec{r} + v_{XC}(\vec{r})\right)\varphi_i(\vec{r}) \\ &= \varepsilon_i\varphi_i(\vec{r}) \end{aligned} \tag{2.3.5}$$

There are two non-trivial components to this equation: the external potential $v_{ext}(\vec{r})$ is typically expressed in terms of the pseudopotentials; there exist various approximations to the exchange-correlation potential $v_{XC}(\vec{r})$. Next, the orbitals $\varphi_i(\vec{r})$ are typically not just stored on a discretized grid – a more effective approach involving localized orbitals can be used, as well as plane-wave expansion, typically invoked when studying materials. We discuss those approximations below.

The Kohn-Sham equations have an important caveat in them: the Kohn-Sham effective Hamiltonian depends on the density of fermions in the systems. Therefore, the equations need to

be solved iteratively: an initial approximation to the density of electrons in the system is created, typically as a superposition of atomic densities, the Hamiltonian is formed, the eigenvalue problem is solved for a new set of orbitals φ_i, corresponding density is formed according to $n(\vec{r}) = \sum_i |\varphi_i(\vec{r})|^2$, and a new Hamiltonian is formed. This process is continued until the density which was used to create the Hamiltonian matches the density corresponding to the orbitals induced by that Hamiltonian. This condition is called "self-consistency"; often additional conditions for convergence of the total energy are imposed. It is important to note that the iterations can be more involved; for example, the creation of the new density involves creating a linear combination of densities from the previous steps. One of the widely used methods for that is the direct inversion of the iterative subspace (DIIS) algorithm.[16,17] Another optimization technique relies on the fact that the number of occupied orbitals is often much smaller than the size of the Kohn-Sham Hamiltonian matrix; therefore, an iterative diagonalization techniques for lowest eigenpairs, such as Davidson algorithm[18,19] or residual minimization (RMM-DIIS),[20] can be used.

2.3.2. DFT functionals

The main uncertainty in the Kohn-Sham equations is the definition of the exchange-correlation functional, which is typically specified as a derivative of the exchange-correlation energy: $V_{XC}(\vec{r}) = \frac{\delta E_{XC}}{\delta n(\vec{r})}$. Since Hohenberg-Kohn theorems only specify that the functionals exist and don't provide any guidance for obtaining the actual expressions, a number of approximations are used these days. Generally, the functional is split into two components, exchange and correlation: $E_{XC}[n(\vec{r})] = E_X[n(\vec{r})] + E_C[n(\vec{r})]$. The approximations for those components can be categorized in three ranks.

First is the local density approximation – the underlying assumption is that the electron density is a homogeneous electron gas, and so the exchange and correlation energies are functions of slowly changing electron density. With those assumptions, the exchange term can be written as:[4]

$$E_X[n(\vec{r})] = -2^{\frac{1}{3}} C_X n(\vec{r})^{\frac{1}{3}} \left((1+\xi)^{\frac{4}{3}} + (1-\xi)^{\frac{4}{3}} \right) \quad (2.3.6)$$

where $\xi = \frac{n_\alpha(\vec{r}) - n_\beta(\vec{r})}{n_\alpha(\vec{r}) + n_\beta(\vec{r})}$ is the spin polarization density, and $C_X = \frac{3}{4}\left(\frac{3}{\pi}\right)^{\frac{1}{3}}$. The correlation energy is obtained by an interpolation formula.[21]

The second rank is the general gradient approximation, which includes terms depending on density gradient. One of the first functionals of this type is BP86, and the most used version currently is PBE functional, which does not contain any empirical parameters.[22-24]

The next step are so-called "hybrid" methods, for which the exchange energy is computed analogously to Fock exchange term in the Hartree-Fock method.[25,26] The premise is that the main problem of DFT is the self-interaction of electrons, which is absent in the Hartree-Fock method; therefore, supposedly the inclusion of the "exact" exchange will correct for the double-counting. The functionals are typically parameterized using data from available experimental compounds. An additional modification often done is the splitting of exchange component into short-and long-range terms,[27-29] which allows to recover the correct $\frac{1}{r}$ asymptotical behavior of the electron-electron interaction; such modifications are especially relevant for computing excited states, since it allows to mitigate the appearance of low-energy charge-transfer states.

2.3.3. Pseudopotential method

The pseudopotential method is based on the idea that the electrons in the system under consideration can be partitioned into the valence electrons, actively participating in the formation of chemical bonds, and core electrons, which undergo only slight rearrangements.[29-30] That allows to simplify the electronic structure problem by "packaging" the given atomic nucleus and its core electrons together, and considering explicitly only the valence electrons moving in the field of nuclear potential screened by core electrons. Apart from a reduction of the size of the problem in question, the main benefit of this operation is that the resulting wavefunctions for valence electrons are much smoother, since the condition of orthogonality to the core orbitals does not have to be met anymore. More specifically, if we denote the valence orbitals as $|\psi_v\rangle$, and core orbitals as $|\psi_c\rangle$, then the valence orbitals can be expressed as a sum of a smooth "pseudo-orbital" $|\varphi_v\rangle$ and the term arising from condition of orthogonality to core orbitals:

$$|\psi_v\rangle = |\varphi_v\rangle - \sum_c |\psi_c\rangle\langle\psi_c|\varphi_v\rangle \tag{2.3.7}$$

With that, the Schrodinger equation for the smooth $|\varphi_v\rangle$ can be written as:

$$\mathcal{H}_{KS}|\varphi_v\rangle = E_v|\varphi_v\rangle + \sum_c (E_c - E_v)\langle\psi_c|\varphi_v\rangle|\psi_c\rangle \tag{2.3.8}$$

from where it follows that the smooth function satisfies Shrodinger equation with energy-dependent "pseudo-Hamiltonian"[31] $\mathcal{H}_{PK} = \mathcal{H}_{KS} - \sum_c (E_c - E)|\psi_c\rangle\langle\psi_c|$. Analogously, the pseudopotential can be defined as

$$\widehat{\omega}_{PK}(E) = \hat{v} - \sum_c (E_c - E)|\psi_c\rangle\langle\psi_c| \tag{2.3.9}$$

where $\hat{v}$ is the Coulomb potential of the nucleus. The terms in the projector sum decay at a sufficient distance from the nucleus, and there the pseudo-wavefunction $|\varphi_v\rangle$ matches the actual

valence orbital $|\psi_v\rangle$. The energy of interaction with the nuclei can be expressed with such approximations as:

$$\int \frac{Z_A}{r_{1A}} |\psi_v(\overrightarrow{r_1})|^2 d\overrightarrow{r_1} = \int \varphi_v^*(\overrightarrow{r_1}) \omega(\overrightarrow{r_1}, \overrightarrow{r_2}) \varphi_v(\overrightarrow{r_2}) d\overrightarrow{r_1} d\overrightarrow{r_2} \tag{2.3.10}$$

where $\omega(\overrightarrow{r_1}, \overrightarrow{r_2})$ is the nonlocal pseudopotential, which can be further broken down into a sum of components acting on different angular momentum channels: $\omega(\overrightarrow{r_1}, \overrightarrow{r_2}) = \sum_l \sum_{m=-l}^{l} Y_{lm}^*(\hat{r_1}) \omega_l(r_1, r_2) Y_{lm}(\hat{r_2})$, where Y_{lm} are spherical harmonics, and $\omega_l(r_1, r_2)$ the pseudopotential component for a given angular momentum channel l.

The typical approach to the creation of the pseudopotentials goes as following: First, the atomic Schrodinger equation is solved with explicit consideration of all electrons. Then, the pseudo-wavefunction is parameterized using one of the available ansatzs[32-33] by imposing the following conditions: it has to match the full-electron solution outside a sphere of a pre-specified radius r_c, and has to preserve the norm of the all-electron solution within that sphere:

$$\begin{cases} R_l^{PP}(r_1) = R_{nl}^{AE}(r_1), & \text{for } r_1 > r_l \\ \int_0^{r_l} r^2 |R_l^{PP}(r_1)|^2 \, dr_1 = \int_0^{r_l} r^2 \left| R_l^{AE}(r_1) \right|^2 dr_1, & \text{for } r_1 < r_l \end{cases} \tag{2.3.11}$$

where R_l^{PP} is the radial part of the pseudo-wavefunction with angular momentum l, and R_{nl}^{AE} is the all-electron atomic orbital. Finally, the pseudopotential is obtained by inverting the Schrodinger equation for a given reference energy.

The norm-conserving potentials are written in the Kleinman-Bylander[34-35] separable form for higher computational efficiency. For that, a "local" channel of angular momentum $\omega_{local}(r_1) = -Z_{core}/r_1$ is chosen, and for the remaining angular momentum channels the difference term is defined: $\Delta\omega_l(r_1) = \omega_l(r_1) - \omega_{local}(r_1)$. With that, the pseudopotential is written as

$$\omega_l^{KB}(\overrightarrow{r_1},\overrightarrow{r_2}) = \omega_{local}(r_1) + \sum_l \Delta\omega_l^{KB}(\overrightarrow{r_1},\overrightarrow{r_2})$$

$$= \omega_{local}(r_1) \tag{2.3.12}$$

$$+ \sum_l \sum_{m=-l}^{l} \frac{\varphi_{lm}(\overrightarrow{r_1})\Delta\omega_l(r_1)\Delta\omega_l(r_2)\varphi_{lm}(\overrightarrow{r_2})}{\int \Delta\omega_l(r_1)|\varphi_{lm}(\overrightarrow{r_1})|^2 d\overrightarrow{r_1}}$$

However, as was shown by Vanderbilt,[36] the norm conservation condition is not mandatory, since the "missing" charge can be added back to the valence electron density. The resulting "ultra-soft" pseudopotentials produce much smoother pseudo-wavefunctions, reducing the computational requirements and improving the transferability of the generated pseudopotentials. However, transferable pseudopotentials require several potential components for an angular momentum channel which makes the construction and testing of such potentials a highly non-trivial process.

A further development is the widely used projector augmented wave (PAW) method.[37-39] This method preserves the transferability of the ultrasoft pseudopotentials while avoiding the complicated step of "pseudization" of the additional charge. This is achieved by connecting the pseudo-wavefunctions $|\varphi_n\rangle$ directly to the all-electron wavefunctions $|\psi_n\rangle$ through a linear transformation:

$$|\psi_n\rangle = |\varphi_n\rangle + \sum_i \left(|\phi_i\rangle - |\tilde{\phi}_i\rangle\right)\langle\tilde{p}_i|\tilde{\psi}_n\rangle \tag{2.3.13}$$

where $|\phi_i\rangle$ are all-electron partial waves, typically obtained from solving radial Schrodinger equation for a given atom, $|\tilde{\phi}_i\rangle$ are the pseudized partial waves, orthogonal to the core states and matching the all-electron partial waves outside a sphere of a given core radius r_c^l, and the projector functions $\langle\tilde{p}_i|$ are defined through $\langle\tilde{p}_i|\tilde{\phi}_j\rangle = \delta_{ij}$. With that, the all-electron charge density can be exressed as a sum of pseudo-orbital density and the atom-centered terms.

Analogously, the expression for the total energy of all-electron system can be written down explicitly. This possibility of working directly with all-electron energy and forces while explicitly treating only pseudized soft wavefunctions, is a substantial advantage of the PAW method. It can also be shown that ultrasoft pseudopotential method can be obtained from the PAW energy expression by expansion in terms of on-site charge densities to first order.[37]

2.3.4. Hubbard correction

The DFT approach incorrectly describes the electronic structure of the strongly correlated transition-metal compounds.[40] The reason for that is the incomplete cancellation of electron self-interaction through correlation and exchange terms; another interpretation is that the total energy as a function of the number of electrons in the system is a concave function in the semi-local functional approximation.[41] One of the potential solutions to this problem is to partition the electronic system into the delocalized bands with predominant *s*- and *p*-character and localized *d*-orbitals of the transition metal. Then a Hubbard-like model Hamiltonian can be written for the localized electrons, which results in an additional energy term or the type:

$$E^U = \frac{U}{2}\sum_{i\neq j} n_{ii}n_{jj} \tag{2.3.14}$$

where n_{ii} is the population of i^{th} localized level,[42-43] that is, the corresponding diagonal element of the density matrix. This expression, however, is not rotationally invariant; a version accounting for that was shown to look like:[44-45]

$$E^U = \frac{1}{2}\sum_{ijkl} U_{ijkl}\left[n_{ik}n_{jl} - n_{il}n_{jk}\right] \tag{2.3.15}$$

where n_{ik} is the density matrix element corresponding to i^{th} and k^{th} localized levels; U_{ijkl} is the matrix element of the two-body screened Coulomb operator.

Another issue is that a part of the screened Coulomb interaction is already included in the DFT exchange-correlation functional; therefore, a double-counting correction needs to be introduced.[46-47] With an additional approximation of spherically-symmetric Coulomb interaction potential, the correction results in additional term to the system Hamiltonian:

$$V_{ij}^{sym} = \frac{\delta E^U}{\delta n_{ij}} = U_{eff}\left[\frac{\delta_{ij}}{2} - n_{ij}\right] \tag{2.3.16}$$

where U_{eff} a model parameter typically chosen by the user in a way to reproduce a certain property of the system under consideration, typically the bandgap. Alternatively, it can be obtained by studying the response of the system to external electric field.[41]

2.3.5. Density functional theory for solids – periodic boundary conditions

The periodic boundary conditions is typically invoked when the electronic structure and properties of the extended systems need to be studied.[48-50] This allows to remove the finite-size effects, and at the same time treat the irregularities in the crystal structure through the usage of the supercell method. We will consider the implications of imposing periodic boundary conditions on a system with Bravais lattice basis vectors $\{\overrightarrow{a_1}, \overrightarrow{a_2}, \overrightarrow{a_3}\}$ and corresponding reciprocal lattice vectors $\{\overrightarrow{b_1}, \overrightarrow{b_2}, \overrightarrow{b_3}\}$. Upon a translation $\vec{R} = n_1\overrightarrow{a_1} + n_2\overrightarrow{a_2} + n_3\overrightarrow{a_3}$, where n_i are integers, the potential induced by the crystal lattice is not changed: $v_{ext}(\vec{r} + \vec{R}) = v_{ext}(\vec{r})$. Therefore, the effective Hamiltonian commutes with the translation operator and the eigenstates can be classified according to the irreducible representations of translation group. Such representations are denoted by a vector $\vec{k}$, which in this case serves as an additional label together with the normal ordinal number n. Moreover, according to Bloch theorem the single-particle eigenstates can be written in the form

$$\psi_{n\vec{k}}(\vec{r}) = \exp\left(i\vec{k}\cdot\vec{r}\right) u_{n\vec{k}}(\vec{r}) \tag{2.3.17}$$

where $u_{n\vec{k}}\left(\vec{r}+\vec{R}\right) = u_{n\vec{k}}(\vec{r})$ is a periodic function. In principle, the translation group is infinite, and the number of allowed vectors $\vec{k}$ is infinite as well; in practice, however, a finite Born-Karman periodic boundary condition is introduced: $\psi(\vec{r}+N_i\overrightarrow{a_i}) = \psi(\vec{r})$, where N_i are integers, and $N = N_1N_2N_3$ is the number of cells in the resulting model crystal. It can be shown that the allowed label vectors in the first Brillouin zone are $\vec{k} = \sum_{i=1}^{3}\frac{m_i}{N_i}\overrightarrow{b_i}$, with integers m_i and reciprocal lattice vectors $\overrightarrow{b_i}$. It is important to note that the number of the allowed k-points is equal to the number of unit cells in the model crystal on which the periodic boundary conditions are imposed.

Periodic wavefunctions $\psi_{nk}(\vec{r})$ can be expanded in terms of plane waves with $\vec{k}$ vectors of the form shown above: $\psi_{n\vec{k}}(\vec{r}) = \sum_{\vec{q}} c_{\vec{k},n}(\vec{q})\exp\left(i\vec{k}\cdot\vec{r}\right)$. However, upon plugging this ansatz in the single-body Schrodinger equation: $-\frac{\hbar^2}{2m}\nabla^2\psi + V_{n-e}(\vec{r})\psi = \varepsilon\psi$ it can be shown that for periodic potential only the terms $\vec{q}$ differing by a reciprocal lattice vector $\vec{K} = n_1\overrightarrow{b_1} + n_2\overrightarrow{b_2} + n_3\overrightarrow{b_3}$ remain in the expression and it takes the following form: $\psi_{nk}(\vec{r}) = \exp\left(i\vec{k}\cdot\vec{r}\right)\sum_{\vec{K}} c_{\vec{k}-\vec{K}}\ \exp\left(-i\vec{K}\cdot\vec{r}\right)$. This is the "plane-wave" expansion oftentimes used in the modern software packages.[51] The pseudopotential method described above serves to make this task easier by considering "smooth" pseudo-wavefunctions with rapidly converging plane-wave expansions. The electronic structure problem for the model crystal is then separated in a number of separate problems for the allowed vectors $\vec{k}$, called "*k*-points":

$$\sum_{\vec{K'}}\left(\frac{1}{2}\left|\vec{k}+\vec{K}\right|^2\delta_{\vec{K},\vec{K'}}+\omega\left(\vec{k}+\vec{K},\vec{k}+\vec{K'}\right)+v_H\left(\vec{K}-\vec{K'}\right)\right.$$
$$\left.+\,v_{XC}\left(\vec{K}-\vec{K'}\right)\right)c_{\vec{k},n}\left(\vec{K'}\right)=\varepsilon_{\vec{k},n}\,c_{\vec{k},n}\left(\vec{K}\right) \tag{2.3.18}$$

Overall system properties, for example, total energy or electronic density, are then computed as a weighted sum over the *k*-points used in the calculation; the schemes for optimal sampling of the first Brillouin zone have been developed.[52] Overall, the choice of density of *k*-point sampling and the cutoffs for the expansion of the wavefunction are the two most important parameters determining the accuracy of the calculation in the density functional theory computation under periodic boundary conditions.

2.3.6. Localized basis set

Oftentimes, however, the plane-wave computations become prohibitive due to the size of the system in question – this happens since solution of the Kohn-Sham equations in such codes involves diagonalization of a matrix of size $N_{PW} \times N_{PW}$, where N_{PW} is the number of plane waves used, which in turn scales as a product of system size for a given cutoff. Most modern codes use fairly advanced techniques, such as projection on the subspace of the size $N_{bnd} \times N_{bnd}$, where N_{bnd} is the number of bands, occupied and empty, in the system, and an iterative diagonalization of the resulting smaller problem; however for the number of atoms exceeding several hundreds to a thousand the computational resources required become prohibitive. A workaround is to make the involved Hamiltonian matrices sparse, which allows to take advantage of more favorably scaling algorithms. This is achieved by introducing the localized basis set, as implemented in SIESTA[53] software. In the following, we briefly summarize the most important aspects of that approach.

The sparsity of the system Hamiltonian and overlap matrix is achieved by using localized basis functions, which are enforced to being zero beyond a certain cutoff threshold. Such strict localization is achieved by splitting the basis orbitals into an angular and radial parts, and explicitly storing the numerical radial part (which goes to zero beyond a strict cutoff) on a grid. The values in between the radial mesh points are interpolated by cubic splines. The desired cutoff radius and the shape of the orbital can be specified by the user. The commonly used in quantum chemistry split-valence basis sets are supported, as well as polarization orbitals.

The pseudopotentials are used in the norm-conserving non-local Kleinman-Bylander[34] form mentioned previously, except for the local part of the pseudopotential being set to a specifically optimized function rather than to one of the angular momentum channels. One projector per angular momentum channel is used in most cases, pseudized with the energy equal to atomic orbital eigenvalue.

The terms in the Kohn-Sham equation corresponding to the kinetic energy and the local part of the pseudopotential are computed in the reciprocal space, and the rest of the terms – on a real-space grid, with grid finesse being controlled by one of the parameters in the computation. The key difference of this code is the computation of the electronic density from the density matrix in the basis of localized atomic orbitals – which is sparse due to their strict localization. Analogously, the total energy of the system is also expressed through the sparse density matrix. Using such techniques, even a linear scaling of the computational intensity with system size can be achieved.

2.3.7. Wannier functions

Under periodic boundary conditions, the Kohn-Sham eigenstates have the translational symmetry of the system under consideration, in addition to being extended throughout the

simulation cell. On the contrary, most of the chemical intuition is generally derived from the localized orbital notation (σ-, π-bonds, etc.). One of the possible approaches to bridging this gap is the localized Wannier functions representation. This approach is implemented under the name of "maximally localized Wannier functions" – MLWF.[54-57] In further discussion, we briefly review the basics of the method to provide ground for understanding of its applications throughout the rest of the work.

Consider a model crystal with periodic boundary conditions imposed over $N_x N_y N_z$ unit cells (and with equal number of k-points). For a single band $\psi_{n\vec{k}}(\vec{r}) = u_{n\vec{k}}(\vec{r}) \exp(i\vec{k} \cdot \vec{r})$, the construction of localized orbital $w_{\vec{R}n}(\vec{r})$ centered in one of the model unit cells translated by a Bravais lattice vector $\vec{R}$ from the origin can be accomplished by an integration over the Brillouin zone with appropriate phase factors:

$$w_{\vec{R}n}(\vec{r}) = \frac{V}{(2\pi)^3} \int_{BZ} \exp(-i\vec{k} \cdot \vec{r})\, \psi_{n\vec{k}}(\vec{r}) d\vec{k} \tag{2.3.19}$$

The resulting set of Wannier functions $w_{\vec{R}n}(\vec{r})$ for $N_x N_y N_z$ values of $\vec{R}$ form an equivalent basis for Schrodinger equation, since the total energy of the model system is invariant under unitary transformations of the basis. An arising complication, however, is the fact that transformation is defined only up to a gauge function $\exp\left(i\phi_n(\vec{k})\right)$, where ϕ_n is periodic in reciprocal space. From the elementary considerations of Fourier transforms theory, one can infer that in order to achieve satisfactory localization in the real space, the Bloch wavefunctions have to be as smooth as possible in k-space; one of the criteria for that can be well-defined gradient in the reciprocal space: $\nabla_{\vec{k}} \psi_{n\vec{k}}(\vec{r})$. Unfortunately, the Kohn-Sham eigenstates don't fulfill this requirement at the points of bands crossing. Therefore, the main challenge is to define a unitary transformation of Bloch eigenstates $\psi_{n\vec{k}}(\vec{r})$ into a group of functions smooth in k-space $\tilde{\psi}_{n\vec{k}}(\vec{r})$

for using in the definition of Wannier function. For a group of J bands this required transformation can be written as:

$$\tilde{\psi}_{n\vec{k}}(\vec{r}) = \sum_{m=1}^{J} U_{mn}^{\vec{k}} \psi_{m\vec{k}}(\vec{r}) \tag{2.3.20}$$

The procedure of finding the Wannier functions consists of defining the appropriate unitary transformation matrix $\mathbf{U}^{\vec{k}}$, and the utilizing the definition of Wannier transformation to get the localized orbitals. There are many possible definitions of this transformation; the one implemented in Wannier90[56] software package is the requirement that the transformation minimizes the spread of Wannier functions obtained after applying the definition of Wannier transformation to the resulting $\tilde{\psi}_{n\vec{k}}(\vec{r})$. The spread is defined as:

$$\Omega = \sum_{\mathrm{n}} \left[\langle w_{\vec{0}n} | r^2 | w_{\vec{0}n} \rangle - \langle w_{\vec{0}n} | r | w_{\vec{0}n} \rangle^2 \right] \tag{2.3.21}$$

where the terms involved can be expressed through matrix elements of the type $\langle \psi_{m\vec{k}} | \nabla_{\vec{k}} | \psi_{n\vec{k}} \rangle$, which can be in turn expressed using the finite-difference approximation through the overlap integrals of the Bloch wavefunctions for neighboring k-points, $\langle \psi_{m\vec{k}} | \psi_{n,\vec{k}+\vec{b}} \rangle$. The minimum of the spread functional with respect to the unitary transformations of Bloch states is found by iterative optimization. The resulting Wannier functions have a lot of resemblance to the traditional σ, π bonds and lone pairs familiar from general chemistry.

2.3.8. Electronic excitations – the delta-SCF (Δ-SCF) method

So far we have discussed the modeling of the electronic ground state of the system of interest. Modeling the excited states of chemical systems has been a subject of active research for decades, and the available methods for the isolated molecules range from fairly simple configuration interaction with single excitations [58] to very sophisticated and unfavorably scaling

approaches like multi-configuration self-consistent field.[59] However, those methods are practically intractable computationally for systems with more than several dozen atoms. For simulation of large systems the most widespread method is linear-response time-dependent density-functional theory (LR-TDDFT).[60-61] In a nutshell, this method amounts to studying the response of the electronic system in question to a small external electric field as a function of field frequency. The spectrum of the excited states of the system can be obtained in such way; the density corresponding to the excited states is constructed of a combination of the occupied and unoccupied Kohn-Sham orbitals of the ground state. The widespread problem of such method is inability to model correctly the charge-transfer excitations due to the incorrect asymptotic behavior of popular exchange-correlation functionals. This shortcoming can be addressed by using so-called range-separated functionals,[61-62] but only to a limited extent.

Another approach is the Δ-SCF method,[63] which involves enforcing a different occupation numbers pattern during the self-consistent calculation. For the ground state of non-metallic systems, the states are fully occupied from the lowest to the highest (HOMO), and the rest left vacant; for metals, Fermi distribution corresponding to the model temperature is enforced. In the case of Δ-SCF, the occupations can be changed at user's discretion; most often, the lowest excited state of the system is modeled by setting the HOMO occupancy to zero and LUMO to one. For metallic systems, more elaborate constructs can be used. This approach has two important advantages: first, the model of the excited state does not depend on the shape of the ground state orbitals – the Kohn-Sham wavefunctions are variationally optimized, which makes it possible to describe the substantial rearrangements of the electronic system; second, with an exact functional this method produces the true densities of the excited states.[64-65] Finally,

this method can be applied to the systems under periodic boundary conditions and coupled with real-time ion and electron dynamics, as we discuss below.

2.4. Implementation of Ehrenfest dynamics with excitations

For the simulation of excited states dynamics we use TDAP-2.0[67] code, which combines the capabilities for creating excitations using the Δ-SCF method and simulating electron-ion dynamics with the Ehrenfest method described above. This code is based on the SIESTA package,[54] thus having support for periodic boundary conditions and localized basis sets. In this chapter we discuss the methodology and implementation details for the case of a localized basis.

2.4.1. Electronic dynamics in the density functional theory approximation

The extension of standard density functional theory to time domain (TDDFT) is based on the Runge-Gross theorems, which establish one-to-one correspondence between the time-dependent electronicdensity and the external potential.[67-68] In this approach instead of minimization of the total energy, the quantum-mechanical action has to be minimized in order to obtain time-dependent Schrodinger equation:

$$A[\Psi] = \int_{t_i}^{t_f} \left\langle \Psi(\mathrm{t}) \middle| i\hbar \frac{\partial}{\partial t} - \mathcal{H}_e(t) \middle| \Psi(t) \right\rangle dt \qquad (2.4.1)$$

The relevant quantity is the action density functional $A[n]$, which has a stationary point at the correct electronic density. The stationary point condition, with utilization of auxiliary reference system of non-interacting particles, leads to the TDDFT equations for the Kohn-Sham orbitals:

$$\mathcal{H}_{KS}(t)\phi_i(\vec{r},t)$$

$$= \left(-\frac{1}{2}\nabla^2 + V_{ext}(\vec{r},t) + V_H(\vec{r},t) + \frac{\delta A_{xc}[n]}{\delta n(\vec{r},t)}\right)\phi_i(\vec{r},t) \tag{2.4.2}$$

$$= i\frac{\partial}{\partial t}\phi_i(\vec{r},t)$$

where the nuclei positions are treated as parameters in the external potential:

$$V_{ext}(\vec{r},t) = -\sum_I \frac{Z_I}{\left|\vec{R}_I(t) - \vec{r}\right|} + \sum_{I<J} \frac{Z_I Z_J}{\left|\vec{R}_I(t) - \vec{R}_J(t)\right|} \tag{2.4.3}$$

In principle, the time-dependent exchange-correlation functional has explicit dependence on time (that is, history of the system), and is nonlocal in space:

$$\frac{\delta A_{xc}[n]}{\delta n(\vec{r},t)} = V_{xc}(\vec{r},t) \tag{2.4.4}$$

In practice, the adiabatic local density approximation is made, which allows to write the exchange-correlation potential as:

$$A_{xc}^{ALDA}[n] = \int_{t_i}^{t_f}\int n(\vec{r}',t')\epsilon_{xc}^{LDA}\big(n(\vec{r}',t')\big)d\vec{r}'dt' \tag{2.4.5}$$

Similar expressions exist for the extension of gradient-corrected functionals into time domain.

Once the initial density and the Kohn-Sham states comprising it $\{\phi_i(t_0)\}$ are known, the system can be propagated in time using the TDKS equation (2.4.2) and the density reconstructed from the Kohn-Sham orbitals:

$$\phi_i(t) = \hat{U}(t,t_0)\phi_i(t_0) \tag{2.4.6}$$

$$\hat{U}(t,t_0) = \hat{T}\exp\left(-i\int_{t_0}^{t}\mathcal{H}_{KS}(t')dt'\right) \tag{2.4.7}$$

where $\hat{T}$ is time-ordering operator, and $\hat{U}(t,t_0)$ is the time evolution operator. We describe the peculiarities associated with implementing those expressions in the localized basis set in the

following section. The nuclei can be propagated classically according to the forces exerted on them by the electronic subsystem.[69-70]

2.4.2. Localized basis set implementation

In the localized basis set dependent on the nuclei positions $\{\chi_j(\{\vec{R}_I(t)\})\}$ the Hamiltonian matrix **H** elements are given by $H_{ij} = \langle\chi_i|\widehat{H}_{KS}|\chi_j\rangle$; analogously, the Kohn-Sham orbitals $\{\phi_i\}$ can be expanded into the basis as vectors $\mathbf{c}_i$, with $|\phi_i\rangle = \sum_k c_i^k|\chi_k\rangle$. In such basis, the Kohn-Sham equations are given by:[66, 71]

$$i\frac{\partial \mathbf{c}_\mathrm{n}}{\partial t} = \mathbf{S}^{-1}(\mathbf{H}+\mathbf{P})\mathbf{c}_n \tag{2.4.8}$$

where **S** is the overlap matrix for that basis, $S_{ij} = \langle\chi_i|\chi_j\rangle$, and **P** is the matrix term due to the motion of the basis functions, $P_{ij} = -i\left\langle\chi_i\middle|\frac{\partial\chi_j}{\partial t}\right\rangle$.

In our work we use the discretized form of $\widehat{U}$ given by

$$\mathbf{U}(t+\Delta t,t) = \exp\left(-i\Delta t\left[\mathbf{S}_{1/2}^{-1}(\mathbf{H}_{1/2}+\mathbf{P}_{1/2})\right]\right) \tag{2.4.9}$$

where subscript ½ indicates matrices taken at the half time step $t + \Delta t/2$. We use the Lanczos iterative procedure in order to obtain the matrix exponential. Each time-propagation step is iterated to self-consistency as described elsewhere.[72]

In Ehrenfest (mean-field) dynamics the nuclei (indexed by *J*) are classical particles moving in a potential approximated by the electronic energy expectation value E_{el}:

$$M_J\frac{\partial^2\vec{R}_J}{\partial t^2} = \overrightarrow{F_J} = -\nabla_{\vec{R}_J}E_{el} \tag{2.4.10}$$

where $\vec{R}_J$ is the position of of J^{th} ion. The forces in this expression conserve the total energy of the system. Within TDDFT and an arbitrary basis, the final expression for the forces is given by:[2]

$$\overrightarrow{F_J} = -\nabla_{\vec{R}_J} \sum_{K \neq J} \frac{Z_J Z_K}{\left|\vec{R}_J - \vec{R}_K\right|} - \sum_n f_n \mathbf{c}_n^* \left(\nabla_{\vec{R}_J} [\mathbf{H} - \mathbf{V}_{dc}]\right) \mathbf{c}_n + \sum_n f_n \mathbf{c}_n^* (\mathbf{H}\mathbf{S}^{-1}\mathbf{D}^J + h.c.)\mathbf{c}_n \quad (2.4.11)$$

where $\mathbf{V}_{dc}$ is the double-counting correction term, $\mathbf{D}_{ij}^J = \left\langle \chi_i \middle| \nabla_{\vec{R}_J} \chi_j \right\rangle$, and f_n are occupation numbers. In the case of adiabatic dynamics the last term reduces to the ordinary Pulay force.

Overall, our approach to the simulation of the photo-induced processes consists of creating an excitation in the model system using the Δ-SCF approach and propagating electrons and ions in time according to the equations above. Such method is expected to work well for systems with high densities of states, such as semiconductor materials surface – which we successfully demonstrate in the next chapter.

2.5. References

1. Marx, D.; Hutler, J., *Ab Initio Molecular Dynamics*; Cambridge University Press: New York, 2009.

2. Doltsinis, N. L. M., D., First Principles Molecular Dynamics Involving Excited States and Nonadiabatic Transitions. *Journal of Theoretical and Computational Chemistry* **2002**, *1*, 319-349.

3. Doltsinis, N. L.; Marx, D., Nonadiabatic Car-Parrinello Molecular Dynamics. *Phys Rev Lett* **2002**, *88*, 166402.

4. Dirac, P. A. M., Note on Exchange Phenomena in the Thomas Atom. *P Camb Philos Soc* **1930**, *26*, 376-385.

5. Barbatti, M.; Granucci, G.; Persico, M.; Ruckenbauer, M.; Vazdar, M.; Eckert-Maksic, M.; Lischka, H., The on-the-Fly Surface-Hopping Program System Newton-X: Application to Ab Initio Simulation of the Nonadiabatic Photodynamics of Benchmark Systems. *J Photoch Photobio A* **2007**, *190*, 228-240.

6. Du, L. K.; Lan, Z. G., An on-the-Fly Surface-Hopping Program Jade for Nonadiabatic Molecular Dynamics of Polyatomic Systems: Implementation and Applications. *J Chem Theory Comput* **2015**, *11*, 1360-1374.

7. Akimov, A. V.; Prezhdo, O. V., The PYXAID Program for Non-Adiabatic Molecular Dynamics in Condensed Matter Systems. *J Chem Theory Comput* **2013**, *9*, 4959-4972.

8. Richter, M.; Marquetand, P.; Gonzalez-Vazquez, J.; Sola, I.; Gonzalez, L., Sharc: Ab Initio Molecular Dynamics with Surface Hopping in the Adiabatic Representation Including Arbitrary Couplings. *J Chem Theory Comput* **2011**, *7*, 1253-1258.

9. Jensen, F., *Introduction to Computational Chemistry*; John Wiley & Sons: Chichester, 2007.

10. Cramer, C., *Essentials of Computational Chemistry*; John Wiley & Sons: Chichester, 2004.

11. Thomas, L. H., The Calculation of Atomic Fields. *P Camb Philos Soc* **1927**, *23*, 542-548.

12. Jones, R. O.; Gunnarsson, O., The Density Functional Formalism, Its Applications and Prospects. *Reviews of Modern Physics* **1989**, *61*, 689-746.

13. Hohenberg, P.; Kohn, W., Inhomogeneous Electron Gas. *Physical Review B* **1964**, *136*, B864-B871.

14. Teller, E., On Stability of Molecules in Thomas-Fermi Theory. *Reviews of Modern Physics* **1962**, *34*, 627-631.

15. Kohn, W.; Sham, L. J., Self-Consistent Equations Including Exchange and Correlation Effects. *Phys Rev* **1965**, *140*, 1133-1138.

16. Pulay, P., Improved Scf Convergence Acceleration. *Journal of Computational Chemistry* **1982**, *3*, 556-560.

17. Pulay, P., Convergence Acceleration of Iterative Sequences. The Case of Scf Iteration. *Chemical Physics Letters* **1980**, *73*, 393-398.

18. Davidson, E. R., Super-Matrix Methods. *Comput Phys Commun* **1989**, *53*, 49-60.

19. Davidson, E. R., The Iterative Calculation of a Few of the Lowest Eigenvalues and Corresponding Eigenvectors of Large Real-Symmetric Matrices. *Journal of Computational Physics* **1975**, *17*, 87-94.

20. Wood, D. M.; Zunger, A., A New Method for Diagonalizing Large Matrices. *J. Phys. A-Math. Gen.* **1985**, *18*, 1343-1359.

21. Vosko, S. H.; Wilk, L.; Nusair, M., Accurate Spin-Dependent Electron Liquid Correlation Energies for Local Spin-Density Calculations - a Critical Analysis. *Can J Phys* **1980**, *58*, 1200-1211.

22. Perdew, J. P., Density-Functional Approximation for the Correlation-Energy of the Inhomogeneous Electron-Gas. *Physical Review B* **1986**, *33*, 8822-8824.

23. Perdew, J. P.; Burke, K.; Ernzerhof, M., Generalized Gradient Approximation Made Simple. *Physical Review Letters* **1996**, *77*, 3865-3868.

24. Perdew, J. P.; Wang, Y., Accurate and Simple Analytic Representation of the Electron-Gas Correlation-Energy. *Physical Review B* **1992**, *45*, 13244-13249.

25. Kummel, S.; Kronik, L., Orbital-Dependent Density Functionals: Theory and Applications. *Reviews of Modern Physics* **2008**, *80*, 3-60.

26. Becke, A. D., Density-Functional Thermochemistry. 3. The Role of Exact Exchange. *Journal of Chemical Physics* **1993**, *98*, 5648-5652.

27. Heyd, J.; Scuseria, G. E., Efficient Hybrid Density Functional Calculations in Solids: Assessment of the Heyd-Scuseria-Ernzerhof Screened Coulomb Hybrid Functional. *Journal of Chemical Physics* **2004**, *121*, 1187-1192.

28. Heyd, J.; Scuseria, G. E.; Ernzerhof, M., Hybrid Functionals Based on a Screened Coulomb Potential. *Journal of Chemical Physics* **2003**, *118*, 8207-8215.

29. Hamann, D. R.; Schluter, M.; Chiang, C., Norm-Conserving Pseudopotentials. *Physical Review Letters* **1979**, *43*, 1494-1497.

30. Yin, M. T.; Cohen, M. L., Theory of Abinitio Pseudopotential Calculations. *Physical Review B* **1982**, *25*, 7403-7412.

31. Phillips, J. C.; Kleinman, L., New Method for Calculating Wave Functions in Crystals and Molecules. *Phys Rev* **1959**, *116*, 287-294.

32. Rappe, A. M.; Rabe, K. M.; Kaxiras, E.; Joannopoulos, J. D., Optimized Pseudopotentials. *Physical Review B* **1990**, *41*, 1227-1230.

33. Kresse, G.; Hafner, J., Norm-Conserving and Ultrasoft Pseudopotentials for First-Row and Transition-Elements. *J. Phys.-Condes. Matter* **1994**, *6*, 8245-8257.

34. Kleinman, L.; Bylander, D. M., Efficacious Form for Model Pseudopotentials. *Physical Review Letters* **1982**, *48*, 1425-1428.

35. Blochl, P. E., Generalized Separable Potentials for Electronic-Structure Calculations. *Physical Review B* **1990**, *41*, 5414-5416.

36. Vanderbilt, D., Soft Self-Consistent Pseudopotentials in a Generalized Eigenvalue Formalism. *Physical Review B* **1990**, *41*, 7892-7895.

37. Kresse, G.; Joubert, D., From Ultrasoft Pseudopotentials to the Projector Augmented-Wave Method. *Physical Review B* **1999**, *59*, 1758-1775.

38. Blochl, P. E., Projector Augmented-Wave Method. *Physical Review B* **1994**, *50*, 17953-17979.

39. Blochl, P. E.; Forst, C. J.; Schimpl, J., Projector Augmented Wave Method: Ab Initio Molecular Dynamics with Full Wave Functions. *B Mater Sci* **2003**, *26*, 33-41.

40. Chan, M. K. Y.; Ceder, G., Efficient Band Gap Prediction for Solids. *Physical Review Letters* **2010**, *105*, 196403.

41. Cococcioni, M.; de Gironcoli, S., Linear Response Approach to the Calculation of the Effective Interaction Parameters in the LDA+*U* Method. *Physical Review B* **2005**, *71*, 035105.

42. Anisimov, V. I.; Aryasetiawan, F.; Lichtenstein, A. I., First-Principles Calculations of the Electronic Structure and Spectra of Strongly Correlated Systems: The LDA+*U* Method. *J. Phys.-Condes. Matter* **1997**, *9*, 767-808.

43. Anisimov, V. I.; Zaanen, J.; Andersen, O. K., Band Theory and Mott Insulators - Hubbard-*U* Instead of Stoner-*I*. *Physical Review B* **1991**, *44*, 943-954.

44. Solovyev, I. V.; Liechtenstein, A. I.; Terakura, K., Is Hund's Second Rule Responsible for the Orbital Magnetism in Solids? *Physical Review Letters* **1998**, *80*, 5758-5761.

45. Shick, A. B.; Drchal, V.; Havela, L., Coulomb-*U* and Magnetic-Moment Collapse in Delta-Pu. *Europhys Lett* **2005**, *69*, 588-594.

46. Dudarev, S. L.; Botton, G. A.; Savrasov, S. Y.; Humphreys, C. J.; Sutton, A. P., Electron-Energy-Loss Spectra and the Structural Stability of Nickel Oxide: An LSDA+*U* Study. *Physical Review B* **1998**, *57*, 1505-1509.

47. Liechtenstein, A. I.; Anisimov, V. I.; Zaanen, J., Density-Functional Theory and Strong-Interactions - Orbital Ordering in Mott-Hubbard Insulators. *Physical Review B* **1995**, *52*, R5467-R5470.

48. Ihm, J.; Zunger, A.; Cohen, M. L., Momentum-Space Formalism for the Total Energy of Solids. *Journal of Physics C-Solid State Physics* **1979**, *12*, 4409-4422.

49. Denteneer, P. J. H.; Vanhaeringen, W., The Pseudopotential-Density-Functional Method in Momentum Space - Details and Test Cases. *Journal of Physics C-Solid State Physics* **1985**, *18*, 4127-4142.

50. Payne, M. C.; Teter, M. P.; Allan, D. C.; Arias, T. A.; Joannopoulos, J. D., Iterative Minimization Techniques for Ab Initio Total-Energy Calculations - Molecular-Dynamics and Conjugate Gradients. *Reviews of Modern Physics* **1992**, *64*, 1045-1097.

51. Giannozzi, P., et al., Quantum Espresso: A Modular and Open-Source Software Project for Quantum Simulations of Materials. *J. Phys.-Condes. Matter* **2009**, *21,* 395502.

52. Monkhorst, H. J.; Pack, J. D., Special Points for Brillouin-Zone Integrations. *Physical Review B* **1976**, *13*, 5188-5192.

53. Soler, J. M.; Artacho, E.; Gale, J. D.; Garcia, A.; Junquera, J.; Ordejon, P.; Sanchez-Portal, D., The SIESTA Method for Ab Initio Order-N Materials Simulation. *J. Phys.-Condes. Matter* **2002**, *14*, 2745-2779.

54. Marzari, N.; Mostofi, A. A.; Yates, J. R.; Souza, I.; Vanderbilt, D., Maximally Localized Wannier Functions: Theory and Applications. *Reviews of Modern Physics* **2012**, *84,* 1419-1475.

55. Marzari, N.; Vanderbilt, D., Maximally Localized Generalized Wannier Functions for Composite Energy Bands. *Physical Review B* **1997**, *56*, 12847-12865.

56. Mostofi, A. A.; Yates, J. R.; Lee, Y. S.; Souza, I.; Vanderbilt, D.; Marzari, N., Wannier90: A Tool for Obtaining Maximally-Localised Wannier Functions. *Comput Phys Commun* **2008**, *178*, 685-699.

57. Souza, I.; Marzari, N.; Vanderbilt, D., Maximally Localized Wannier Functions for Entangled Energy Bands. *Physical Review B* **2002**, *65*, 035109.

58. Sherrill, C. D.; Schaefer, H. F., The Configuration Interaction Method: Advances in Highly Correlated Approaches. In *Advances in Quantum Chemistry, Vol 34*, Lowdin, P. O.; Sabin, J. R.; Zerners, M. C.; Brandas, E., Eds. Elsevier Academic Press Inc: San Diego, 1999; Vol. 34, pp 143-269.

59. Schmidt, M. W.; Gordon, M. S., The Construction and Interpretation of MSCSF Wavefunctions. *Annu Rev Phys Chem* **1998**, *49*, 233-266.

60. Casida, M. E., Time-Dependent Density-Functional Theory for Molecules and Molecular Solids. *Journal of Molecular Structure-Theochem* **2009**, *914*, 3-18.

61. Dreuw, A.; Head-Gordon, M., Failure of Time-Dependent Density Functional Theory for Long-Range Charge-Transfer Excited States: The Zincbacteriochlorin-Bacterlochlorin and Bacteriochlorophyll-Spheroidene Complexes. *Journal of the American Chemical Society* **2004**, *126*, 4007-4016.

62. Neugebauer, J.; Gritsenko, O.; Baerends, E. J., Assessment of a Simple Correction for the Long-Range Charge-Transfer Problem in Time-Dependent Density-Functional Theory. *Journal of Chemical Physics* **2006**, *124*, 214102.

63. Gavnholt, J.; Olsen, T.; Engelund, M.; Schiotz, J., Delta Self-Consistent Field Method to Obtain Potential Energy Surfaces of Excited Molecules on Surfaces. *Physical Review B* **2008**, *78*, 075441.

64. Gorling, A., Density-Functional Theory Beyond the Hohenberg-Kohn Theorem. *Physical Review A* **1999**, *59*, 3359-3374.

65. Cullen, J.; Krykunov, M.; Ziegler, T., The Formulation of a Self-Consistent Constricted Variational Density Functional Theory for the Description of Excited States. *Chem Phys* **2011**, *391*, 11-18.

66. Kolesov, G.; Granas, O.; Hoyt, R.; Vinichenko, D.; Kaxiras, E., Real-Time Td-Dft with Classical Ion Dynamics: Methodology and Applications. *J Chem Theory Comput* **2016**, *12*, 466-476.

67. Marques, M. A. L.; Gross, E. K. U., Time-Dependent Density Functional Theory. *Annu Rev Phys Chem* **2004**, *55*, 427-455.

68. Runge, E.; Gross, E. K. U., Density-Functional Theory for Time-Dependent Systems. *Physical Review Letters* **1984**, *52*, 997-1000.

69. Saalmann, U.; Schmidt, R., Excitation and Relaxation in Atom-Cluster Collisions. *Physical Review Letters* **1998**, *80*, 3213-3216.

70. Kunert, T.; Schmidt, R., Non-Adiabatic Quantum Molecular Dynamics: General Formalism and Case Study H_2^+ in Strong Laser Fields. *European Physical Journal D* **2003**, *25*, 15-24.

71. Kolesov, G.; Vinichenko, D.; Tritsaris, G. A.; Friend, C. M.; Kaxiras, E., Anatomy of the Photochemical Reaction: Excited-State Dynamics Reveals the C–H Acidity Mechanism of Methoxy Photo-Oxidation on Titania. *Journal of Physical Chemistry Letters* **2015**, *6*, 1624-1627.

72. Castro, A.; Marques, M. A. L.; Rubio, A., Propagators for the Time-Dependent Kohn-Sham Equations. *Journal of Chemical Physics* **2004**, *121*, 3425-3433.

CHAPTER 3. WATER PHOTO-OXIDATION ON THE RUTILE TITANIA (110) SURFACE

In this chapter we discuss the simulation of water photo-oxidation; we introduce a framework for the study of photo-catalyzed reactions on semiconductor surfaces based on time-dependent density functional theory that explicitly accounts for the evolution of electronically excited states. Within this approach, we investigate carrier-mediated oxidation of molecularly adsorbed water on the rutile TiO_2 (110) surface. We find that this process is possible in synergy with thermal effects at temperatures between 60 and 100 K only when defects like Ti interstitials are present in the subsurface region. This work was published as G. A. Tritsaris, D. Vinichenko, G. Kolesov, C. M. Friend, E. Kaxiras. Dynamics of the Photogenerated Hole at the Rutile TiO_2 (110)/Water Interface: a Nonadiabatic Simulation Study. Journal of Physical Chemistry C, 2014, V. 118, pp. 27393–27401.

3.1. Introduction

Light-assisted hydrogen production in photo-electrochemical cells (PECs) constitutes an important avenue towards solar energy conversion for the production of carbon-free fuel.[1] Upon illumination the oxygen evolution reaction (water oxidation reaction) occurs on the anode of the cell, which can be described by:

$$2\mathrm{H_2O} + 4h^+ = \mathrm{O_2} + 4\mathrm{H}^+ \tag{3.1.1}$$

with molecular hydrogen being produced on the cathode. The conversion efficiency of the PEC depends critically on the catalytic performance of the electrodes. Nanostructured devices based on titanium dioxide (TiO_2) are promising candidates for wide use in photon-induced reactions because of this material's photochemical stability, nontoxicity, and natural abudance.[2-6]

TiO_2 is a semiconductor metal oxide with an optical band gap of 3.2 eV in its bulk rutile form and band gap edges that straddle the water redox potentials; that is, the involved chemical reactions become thermodynamically accessible upon photon absorption. The catalytic activity of TiO_2 can be enhanced by co-catalysts or by metal dopants or defects due to ambient contamination during catalyst preparation.[7-11] The effect of such structural modifications on the thermochemistry at the surface/adsorbate interface has been discussed extensively in the literature,[3, 5, 12] but a comprehensive atomistic description of dynamic processes, such as the transport of the photo-generated charge carriers in the surface, is lacking even for pure TiO_2. Photon-induced water dissociation on the rutile TiO_2 (110) surface has recently been reported in the scanning tunneling microscope study of Tan et al.,[13] which demonstrated a very low rate (few events per hour) O–H bond-breaking in water upon irradiation with ultraviolet (UV) light under ultrahigh-vacuum conditions at 80 K. However, according to the density functional theory (DFT) calculations reported by Patel et al.,[14] the highest occupied molecular orbitals of molecular water are 1.40 eV below the valence band maximum (VBM) of the surface, which raises the question whether the reported dissociation of water is light-driven.

The measured properties of real materials depend on the preparation methods and conditions, which makes the identification of universal structure−property relationships a challenging task and hinders a systematic approach to catalyst optimization. Atomistic modeling and simulation of surface adsorbate interfaces can provide insights into the microscopic physicochemical processes that control catalysis,[12, 15-17] an inherently atomic-scale phenomenon, as well as interfacial charge-carrier transfer.[18-20] Previous theoretical studies of photon-mediated catalysis on rutile TiO_2 surfaces used ground-state calculations to investigate the effect of excited electronic states on reaction energetics and neglected the details of charge-carrier motion which

can strongly affect the overall reaction rate. This highlights the need for computational methods to investigate charge-carrier dynamics in materials accurately and realistically.[21-23] Here we use *ab initio* electronic structure calculations within DFT[24] and its time-dependent version (TDDFT)[25] to develop a framework based on the concept of localized and delocalized electronic states for the study of photocatalytic reactions on semiconductor surfaces.

Within this framework we evaluate the possibility of splitting of molecularly adsorbed water by photo-generated holes on rutile TiO_2 surfaces with (110) orientation, which provides insights at the femtosecond time scale (10^{-15} s). In addition, the time-domain description of the interaction of the hole with the water molecule could elucidate the reaction mechanism of other photochemical reactions where water is present. From our simulations we find that whether the first step of water dissociation is photon-induced or thermal depends sensitively on the local atomic environment and external parameters such as temperature. Specifically, oxidative dehydrogenation of molecular water by a photo-generated hole is possible in synergy with thermal effects only on surfaces with defects, such as subsurface interstitial Ti atoms which act as trap sites for the hole. Our explicit time-dependent modeling provides useful guidelines for the rational design of efficient light-driven catalysis through careful control of the active site on the surface.

3.2. Models and methods

TiO_2 in its rutile form is a tetragonal crystal in the crystallographic space group $P4_2/mnm$ and has two formula units in the unit cell. Its most stable low-index surface is the one with (110) orientation (the calculated surface energy for the relaxed surface is 15.6 meV/au^2).[26] We used a unit cell of two Ti and four O atoms to model the bulk structure of rutile TiO_2 and an extended, 4-trilayer thick slab containing 162 atoms to model the (110) surface (Figure 3-1). The

latter was constructed by a truncation of the bulk structure along the (110) crystallographic plane. The obtained unit cell was augmented by a vacuum region of 8 Å on each side of the slab in *z* direction.

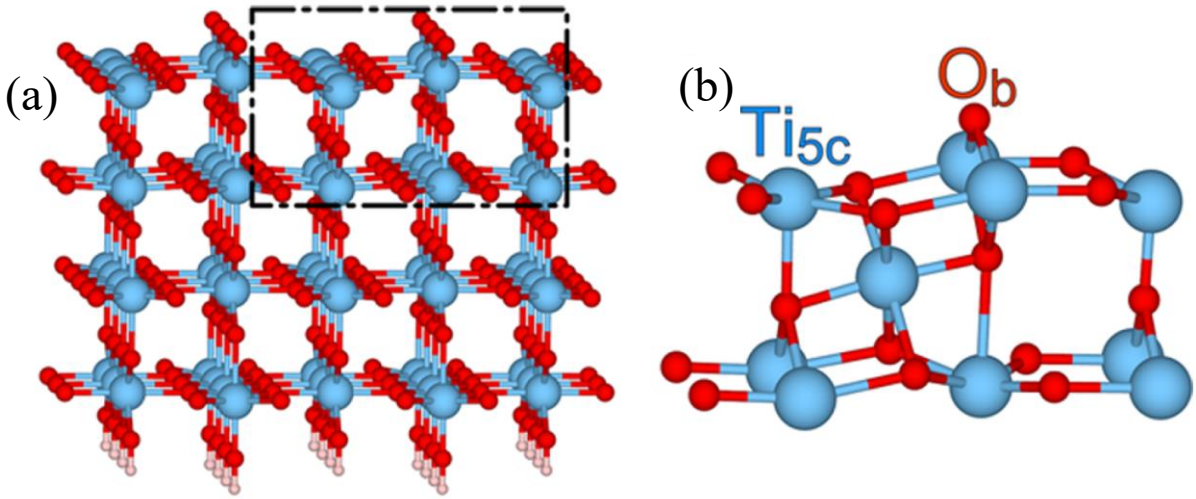

Figure 3-1. The structural model of the rutile TiO_2 surface with (110) orientation.

(a) Simulation cell of the stoichiometric surface with nearest O atoms in the surface plane. Blue, red, and white spheres represent Ti, O, and H atoms, respectively

(b) A portion of the simulation cell, identified by the dashed area in (a), with a Ti interstitial defect in the subsurface region. The surface 5-fold coordinated Ti atom (Ti_{5c}) and bridging O atom (O_b) are identified.

To model the catalyst surface, we used supercells with (3×2) periodicity relative to the primitive unit cell. One side of the slab was terminated with additional hydrogen atoms (bonded to the exposed O atoms) and hydroxyl (OH) ions (bonded to the exposed Ti atoms) in order to eliminate spurious localized states associated with the dangling bonds of O and Ti atoms, respectively, and emulate the bulk region of the catalyst. The positions of all atoms were optimized while keeping the lateral dimensions of the unit cell fixed to the corresponding bulk values and the terminating OH groups constrained in the direction perpendicular to the surface.

We performed total energy calculations using the GPAW code,[27] a grid-based implementation of the projector augmented-wave (PAW) method for all-electron calculations within the frozen core approximation.[28] The PAW setup for Ti (O) accounts for 12 (6) valence electrons and 10 (2) electrons in the frozen core (GPAW setup release 9672). A real-space grid with spacing of 0.18 and 0.24 Å was used for static and time-dependent calculations, respectively, and the reciprocal space was sampled only on the Γ-point because of the large lateral dimensions of the supercell. We used the Perdew – Burke – Ernzerhof (PBE) functional to describe exchange and correlation in the electronic system.[29] Because of the inadequate description of the strong Coulomb repulsion between semicore 3d-electrons localized on Ti, the PBE functional was augmented with the Hubbard-like term[30] $\sum_{\alpha}\left(U_{Ti(d)}/2\right)[Tr(\hat{\rho}^{\alpha}) - Tr([\hat{\rho}^{\alpha}]^2)]$, following the rotationally invariant formulation of Dudarev et al.,[31] where $\hat{\rho}^{\alpha}$ is the atomic orbital occupation matrix of the 3d-orbital of the Ti atom α and $U_{Ti(d)}$ is an effective parameter quantifying the screened Coulomb electron−electron interaction. This so-called DFT+U approach has been shown to describe trends in surface reactivity on TiO_2 with satisfactory accuracy compared to experiment[10, 30, 32] and with much less computational effort compared to approaches that use hybrid exchange and correlation functionals, which apply the same correction to the entire system regardless of where the error originates from. To obtain electronic configurations at the moment of vertical excitation, an electron is promoted to the bottom of the conduction band described by the wave function ψ_{N+1} and we use a generalization of the Δ-self-consistent field method (Δ-SCF) to allow the hole to occupy specifically designed linear combinations of the N occupied Kohn−Sham wave functions.[33] This maintains charge neutrality in the total system. The high dielectric constant of rutile TiO_2 (~100) contributes to the

screening of opposite charges from each other. The electron density, $n(\vec{r})$ at each self-consistency cycle is calculated as:[33]

$$n(\vec{r}) = \sum_{i=1}^{N} \psi_i^*(\vec{r})\psi_i(\vec{r}) - \sum_{i,j=1}^{N} a_i^* a_j \psi_i^*(\vec{r})\psi_i(\vec{r}) + \psi_{N+1}^*(\vec{r})\psi_{N+1}(\vec{r}) \quad (3.2.1)$$

The expansion coefficients, a_j, are optimized in a self-consistent manner so that the orbitals associated with the hole and electron – second and third terms in the equation (3.2.1) – resemble some input wave functions as much as possible. This approach has been used to study hole localization in rutile and anatase TiO_2, giving results in agreement with published photoluminescence measurements.[34] It has the additional benefit of constructing Kohn−Sham wave functions that are variationally optimized and not involving computationally expensive summations over many unoccupied electronic states, in contrast for example to calculations based on the linear-response TDDFT scheme or the solution of the Bethe–Salpeter equation.

We used the Ehrenfest approximation within the framework of real-time TDDFT to couple the electronic and ionic dynamics in the excited-state catalytic system,[25, 35-36] which is evolved from time t to time $t + \Delta t$ using the time evolution operators $\hat{T}_N$ and $\hat{T}_e$ for the nuclear and electronic parts, respectively:[35]

$$\hat{T}_{N,e}(t, t + \Delta t) = \hat{T}_N(t, t + \Delta t/2)\hat{T}_e(t, t + \Delta t)\hat{T}_N(t + \Delta t/2, t + \Delta t) \quad (3.2.2)$$

The electronic subsystem is treated quantum mechanically, propagated using the Crank–Nicholson semi-implicit scheme for $\hat{T}_e$ with a predictor-corrector time step of $\Delta t/2 = 5$ as (1 as $= 10^{-18}$ s).[37] The forces on the classical ions result from a mean-field average over the electronic states, and the ionic subsystem is evolved in time by application of the standard velocity Verlet propagator, $\hat{T}_N$. The occupation of all the wave functions in the system remains fixed during the simulation. Examples of previous successful applications of Ehrenfest dynamics include the

study of charge-carrier transport in dye-sensitized TiO_2 nanowires[38] and hydrogen bombardment of graphene-like nanoflakes,[35] and it is generally expected to work well in situations where the electronic states have similar character. In the present work we use TiO_2 as an illustrative example because of the availability of experimental data, but our computational strategy is general and should be useful for the study of other metal oxides.

In the rutile TiO_2 crystal each Ti (O) atom in the unit cell of the bulk structure has 6 (3) nearest neighbors. After structural optimization of the unit cell of the bulk structure, the calculated lattice constants are $a_0 = 4.65$ Å and $c_0 = 2.96$ Å, in good agreement with experiment (the measured lattice constants at 15 K are $a_0 = 4.586$ Å and $c_0 = 2.954$ Å)[39] and previous DFT calculations at the PBE approximation level ($a_0 = 4.65$ Å and $c_0 = 2.96$ Å).[30] For all calculations in the present work we use $U_{Ti(d)} = 4.2$ eV to improve the description of electronic screening in TiO_2, as proposed in previous studies using Dudarev's DFT+U approach within the PAW formalism.[21, 40] For $U_{Ti(d)} = 0$ the calculated lattice constants change by less than 1%. The outermost atomic layer in the stoichiometric (110) surface exhibits 5-fold coordinated Ti atoms (Figure 3-1(b), Ti_{5c}) and 2-fold coordinated O atoms, hereafter referred to as "bridging" O atoms (Figure 3-1(b), O_{br}).

In a photo-electrochemical cell, a photon of adequate energy excites an electron in the semiconductor catalytic material to the conduction band, leaving a hole in the valence band. After the photo-generated hole reaches the surface it is transferred to the water molecule, which is oxidized to a hydroxyl species, assuming that recombination of charge carriers does not occur first (these radiative relaxation processes are not considered in the present work). The lowest energy needed for the creation of an electron–hole pair is determined by the material's optical adsorption edge: according to calculations by Kang and Hybertsen based on the GW

approximation to the self-energy and the solution of the two-particle Bethe–Salpeter equation, the energy of the first dipole-allowed singlet exciton is 3.35 eV for bulk rutile TiO_2, almost equal to the quasiparticle band gap, 3.34 eV.[41] Within the framework of the density functional theory, the fundamental band gap of semiconductors and insulators is underestimated, but the DFT+U approach partially alleviates this problem.[12] We calculate the Kohn–Sham band gap of the bulk structure of rutile TiO_2 to be 1.69 eV for $U_{Ti(d)} = 0$ and 2.33 eV for $U_{Ti(d)} = 4.2$ eV. For the stoichiometric TiO_2 (110) surface, the band gap is calculated to be 1.72 and 2.38 eV for $U_{Ti(d)} = 0$ and $U_{Ti(d)} = 4.2$ eV, respectively. Because the conduction band has mainly titanium 3d-orbital character and the valence band oxygen 2p -orbital character, the band gap widens in an asymmetric fashion with increasing $U_{Ti(d)}$ as the CBM shifts to a higher energy.

A single water molecule preferably binds to the surface directly above a Ti_{5c} site; this is supported by experimental evidence from scanning tunneling microscopy studies under ultrahigh-vacuum conditions and theoretical work based on DFT investigations of the interaction between water and the rutile TiO_2 (110) surface.[13, 32] Upon oxidation of water, one of its H atoms is transferred to TiO_2 where it binds to a bridging O atom, leaving an OH group adsorbed at the Ti_{5c} site. To quantify the stability of water on the surface, we calculate the binding energy with respect to the clean surface and an isolated molecule. In this scheme, the lower the calculated value of the binding energy, the stronger the adsorption. For the coverage of 1/6 monolayer, adsorption is exothermic, and the calculated binding energy for molecular water is -0.96 eV, 0.22 eV higher than that for dissociated water with a binding energy of -1.18 eV. The lateral dimensions of the supercell ensure no interaction between water molecules adsorbed at neighboring supercells, as it would be relevant to the low-coverage limit. Using nudged elastic band calculations,[42] we find that the minimum-energy barrier for thermal dissociation is

relatively low, 0.15 eV (14 kJ/mol). The relative stability of the two modes of adsorption, molecular and dissociative, and the energy barrier depend on the computational method and the details of the structural model employed, for instance, the thickness of the slab, as discussed in the work of Kowalski et al.[43] For the structural model of the TiO_2 (110) surface employed in our calculations, going from 4 to 5 trilayer thick slabs changes the binding energy for adsorption by only 0.06 eV and the center of the water O *p*-states, calculated as the first moment of the projected density of states on the O atom, is shifted to lower energy by 0.06 eV. From this test, we do not expect the conclusions of the present work to be affected by computational details such as the thickness of the slab.

After ground-state structural optimization, the temperature of the catalytic system was raised from 0 K to a temperature relevant to experimental conditions.[13] We first assigned initial velocities to the ions of the optimized surface/adsorbate configuration according to the equilibrium Boltzmann–Maxwell distribution at 100 K, and we subsequently performed a molecular dynamics simulation for 300 fs in the microcanonical ensemble (N, V, E) using a time step of 1 fs. The positions of the atoms in the bottom trilayer were kept fixed during the simulation. We then used an equilibrated configuration from this simulation as the ground state of the surface/adsorbate interface at the moment of the vertical excitation. To model the excited-state surface/adsorbate interface, the electronic structure of the thermalized configuration was augmented by two wave functions: one to which we assign occupancy of -1 (hole) designed as discussed in the next section, and the other corresponds to the wave function associated with the CBM which is assigned occupancy of $+1$ (electron). The obtained electron–ion system corresponds to the excited state of the surface/adsorbate interface at the moment of the vertical excitation, which is defined as the zero of the time scale ($t = 0$). In this work, we concentrate on

the interaction of the hole with the adsorbate, and we do not consider exciton generation and diffusion or the separation of the charge carriers that precede hole transport to the surface. Instead, the excitation is represented by a weighted sum over Kohn–Sham wave functions in an electronic configuration that does not necessarily correspond to the initial excitation of the catalytic system by light which leads to exciton generation. The associated vertical excitation therefore may not contribute strongly to a true excitonic state of the surface/adsorbate interface, but investigation of the optical response of the material is beyond the scope of the current study.

Our computational strategy for the study of the dynamics of the interaction between the surface and the molecule under conditions of photoexcitation can be summarized as follows: (1) After ground-state structural optimization, *ab initio* molecular dynamics in the microcanonical ensemble is used to bring the catalytic system to a desired temperature. (2) An excitation is created by promoting an electron from a state which is an expansion of the occupied Kohn–Sham wave functions to the bottom of the conduction band of the surface. (3) The excited state of the surface/adsorbate interface is evolved in time using Ehrenfest dynamics within the framework of real-time propagation TDDFT. The evolution of the catalytic system is studied by monitoring the difference in the electron densities between the system in the excited state and the ground state of the same ionic configuration at each time instant.

3.3. Results

We used three different orbitals for the hole to investigate its motion in the surface, discussed in order of increasing relevance to standard experimental conditions. The first was modeled after a wavefunction entirely localized at the water molecule adsorbed on the stoichiometric TiO_2 (110) surface, chosen because the vertical excitation energy, 9.60 eV, was such that it ensured the dissociation of the water molecule. The main purpose of this excitation

was to evaluate the suitability of our methodology for the study of photon-induced dynamic processes at surface/adsorbate interfaces. The second hole wavefunction was localized at the surface, and the excitation energy was 3.93 eV, close to the band gap of the bulk material, which should correspond to realistic situations for irradiation with UV light. The third excitation energy was 1.00 eV in a configuration that included a hole wavefunction localized at a subsurface defect near the site of the water molecule absorption. The evolution of these three excited states in time is shown in Figure 3-2. We discuss them in detail in this order.

In the first excitation, the wave function associated with the hole resembles as much as possible the nonbonding orbital of an isolated water molecule. The difference in energy between the excited and ground state at $t = 0$ is 9.60 eV (equal to the vertical-excitation energy). This energy is more than 3 times higher than the band gap, and it corresponds to the extreme UV irradiation wavelength of 130 nm. This is because the molecular orbitals of water where the hole was created are energetically much lower than the TiO_2 band gap edges: the center of the water oxygen *p*-states is 4.72 eV lower than the VBM and 7.09 eV than the conduction band minimum (CBM). It is natural to expect that such high excitation energy can generate a deep hole with the oxidative power to initiate splitting of water. We calculate the difference between the electron density of the excited state and the electron density of the ground state of the same ionic configuration at each instant as a means to monitor the evolution of the excited state electron−ion system (see also supporting information). Figure 3-2(a) shows that in comparison to the ground state, charge is depleted mostly in the region around the water O atom (light green) immediately after the vertical excitation ($t = 0$). It also reveals an image charge in the surface that screens the excitation as well as some charge redistribution within the water molecule and between the molecule and the nearest bridging O atom of the surface. The excess charge introduced by the

electron wave function is mostly localized on Ti atoms in the subsurface region and away from the water molecule.

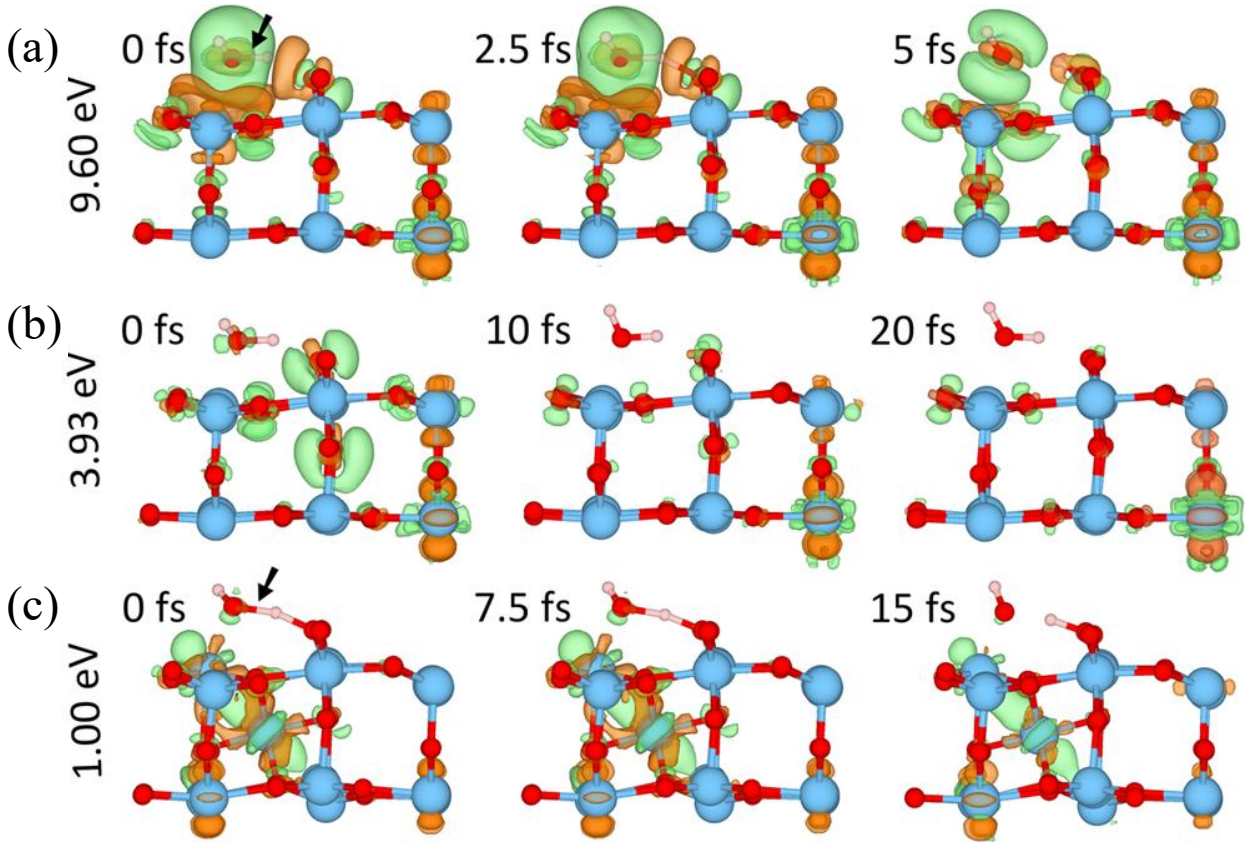

Figure 3-2. The excited-state dynamics of titania-water system.

Difference between the electron densities of the excited state and the ground state of the TiO_2(110) surface/water interface for the same ionic configuration at different time instants during simulation for three different vertical excitations: (a) 9.60 eV and (b) 3.93 eV in the stoichiometric surface and (c) 1.00 eV in a defected surface. The electron density differences are shown by isosurfaces with value 0.01 $e/\text{Å}^3$, where green and orange signify regions of electron charge depletion and charge accumulation, respectively. Black arrows point at the O−H bonds being broken.

After the high-energy vertical excitation the system is evolved in time using Ehrenfest dynamics. Similar to the situation at $t = 0$, at $t = 2.5$ fs the hole is almost entirely localized on the water molecule, although some charge has also been depleted from the O atom in the subsurface region nearest to the water molecule. At $t = 5$ fs the water molecule has dissociated, leaving one OH group bound to the Ti_{5c} atomic site (Figure 3-2(a)) and at $t = 10$ fs the hole has been transferred to the subsurface O atom and the charge distribution associated with the hole resembles the oxygen p_z-orbital (with two lobes aligned in the direction perpendicular to the surface). Part of the potential energy, about 2.7 eV, of the initial excitation is lost through energy transfer to the ionic motion. The distance between the O atom of the water molecule and the H atoms increases from 1.05 Å at $t = 0$ to 1.21 Å at $t = 2.5$ fs and to 1.62 Å at $t = 5$ fs. The distance between the H atom of the water molecule and the nearest bridging O atom decreases from 1.52 Å at $t = 0$ to 1.35 Å at $t = 2.5$ fs and to 0.94 Å at $t = 5$ fs. Inspection of Figure 3-2(a) also reveals that during the entire simulation time the electron is virtually immobile. The next vertical excitation we considered, 3.93 eV (see Figure 3-2(b)), involves electronic states only near the band gap edges of the surface, which is relevant to experimental UV conditions. In photo-induced catalysis the charge carrier needs to reach the surface/adsorbate interface and to be stable until transferred to the adsorbate.[5] In this respect, carrier trapping is an important issue in interfacial charge transfer, but the possible location of hole traps in the TiO_2 surface remains a subject of investigation, although bridging and subsurface O atoms have been suggested as likely atomic sites for small polarons.[5, 44-45] We also found these O atomic sites to mediate hole transport (Figure 3-2(a), $t = 5$ fs). In order to better understand the effect of spatial localization of charge on the surface/adsorbate interaction, we analyze the electron localization function (ELF) which has proven useful in studying bonding in extended and molecular systems in a

chemically intuitive way.[46-47]

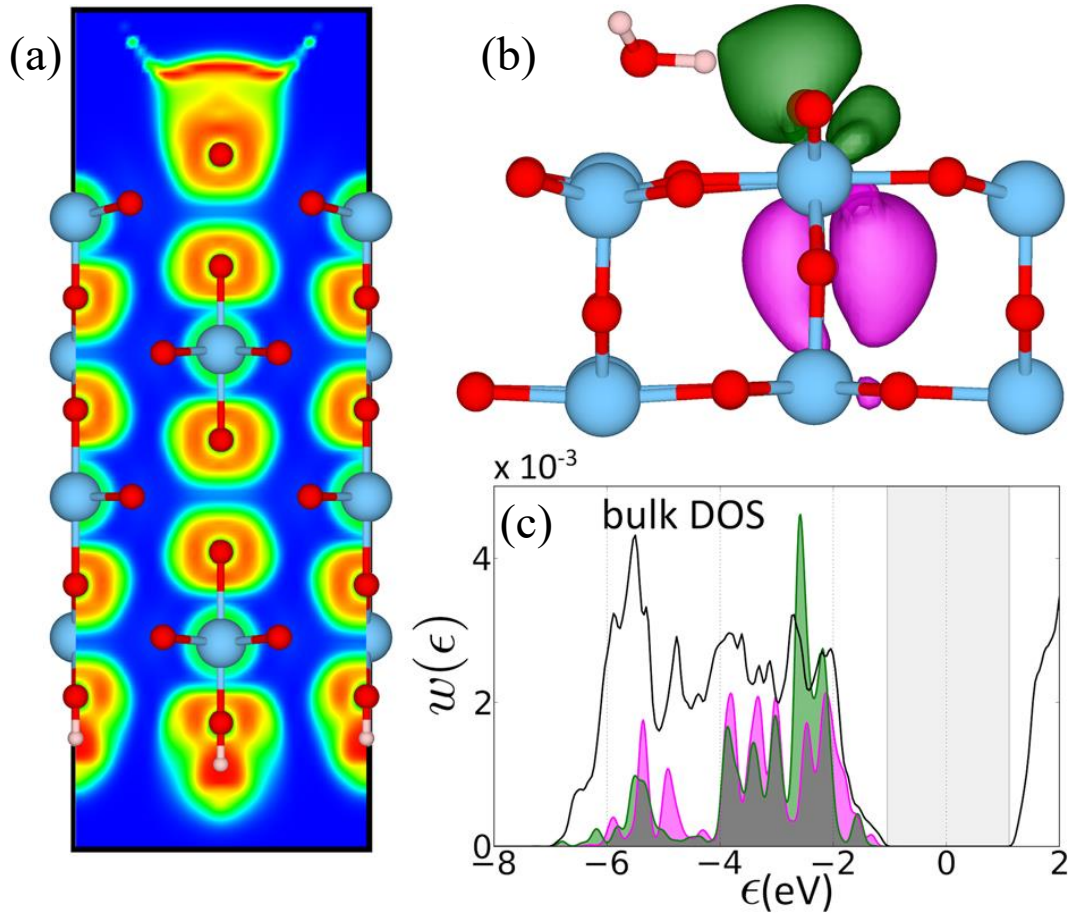

Figure 3-3. The electronic structure of titania–water system.

(a) Two-dimensional cross section of the electron localization function through the bridging oxygen (blue, green, and red for values of 0, 0.5, and 1).

(b) Isosurfaces representing the maximally localized Wannier functions spanning the range of energies closest to the valence band maximum for the bridging (green cloud) and subsurface (violet cloud) oxygen atoms nearest to the water molecule.

(c) Projections of the Wannier functions onto the Kohn−Sham states of the TiO_2 (110) surface (the regions where green and violet weights overlap appear gray). The total area of each shaded plot equals 1. The black line marks the density of states of the bulk, and the middle of its band gap is defined as the zero of the energy scale. The shaded vertical strip delimits the band gap.

The ELF is a positive quantity between 0 (blue regions in Figure 3-3(a)) and 1 (red regions in Figure 3-3(a)), used as a measure of finding an electron in the vicinity of another electron of the same spin (and therefore of the probability of opposite-spin pairs). Figure 3-3(a) portrays a two-dimensional cross section of the ELF through a bridging O atom. In the interior of the slab, the Ti−O has strong ionic character as evidenced by the high localization values, >0.8, near each O atom, as would be expected in the bulk structure of TiO_2. For the exterior of the slab, in contrast to the bottom surface that is terminated with −OH groups, the top surface also exhibits high ELF values above the bridging O atoms, the signature of a dangling bond. While ELF provides insights into the spatial distribution of chemical bonds, it does not reveal the shape of orbitals where the electrons are expected to exist. To this end, we analyze the maximally localized Wannier functions of the valence band manifold, an approach which allows the representation of extended Bloch states with atomic-like orbitals. We used the Wannier90[47] utility interfaced to the Quantum Espresso[48] code to produce the Wannier functions. The functions we obtained possess the expected character: in the bulk-like region of the slab, there is one Wannier function for each σ-bond between Ti and O atoms and in the subsurface one "lone pair" orbital on each of the O atoms (represented by the purple clouds in Figure 3-3(b)). On the surface, each bridging O atom has a dangling bond which mixes with the lone pair, resulting in two almost equivalent Wannier functions (one of which is shown by the green clouds in Figure 3-3(b)). Because the Wannier functions are linear combinations of the eigenfunctions of the Kohn–Sham Hamiltonian, they can be represented by the weights of the Kohn–Sham orbitals contributing to a given function. We find that the Wannier function corresponding to the dangling bonds of the bridging O atom is comprised mostly of the Kohn–Sham orbitals that span the range of energies near the top of the valence band (Figure 3-3 (c)). These Wannier functions

were used as a guide for modeling a hole wave function that is equally distributed over the bridging and subsurface O atoms nearest to the water molecule. In order to accomplish this spatial distribution of the hole wave function in Δ-SCF, we calculated the projections of the Wannier function on the atomic orbitals; the largest coefficients that correspond to the bridging oxygen *s*- and *p*-orbitals are: $s = 0.32,\ p_x = 0.10,\ p_y = 0.10,\ p_z = 0.05$.

Using the previously thermalized ionic configuration, we model the excitation at $t = 0$ by augmenting the electronic system with a wave function which is a linear combination of the PAW atomic partial waves with the coefficients presented above and which resembles as much as possible the calculated Wannier functions shown in Figure 3-3(b). As in the case of the high-energy excitation (vertical-excitation energy of 9.60 eV), the electron wave function corresponds to the state at the CBM. Upon excitation, charge is depleted from the region in the vicinity of the bridging and subsurface O atoms that are nearest to the water molecule, and there is also some redistribution of charge in the molecule and at the surface oxygen atoms (Figure 3-2(b), $t = 0$). The vertical excitation energy, 3.93 eV, corresponds to photon wavelength of 313 nm, which lies within the typical range for irradiation with UV light in experiment.[6] This low-energy excitation involves no bond-breaking events during the first 30 fs. We do not expect that propagation of the electron-ion system at longer times would lead to oxidation of water since already at 10 fs the hole has diffused to the interior of the slab. This is because the net positive charge at the surface/adsorbate interface gives rise to a gradient in the electrostatic potential that drives the hole away from the water molecule.[45] Irrespective of the role of ionic motion, hole transport is mediated by atomic oxygen sites through transitions involving electronic states with strong oxygen *p*-level contribution.

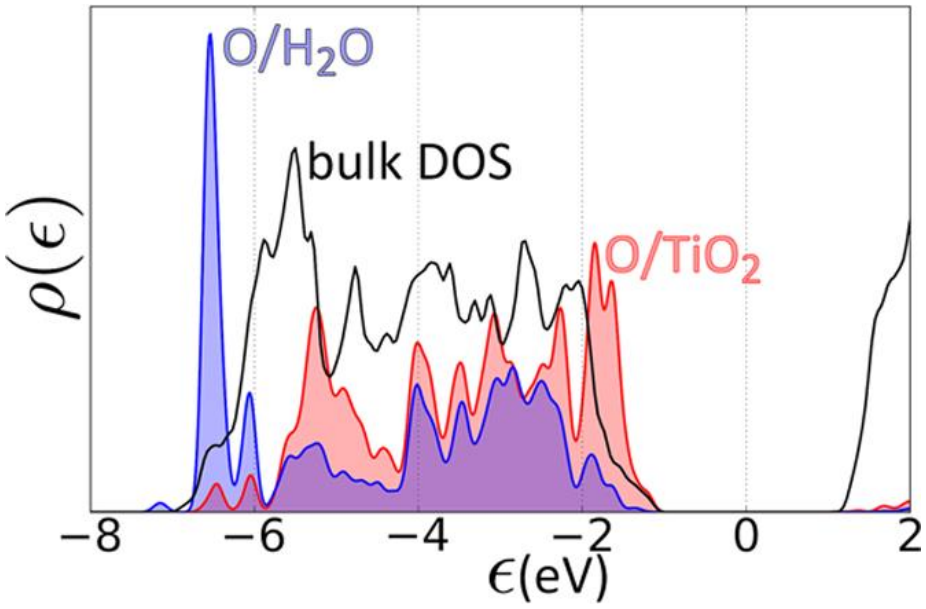

Figure 3-4. Density of states projected onto the *p* orbitals of the oxygen atoms in the top atomic trilayer of the TiO_2 (110) surface (red) and the oxygen atom of the water molecule (blue) at simulation time of 30 fs after low-energy excitation (3.93 eV). The black line marks the density of states of the bulk (scaled by a factor of 1/72), and the middle of its band gap is defined as the zero of the energy scale.

The strong coupling between these atomic sites prevents the spatial confinement of the hole near the water molecule, and the initially localized redistribution of the electron density at $t = 0$ spreads over the entire lattice supercell (which comprises several primitive unit cells) within a few femtoseconds. Most importantly, although there is considerable overlap between the oxide surface and molecular O states at lower energies in the -6 to -2 eV range relative to the VBM, there are no molecular states near the band gap edge where the hole could reside as the projected density of states plot in Figure 3-4 reveals.[14] For the calculation of the projected density of states we used the time-dependent wave functions, which although do not have a rigorous physical meaning, we used to obtain a qualitative description of the effect of the moving ions on the energy of molecular states. Modeling the orbital associated with the hole after a wave

function that is (a) entirely localized at a single bridging O atom, (b) equally distributed over all 6 bridging O atoms in the supercell, or (c) equally distributed over the water molecule (nonbonding orbital) and the nearest bridging O atom, results in delocalization of the hole within 20 fs.

Calculations with $U_{Ti(d)} = 2.0$ eV lead to the same outcome as for $U_{Ti(d)} = 4.2$ eV. We conclude from our excited-state Ehrenfest dynamics that photon-induced dehydrogenation of molecular water on pure TiO_2 (110) surfaces is not likely. The rutile phase of stoichiometric TiO_2 has for long been of great scientific interest as a prototypical metal oxide. This phase is mostly inert and under standard conditions its surface is non-stoichiometric with excess Ti atoms.[7] Surface and near-surface local imperfections, such as the common Ti interstitials and O vacancies, modify the surface chemistry of this material and can be beneficial to its efficiency as a catalyst. For example, Ti interstitials in the subsurface region can stabilize the bonding of chemical species as in the case of formadelhyde (CH_2O), discussed in the DFT study of Haubrich et al.[9] In what concerns the reactions of water on the rutile TiO_2 surface, increased interaction with the surface induced by the defects should decrease the energy barrier for thermal dissociation. In addition, defect states reduce the minimum energy for photoexcitation from the UV range to the visible range and can act as local trap sites for the photo-generated hole until it is transferred to the water molecule to promote the oxidation reaction.

Motivated by the above considerations, we studied the role of Ti interstitials in photon-induced water splitting on the rutile TiO_2 (110) surface. To model the defected catalytic surface, we introduced a single Ti atom in one of the ⟨110⟩ channels in the subsurface region of our structural model (see Figure 3-1(b)). This configuration is thermodynamically more stable (lower in energy) by 0.22 eV compared to a configuration with the Ti interstitial between two ⟨110⟩

channels. The calculated binding energy for molecular water adsorbed on the Ti_{5c} atomic site is -1.12 eV, less stable by 0.15 eV than dissociated water which has an adsorption energy of -1.27 eV, and more stable by 0.16 eV than adsorption on the stoichiometric surface (adsorption energy of -0.96 eV). Moreover, incorporation of the Ti interstitial in the TiO_2 lattice introduces electronic states in the band gap. The calculated energy of these defect-related states depends on the details of the computational method, for instance the value of the parameter U in the DFT+U approach.[30, 49] For $U_{Ti(d)} = 4.2$ eV, the defect-induced state is 1.00 eV below the CBM, in reasonable agreement with UV photoelectron spectroscopy on bulk reduced rutile TiO_2 which reveals defect-related electronic states at about 0.9 eV below the CBM.[50] Ti interstitials in the subsurface region also result in a lower energy barrier for thermal subtraction of hydrogen from water. During ground-state molecular dynamics in the microcanonical ensemble at 100 K, we find that an adsorbed water molecule spontaneously dissociates in the presence of a subsurface Ti interstitital, which suggests that the first step of water splitting on defect-containing TiO_2 (110) surfaces can be an entirely thermally driven process even at relatively low temperatures. On the other hand, an adsorbed water molecule remains intact at a temperature of 60 K. A resonance between the two limiting structures of molecular and dissociated water occurs at the intermediate temperature of 80 K, and although no thermal dissociation is observed, for 28% of the total simulation time the water O–H bond is elongated more than 10% of its equilibrium value, 1.02 Å, and is elongated by more than 20% of its equilibrium value for 12% of the total simulation time. This extension of the water O–H bond allows the H atom to be shared between the water and bridging O atoms. For comparison, in a simulation at 60 K this bond is never stretched by more than 15%.

To investigate the effect of photo-excitation on water dissociation in the defect-containing surface at 80 K, we considered the situation where an electron is promoted at $t = 0$ from a defect-related state in the band gap to the CBM (shown in Figure 3-2(c)). This corresponds to an excited-state configuration within a polaronic description of a hole trapped by a defect, following irradiation with light of energy equal to or higher than the optical absorption edge of the material. At $t = 0$, the distance between the water O and H atoms is 1.27 Å and the distance between the water H and bridging O atoms is 1.24 Å. Figure 3-2(c) shows that the hole is localized in the subsurface region at the Ti interstitial and partly overlaps the extended form of molecular water. The pinning of the hole near the surface weakens the restorative force on the H atom in the stretched water O–H bond, which leaves the H atom with enough kinetic energy to overcome the energy barrier for approaching the bridging O atom and binding to the surface within 15 fs.

Ehrenfest dynamics is a semi-classical method that cannot capture quantum mechanical processes in the ionic subsystem. Upon including nuclear quantum effects a water H atom could hop onto the nearby bridging O atom at small separations (less than ~0.2 Å) between the water H and surface bridging O atoms. However, a more complete quantum-mechanical treatment, for example using surface hopping techniques,[51-52] remains currently impractical for the study of our structural models because of size and time scale considerations.

3.4. Discussion and conclusions

Our results for the interaction of low concentrations of molecular water with stoichiometric and defected surfaces of rutile TiO_2 with (110) orientation have important implications for the photocatalytic activity of the material. We can draw as a first conclusion that whether water photo-oxidation is possible depends on the local atomic environment and

experimental conditions. On the idealized stoichiometric surface a hole relaxed at the top of the TiO_2 valence band cannot oxidize molecularly adsorbed water as the highest occupied molecular orbitals remain lower in energy than the VBM of the surface by −2 eV (Figure 3-4). Moreover, according to our molecular dynamics simulations, thermal dissociation does not occur at temperatures at least as high as 300 K. Increasing the coverage of water on the surface from 1/6 to 1/3 of a monolayer with two water molecules adsorbed on adjacent Ti_{5c} atomic sites has little effect on the electronic structure of the surface/adsorbate interface due to the weak interaction between the molecules (adsorption is stabilized by 2%). However, our model assumes ultrahigh vacuum, and the effects of complex environments such as electrolytes on the dynamics of the surface/adsorbate interface merit further investigation.[53]

As a second conclusion, we find that the water photo-oxidation reaction could occur on surfaces containing Ti interstitials at temperatures between 60 and 100 K. In the work of Tan et al.,[13] which reported water photo-oxidation, the pertinent scanning tunneling measurements were performed at a temperature of 80 K, a critical factor for photocatalytic dissociation according to our findings, in the presence of Ti interstitial defects (a feature not discussed in that work). Because the Ti interstitial participates in the reaction mechanism by modifying the potential energy landscape and by pinning the hole near the surface, the catalytic active site comprises both the Ti_{5c} and Ti-interstitial atomic sites. We calculated that the hopping of a Ti interstitial between two adjacent equivalent equilibrium sites requires overcoming an energy barrier of 0.52 eV (50 kJ/mol). From this fact, we expect that diffusion of Ti interstitials will be of secondary importance regarding the reaction mechanism of water oxidation at temperatures of practical interest. Care must be taken for the correct interpretation of our findings, which demonstrate the importance of thermal effects in water dissociation but the estimated short time for O–H bond

breaking, 15 fs, is not necessarily in quantitative correspondence with experimentally measured high rate. In contrast, as a number of necessary conditions must be met for the reaction to occur, the dissociation probability should be low, in support of the observation that rutile TiO_2 (110) surfaces are inefficient as catalysts for the water photo-oxidation reaction under ultrahigh-vacuum conditions.[54] In addition, subsurface defects can act as recombination centers with adverse effects on the dissociation probability, but recombination of charge carriers is not considered in this work. In the same vein, the temperature window which favors a resonant state of adsorption will depend on the structural details of the surface, although the estimated temperature of 80 K is in agreement with experiment.[13]

The interstitial-containing surface studied here can be thought of as a structural model of catalytic surfaces where a hole is trapped at the active site facilitating the reaction. This result highlights the importance of defects in photo-catalysis and suggests that careful control of the atomic structure and composition of the surface can enhance the reaction rate. For example, structural defects such as steps or kinks fit a generalized notion of a defect-containing active center,[54-56] but further work is necessary in order to explore how morphologies other than extended flat surfaces mediate photon-induced catalysis.[57]

Our methodology should also be useful for the study of adsorbates other than water. Although the alignment between the electronic levels of the TiO_2 and molecular water is not favorable for the oxidation reaction to occur on the pure and flat (110) surface, the outcome of illumination can be different for species that chemisorb on the surface, such as methoxy (OCH_3). The methoxy radical is a possible reaction intermediate of another important reaction – the methanol oxidation reaction.[58-59] In our initial investigation of this system, the projected density of states calculation reveals that there is a significant population of methoxy O states at the VBM

of TiO_2, and therefore methoxy should act as a hole scavenger enabling interfacial hole transfer. A more detailed investigation and analysis are the subject of the work described in Chapter 3.

In conclusion, we developed a simple methodology for the modeling and simulation of dynamic processes at excited-state interfaces which we used to study charge-carrier motion and interaction with molecularly adsorbed water on rutile TiO_2 (110) surfaces. We found that water photo-oxidation to hydroxyl is possible in synergy with thermal effects. On the stoichiometric surface there are no electronic states associated with water near the VBM of the TiO_2 surface, and therefore oxidation by a photo-generated hole is not possible for excitation energies corresponding to the band gap of TiO_2. By contrast, Ti interstitials in the subsurface region increase the binding strength of water on the surface, lower the minimum energy barrier for thermal dissociation of water, and act as trap sites for holes until they are transferred to water and oxidize it. Moreover, the reaction pathway critically depends on external conditions such as temperature. On the defected surface, low temperatures (<60 K, according to our molecular dynamics simulation) favor molecularly adsorbed water and higher temperatures (>100 K) favor dissociated water. Intermediate temperatures, ~80 K, favor a resonant state of adsorption that enables splitting of the molecule. Under these circumstances thermal and photon-induced effects act in a complementary manner that leave an ambiguity in the causal sequence: water dissociation can be equally considered a photon-induced and thermal process or a thermal and photo-assisted process. Although optimization of catalytic performance is ultimately a combinatorial challenge, this work paves the way for the investigation of the kinetics of photo-electrochemical reactions on semiconductor surfaces and provides useful guidelines for the rational design of efficient photocatalytic materials by careful control of the active sites on the surface.

3.5. References

1. Walter, M. G.; Warren, E. L.; McKone, J. R.; Boettcher, S. W.; Mi, Q. X.; Santori, E. A.; Lewis, N. S., Solar Water Splitting Cells. *Chemical Reviews* **2011**, *111*, 5815-5815.

2. Fujishima, A.; Honda, K., Electrochemical Photolysis of Water at a Semiconductor Electrode. *Nature* **1972**, *238*, 37-38.

3. Fujishima, A.; Zhang, X.; Tryk, D. A., TiO_2 Photocatalysis and Related Surface Phenomena. *Surface Science Reports* **2008**, *63*, 515-582.

4. Ni, M.; Leung, M. K. H.; Leung, D. Y. C.; Sumathy, K., A Review and Recent Developments in Photocatalytic Water-Splitting Using TiO_2 for Hydrogen Production. *Renew Sust Energ Rev* **2007**, *11*, 401-425.

5. Henderson, M. A., A Surface Science Perspective on TiO_2 Photocatalysis. *Surface Science Reports* **2011**, *66*, 185-297.

6. Henderson, M. A.; Lyubinetsky, I., Molecular-Level Insights into Photocatalysis from Scanning Probe Microscopy Studies on TiO_2 (110). *Chemical reviews* **2013**, *113*, 4428-4455.

7. Diebold, U., The Surface Science of Titanium Dioxide. *Surface Science Reports* **2003**, *48*, 53-229.

8. Thompson, T. L.; Yates, J. T., Surface Science Studies of the Photoactivation of TiO_2-New Photochemical Processes. *Chemical Reviews* **2006**, *106*, 4428-4453.

9. Haubrich, J.; Kaxiras, E.; Friend, C. M., The Role of Surface and Subsurface Point Defects for Chemical Model Studies on TiO_2: A First-Principles Theoretical Study of Formaldehyde Bonding on Rutile TiO_2 (110). *Chemistry-A European Journal* **2011**, *17*, 4496-4506.

10. Garcia-Mota, M.; Vojvodic, A.; Abild-Pedersen, F.; Norskov, J. K., Electronic Origin of the Surface Reactivity of Transition-Metal-Doped TiO_2 (110). *Journal of Physical Chemistry C* **2013**, *117*, 460-465.

11. Su, R., et al., Designer Titania-Supported Au-Pd Nanoparticles for Efficient Photocatalytic Hydrogen Production. *ACS Nano* **2014**, *8*, 3490-3497.

12. Liao, P.; Carter, E. A., New Concepts and Modeling Strategies to Design and Evaluate Photo-Electro-Catalysts Based on Transition Metal Oxides. *Chem Soc Rev* **2013**, *42*, 2401-2422.

13. Tan, S. J.; Feng, H.; Ji, Y. F.; Wang, Y.; Zhao, J.; Zhao, A. D.; Wang, B.; Luo, Y.; Yang, J. L.; Hou, J. G., Observation of Photocatalytic Dissociation of Water on Terminal Ti Sites of TiO_2 (110)-1×1 Surface. *Journal of the American Chemical Society* **2012**, *134*, 9978-9985.

14. Patel, M.; Mallia, G.; Liborio, L.; Harrison, N. M., Water Adsorption on Rutile TiO_2 (110) for Applications in Solar Hydrogen Production: A Systematic Hybrid-Exchange Density Functional Study. *Physical Review B* **2012**, *86*, 045302.

15. Tritsaris, G. A.; Greeley, J.; Rossmeisl, J.; Norskov, J. K., Atomic-Scale Modeling of Particle Size Effects for the Oxygen Reduction Reaction on Pt. *Catal Lett* **2011**, *141*, 909-913.

16. Tritsaris, G. A.; Rossmeisl, J., Methanol Oxidation on Model Elemental and Bimetallic Transition Metal Surfaces. *Journal of Physical Chemistry C* **2012**, *116*, 11980-11986.

17. Norskov, J. K.; Bligaard, T.; Rossmeisl, J.; Christensen, C. H., Towards the Computational Design of Solid Catalysts. *Nat Chem* **2009**, *1*, 37-46.

18. Meng, S.; Kaxiras, E., Real-Time, Local Basis-Set Implementation of Time-Dependent Density Functional Theory for Excited State Dynamics Simulations. *Journal of Chemical Physics* **2008**, *129*, 054110.

19. Duncan, W. R.; Prezhdo, O. V., Theoretical Studies of Photoinduced Electron Transfer in Dye-Sensitized TiO_2. *Annu Rev Phys Chem* **2007**, *58*, 143-184.

20. Onda, K.; Li, B.; Zhao, J.; Jordan, K. D.; Yang, J. L.; Petek, H., Wet Electrons at the H_2O/TiO_2 (110) Surface. *Science* **2005**, *308*, 1154-1158.

21. Ji, Y. F.; Wang, B.; Luo, Y., Location of Trapped Hole on Rutile- TiO_2 (110) Surface and Its Role in Water Oxidation. *Journal of Physical Chemistry C* **2012**, *116*, 7863-7866.

22. Valdes, A.; Kroes, G. J., Cluster Study of the Photo-Oxidation of Water on Rutile Titanium Dioxide (TiO_2). *Journal of Physical Chemistry C* **2010**, *114*, 1701-1708.

23. Bernardi, M.; Vigil-Fowler, D.; Lischner, J.; Neaton, J. B.; Louie, S. G., Ab Initio Study of Hot Carriers in the First Picosecond after Sunlight Absorption in Silicon. *Physical Review Letters* **2014**, *112*, 257402.

24. Hohenberg, P.; Kohn, W., Inhomogeneous Electron Gas. *Physical Review B* **1964**, *136*, B864-B871.

25. Runge, E.; Gross, E. K. U., Density-Functional Theory for Time-Dependent Systems. *Physical Review Letters* **1984**, *52*, 997-1000.

26. Ramamoorthy, M.; Vanderbilt, D.; Kingsmith, R. D., 1st-Principles Calculations of the Energetics of Stoichiometric TiO_2 Surfaces. *Physical Review B* **1994**, *49*, 16721-16727.

27. Enkovaara, J., et al., Electronic Structure Calculations with GPAW: A Real-Space Implementation of the Projector Augmented-Wave Method. *J. Phys.-Condes. Matter* **2010**, *22*, 253202.

28. Blochl, P. E., Projector Augmented-Wave Method. *Physical Review B* **1994**, *50*, 17953-17979.

29. Perdew, J. P.; Burke, K.; Ernzerhof, M., Generalized Gradient Approximation Made Simple. *Physical Review Letters* **1996**, *77*, 3865-3868.

30. Stausholm-Moller, J.; Kristoffersen, H. H.; Hinnemann, B.; Madsen, G. K. H.; Hammer, B., DFT+*U* Study of Defects in Bulk Rutile TiO_2. *Journal of Chemical Physics* **2010**, *133*.

31. Dudarev, S. L.; Botton, G. A.; Savrasov, S. Y.; Humphreys, C. J.; Sutton, A. P., Electron-Energy-Loss Spectra and the Structural Stability of Nickel Oxide: An LSDA+*U* Study. *Physical Review B* **1998**, *57*, 1505-1509.

32. Hammer, B.; Wendt, S.; Besenbacher, F., Water Adsorption on TiO_2. *Topics in Catalysis* **2010**, *53*, 423-430.

33. Gavnholt, J.; Olsen, T.; Engelund, M.; Schiotz, J., Delta Self-Consistent Field Method to Obtain Potential Energy Surfaces of Excited Molecules on Surfaces. *Physical Review B* **2008**, *78*, 075441.

34. Zawadzki, P.; Jacobsen, K. W.; Rossmeisl, J., Electronic Hole Localization in Rutile and Anatase TiO_2 – Self-Interaction Correction in Delta-SCF DFT. *Chemical Physics Letters* **2011**, *506*, 42-45.

35. Ojanpera, A.; Havu, V.; Lehtovaara, L.; Puska, M., Nonadiabatic Ehrenfest Molecular Dynamics within the Projector Augmented-Wave Method. *Journal of Chemical Physics* **2012**, *136*.

36. Qian, X. F.; Li, J.; Lin, X.; Yip, S., Time-Dependent Density Functional Theory with Ultrasoft Pseudopotentials: Real-Time Electron Propagation across a Molecular Junction. *Physical Review B* **2006**, *73*, 035408.

37. Castro, A.; Marques, M. A. L.; Rubio, A., Propagators for the Time-Dependent Kohn-Sham Equations. *Journal of Chemical Physics* **2004**, *121*, 3425-3433.

38. Meng, S.; Ren, J.; Kaxiras, E., Natural Dyes Adsorbed on TiO_2 Nanowire for Photovoltaic Applications: Enhanced Light Absorption and Ultrafast Electron Injection. *Nano Letters* **2008**, *8*, 3266-3272.

39. Burdett, J. K.; Hughbanks, T.; Miller, G. J.; Richardson, J. W.; Smith, J. V., Structural Electronic Relationships in Inorganic Solids - Powder Neutron-Diffraction Studies of the Rutile and Anatase Polymorphs of Titanium-Dioxide at 15 and 295-K. *Journal of the American Chemical Society* **1987**, *109*, 3639-3646.

40. Morgan, B. J.; Watson, G. W., Intrinsic *n*-Type Defect Formation in TiO_2: A Comparison of Rutile and Anatase from GGA+*U* Calculations. *Journal of Physical Chemistry C* **2010**, *114*, 2321-2328.

41. Kang, W.; Hybertsen, M. S., Quasiparticle and Optical Properties of Rutile and Anatase TiO_2. *Physical Review B* **2010**, *82*, 085203.

42. Henkelman, G.; Uberuaga, B. P.; Jonsson, H., A Climbing Image Nudged Elastic Band Method for Finding Saddle Points and Minimum Energy Paths. *Journal of Chemical Physics* **2000**, *113*, 9901-9904.

43. Kowalski, P. M.; Meyer, B.; Marx, D., Composition, Structure, and Stability of the Rutile TiO_2 (110) Surface: Oxygen Depletion, Hydroxylation, Hydrogen Migration, and Water Adsorption. *Physical Review B* **2009**, *79*, 115410.

44. Zawadzki, P.; Laursen, A. B.; Jacobsen, K. W.; Dahl, S.; Rossmeisl, J., Oxidative Trends of TiO_2-Hole Trapping at Anatase and Rutile Surfaces. *Energy & Environmental Science* **2012**, *5*, 9866-9869.

45. Kerisit, S.; Deskins, N. A.; Rosso, K. M.; Dupuis, M., A Shell Model for Atomistic Simulation of Charge Transfer in Titania. *Journal of Physical Chemistry C* **2008**, *112*, 7678-7688.

46. Savin, A.; Nesper, R.; Wengert, S.; Fassler, T. F., ELF: The Electron Localization Function. *Angew. Chem.-Int. Edit.* **1997**, *36*, 1809-1832.

47. Marzari, N.; Mostofi, A. A.; Yates, J. R.; Souza, I.; Vanderbilt, D., Maximally Localized Wannier Functions: Theory and Applications. *Reviews of Modern Physics* **2012**, *84*, 1419-1475.

48. Giannozzi, P., et al., Quantum Espresso: A Modular and Open-Source Software Project for Quantum Simulations of Materials. *J. Phys.-Condes. Matter* **2009**, *21*, 395502.

49. Calzado, C. J.; Hernandez, N. C.; Sanz, J. F., Effect of on-Site Coulomb Repulsion Term *U* on the Band-Gap States of the Reduced Rutile (110) TiO_2 Surface. *Physical Review B* **2008**, *77*, 045118.

50. Wendt, S., et al., The Role of Interstitial Sites in the Ti 3*d* Defect State in the Band Gap of Titania. *Science* **2008**, *320*, 1755-1759.

51. Tully, J. C., Molecular-Dynamics with Electronic-Transitions. *Journal of Chemical Physics* **1990**, *93*, 1061-1071.

52. Tapavicza, E.; Tavernelli, I.; Rothlisberger, U., Trajectory Surface Hopping within Linear Response Time-Dependent Density-Functional Theory. *Physical Review Letters* **2007**, *98*, 023001.

53. Liu, L. M.; Zhang, C. J.; Thornton, G.; Michaelides, A., Structure and Dynamics of Liquid Water on Rutile TiO_2 (110). *Physical Review B* **2010**, *82*, 161415.

54. Henderson, M. A., Structural Sensitivity in the Dissociation of Water on TiO_2 Single-Crystal Surfaces. *Langmuir* **1996**, *12*, 5093-5098.

55. Shevlin, S. A.; Woodley, S. M., Electronic and Optical Properties of Doped and Undoped TiO_2 Nanoparticles. *Journal of Physical Chemistry C* **2010**, *114*, 17333-17343.

56. Martinez, U.; Vilhelmsen, L. B.; Kristoffersen, H. H.; Stausholm-Moller, J.; Hammer, B., Steps on Rutile TiO_2 (110): Active Sites for Water and Methanol Dissociation. *Physical Review B* **2011**, *84*, 205434.

57. Gong, X. Q.; Selloni, A.; Batzill, M.; Diebold, U., Steps on Anatase TiO_2 (101). *Nat. Mater.* **2006**, *5*, 665-670.

58. Shen, M.; Henderson, M. A., Identification of the Active Species in Photochemical Hole Scavenging Reactions of Methanol on TiO_2. *Journal of Physical Chemistry Letters* **2011**, *2*, 2707-2710.

59. Phillips, K. R.; Jensen, S. C.; Baron, M.; Li, S.-C.; Friend, C. M., Sequential Photo-Oxidation of Methanol to Methyl Formate on TiO_2 (110). *Journal of the American Chemical Society* **2013**, *135*, 574-577.

CHAPTER 4. PHOTO-INDUCED C–H ACIDITY REACTION MECHANISM

We conclude Part I by demonstrating another example application of the Ehrenfest dynamics method to the case of methoxy photo-oxidation on rutile TiO_2 (110) surface. We compare the ground-state and excited-state pathways of the reaction, demonstrate a way to translate the results of the computational simulations into Lewis structures, and establish a new reaction mechanism – "photo-induced C–H acidity". This chapter is based on the work described in: G. Kolesov, D. Vinichenko, G. A. Tritsaris, C. M. Friend, E. Kaxiras. Anatomy of the Photochemical Reaction: Excited-State Dynamics Reveals the C–H Acidity Mechanism of Methoxy Photo-oxidation on Titania. Journal of Physical Chemistry Letters, V. 6, pp. 1624 – 1627, to which Dmitry Vinichenko and Grigory Kolesov contributed equally.

4.1. Introduction

There is a great interest in the design of materials that can efficiently utilize solar energy to drive important chemical processes, including the synthesis of chemicals and fuels and the degradation of environmental pollutants. Most often, the choice of a photo-catalyst is guided by thermodynamic considerations, that is, a comparison of the energy of charge carriers (electrons and holes) created by photo-excitation to the energetic requirements for a specific reaction.[1-5] Although this is a necessary condition, it is not sufficient because the evolution of a system on the excited-state potential energy surface can have completely different features and driving forces compared to the ground-state. Thus, understanding heterogeneous photochemistry requires detailed insight into interactions of the key molecular intermediates with the semiconductor photo-catalyst, and their time evolution after excitation to produce the desired

reaction. To achieve this understanding, surface-bound intermediates and the reaction path to products must be identified. Such fundamental insights can provide a generalizable model for guiding the design of photo-catalysts for specific reactions.

4.2. Rutile titania as a model photo-catalyst

Titanium dioxide is among the most widely studied materials among metal oxides, with numerous proposed applications, from catalyst for water photo-oxidation to optical coating catalyst for environmental cleanup, and corrosion-protective coating.[1, 6-7] Rutile is the most thermodynamically stable phase of titanium dioxide. It has tetragonal crystal structure (space group $P4_2/mnm$) with distorted octahedron coordination of titanium atoms and trigonal coordination of oxygen atoms; the system does not have any localized magnetic moments.[6, 8-9] In a localized-bonds description which can be obtained by carrying out the Wannier function analysis,[10] each oxygen atom has 3 σ-bonds with neighboring titanium atoms and an out-of-plane p-type lone pair; each titanium atom has 6 σ-bonds to oxygen atoms in its coordination sphere. Due to an electronegativity difference between oxygen and titanium atoms, the bonds are polarized towards oxygen, and the structure has large ionic character. Therefore, the valence band is mostly spanned by oxygen p states, and the conduction band is spanned by titanium d states. The top of the valence band has large contributions from oxygen lone pairs. On the (110) surface the main features are the rows of five-fold coordinated titanium atoms, $\mathrm{Ti_{5c}}$, and the rows of two-fold coordinated "bridging" oxygen atoms, $\mathrm{O_b}$.[6, 11] The former typically act as adsorption sites for Lewis or Bronsted base adsorbates, such as water and alcohols, whereas the latter serve as Bronsted bases themselves and are prone to abstracting hydrogen atom with formation of the "bridging" hydroxyl species, $\mathrm{OH_b}$; the associated extra electron can get trapped on the lattice distortion in the subsurface region, forming a polaron.[12] Another prominent feature is the

possibility of bridging oxygen abstraction with formation of a vacancy, $V_{O_{br}}$. Such vacancies give rise to polarons in the subsurface region of the material,[13-17] often act as chemisorption sites,[18-21] and alter the catalytic properties of the material.[22]

4.3. Thermal chemistry of methanol on the rutile titania (110) surface

The preferred adsorption pathway for methanol is binding to an oxygen vacancy according to $CH_3OH + V_O + O_b = CH_3O_b + OH_b$. The resulting products are fairly stable – the temperature-programmed desorption peak corresponding to their recombinative decomposition is at 480 K. Adsorption on the undercoordinated Ti rows results in a mixture of molecular adsorption with the desorption peak at 295 K and dissociative adsorption with the cleavage of O–H bond and transfer of hydrogen atoms to neighboring bridging oxygen. Such dissociation of methanol to methoxy and bridging hydrogen has been studied computationally and is characterized by small positive enthalpy (~0.2 eV) and a reaction barrier of about 0.6 eV.[23-24] A more reliable way to prepare dissociated methanol is to put it on the rutile TiO_2 (110) surface with oxygen adatoms at 300 K – this procedure results in a pure layer of methoxy species residing on the 5-fold coordinated titanium rows.[25] The methoxy intermediate is stable thermally, up to 600 K, when thermal C–H bond cleavage occurs, yielding formaldehyde and methanol.[26-27]

From elementary considerations,[28] the electronic structure of the methoxy oxygen can be described as two σ-bonds to carbon and surface titanium, and two lone pairs, one of which lies in the C–O–Ti plane (n_σ), and the other perpendicular to that plane (n_p). Methoxy oxygen fills the coordination sphere of the titanium atom on which it is adsorbed, and thus is analogous to a slab oxygen with lone pair pointing in the [-110] direction. This leads to methoxy states contributing

to the density of states in the region near valence band maximum (Figure 4-1), as observed in earlier computational work.[29-30]

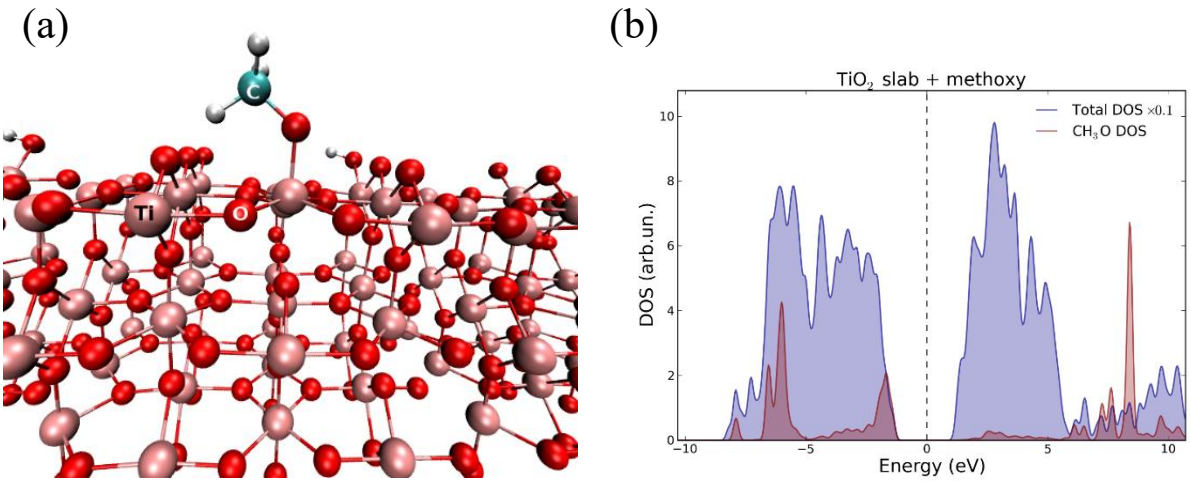

Figure 4-1. Rutile TiO_2 (110) slab and its density of states (DOS).

(a) Geometry of the rutile TiO_2 (110) slab with adsorbed methoxy. Only the top three layers of titanium atoms in the slab are shown; O atoms are dark red, Ti – pink, C – cyan, H – white.

(b) Total DOS and adsorbate-projected DOS. The total DOS is scaled down by a factor of 0.1 and shown in blue. The methoxy DOS is shown in red color. Hybridization of methoxy states with the top of titania valence band can be seen.

In order to get additional insight into the thermal chemistry of C–H bond cleavage, we use first-principles simulation to infer the transition state of the reaction. We use the climbing image nudged elastic band (CI–NEB) method[31-32] as implemented in VASP[33] software package. For the simulation of rutile TiO_2 (110) surface we have used a slab shown in Figure 4-1, which we describe in more details below. The lattice parameters used correspond to optimized bulk lattice constant value, $a_{bulk} = 4.550$ Å, $c_{bulk} = 2.974$ Å. For the (110) surface slab the first lateral dimension, along the Ti_{5c} rows, corresponds to the c direction of bulk material; we include

three Ti_{5c} in a row in our simulation, resulting in the slab size in x direction $a_{slab} = 3c_{bulk} = 8.922$ Å. The second lateral dimension is the [110] direction in the bulk material; the size of the slab along that dimension determines how many rows of undercoordinated Ti atoms are included in the simulation. We use a supercell with two rows of Ti_{5c}, corresponding to the slab size in y direction $b_{slab} = 2\sqrt{2}a_{bulk} = 12.869$ Å. The thickness of the slab in the z direction is determined by the number of layers of Ti atoms in the simulation cell; we have used 4-layer slab. The dangling bonds on the bottom of the slab were saturated with hydrogen atoms and hydroxyl groups. To enforce bulk-like geometric behavior, 48 atoms on the bottom of the slab were frozen. The Brillouin zone was sampled by the Γ point in all simulations. Since the slab contains 6 adsorption sites, having one methoxy particle corresponds to the coverage by adsorbed methoxy molecules of 0.18 monolayer (ML), commensurate with the experimental results.[25]

In modeling the electronic structure of TiO_2, it is important to go beyond standard DFT methods in order to correctly describe the behavior of localized titanium d-shell electrons.[12, 34-37] It has been shown that without DFT+*U* or other methods which treat electron interactions more accurately, the polaronic structures in TiO_2 cannot be resolved.[12, 34] Here we employed DFT+*U* method[38] that adds a correction for the on-site Coulomb interaction and allows to treat localized d-electrons in a computationally efficient way. We applied DFT+*U* correction to the *d*-orbitals of titanium; our benchmarking simulations yielded a value of $U = 4.2$ eV which is in agreement with that found in other works.[12] We used the PBE[39] exchange-correlation functional and standard PAW datasets of VASP program.

We find the reaction enthalpy is 0.9 eV for thermal reaction $CH_3O_{Ti} + O_b = CH_2O_{Ti} + OH_b$. The energy barrier for this reaction occurring is 1.6 eV; the corresponding minimal energy path is shown on Figure 4-2. The electronic structure of the product of the reaction features two

bridging hydroxyls and localized electrons (polarons) associated with them. In the discussion of the excited state trajectory we will see how the final electronic structure will be the same as that of the ground-state trajectory.

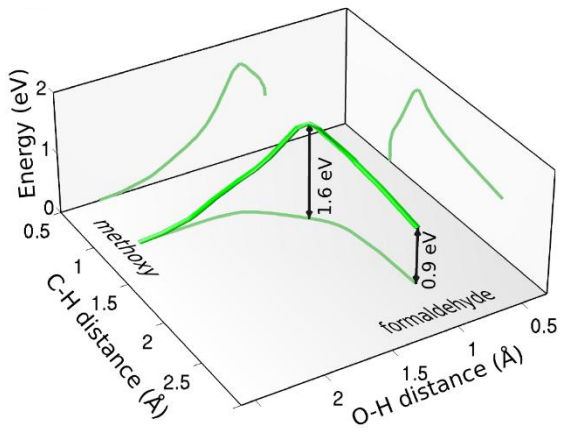

Figure 4-2. Minimal energy pathway for methoxy C–H bond thermal cleavage.

NEB calculation showing the 1.6 eV barrier for the methoxy-to-formaldehyde reaction, and the 0.9 eV reaction enthalpy. The geometry region corresponding to the reagent is of short C–H distance and large O–H distance; in the course of the reaction the O–H bond distance decreases to equilibrium value.

4.4. Excited-state dynamics of methoxy photo-oxidation

The photo-oxidation of methanol is widely studied due to its hole scavenging ability,[40-44] as well as the first step of practically relevant route to synthesizing methyl formate.[45-47] As mentioned before, the cleavage of methoxy C–H bond with formaldehyde formation can be alternatively achieved by creating an excitation across the titania band-gap of 3.1 eV.[48] This reaction – production of formaldehyde from methoxy on rutile TiO_2 (110) is an excellent test case for simulating an elementary reaction driven by creation of an excited state.

4.4.1. Computational methods

We used the self-consistent field (Δ-SCF) approach to create the photo-generated electron–hole pair, a procedure that includes promotion of one electron from the highest

occupied to the lowest unoccupied orbital and simultaneous relaxation of all orbitals. We choose this approach for several reasons: 1) it has been shown to have good accuracy in modeling the lowest-energy excitation within certain symmetry, which is sufficient for the purposes of this work; 2) it can be in principle exact and 3) *most importantly* for the present study, in contrast to linear response TDDFT, the Δ-SCF approach does not build upon the Kohn-Sham (KS) eigenstates of the ground state Hamiltonian; instead, the orbitals are variationally optimized with the excitation. This can be thought of as being similar to a multi-configuration SCF scheme in quantum chemistry applied to a fixed number of singly-excited Slater determinants with fixed coefficients. This latter property of Δ-SCF method partially compensates for the charge-transfer state problem in TDDFT and makes it possible to describe the excited state as a single Slater determinant with two easy to track singly occupied excited orbitals. The resulting lowest excitation energy is 2.5 eV, close to the value of titania band gap. In order to capture the dynamics of the coupled electron-ion system we used non-adiabatic Ehrenfest dynamics in the context of real-time time-dependent density functional theory (TDDFT), as implemented in TDAP-2.0 code.[49] This approach involves no empirical parameters and can reproduce with near chemical accuracy the essential features of the system.

After geometry optimization, we sampled the potential energy surface using ground state Born-Oppenheimer molecular dynamics (BOMD) at 200 K until thermal equilibration to obtain a variety of starting point configurations for excited-state dynamics analogous to that shown in Figure 4-1(a). The majority of randomly selected configurations resulted in methoxy photo-oxidation during the first 200 fs of the simulation, with all successful trajectories having qualitatively similar reaction mechanism. In the following, we discuss one of the most illustrative trajectories in order to highlight the relevant features of the reaction mechanism.

4.4.2. Lewis structures from first principles: the C–H acidity mechanism

Response of the system to the motion of charge carriers can be inferred through the changes of the key bond lengths: the C–H bond being cleaved and the O–H bond being formed, which are presented in Figure 4-3 along with the length of adsorbate C–O bond. We can immediately infer the presence of an intermediate with bond order 1.5 and lifetime of ~60 fs, which is created after C–H bond cleavage; however, an analysis of the dynamics of electronic system is necessary to establish its identity.

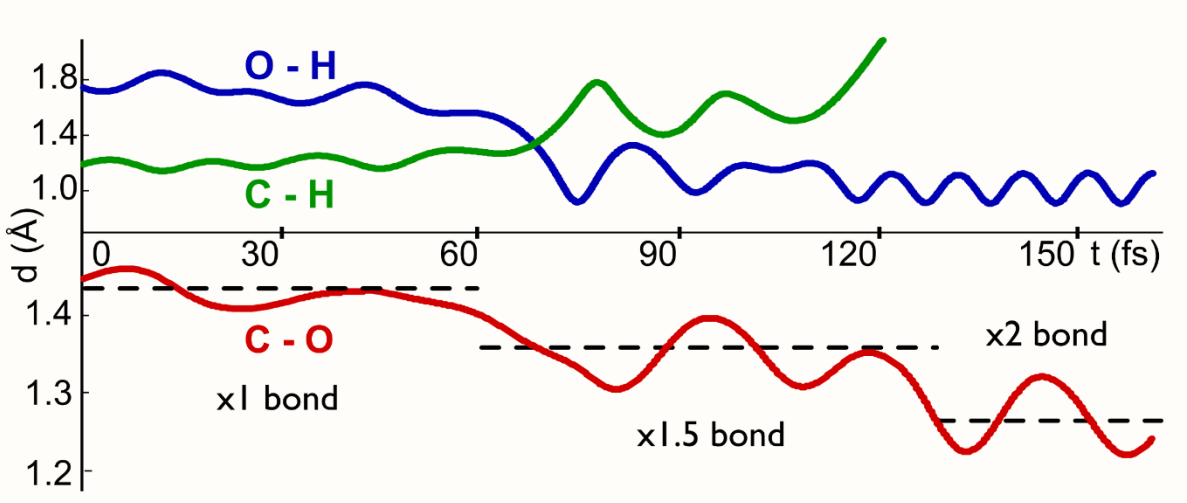

Figure 4-3. Bond length evolution along the excited state path.

Three clearly distinct regions can be inferred based on the average length of C–O bond: single bond of methoxy from 0 to 60 fs; bond of order 1.5 from 60 to 135 fs, and double bond of formaldehyde after 135 fs.

In order to analyze the electronic dynamics, we have tracked the time evolution of frontier orbitals: $\tilde{\phi}_{N_{occ}}$, the "hole" state, from which electron is removed, and $\tilde{\phi}_{N_{occ}+1}$, the previously vacant "electron" state; jointly, those states represent the spin polarization in the system and highlight the chemically active areas. The evolution of the hole state is the most

relevant for a photo-oxidation reaction; it represents depletion of spin-down electrons from previously occupied regions and accumulation of spin-up electrons in previously unoccupied regions.

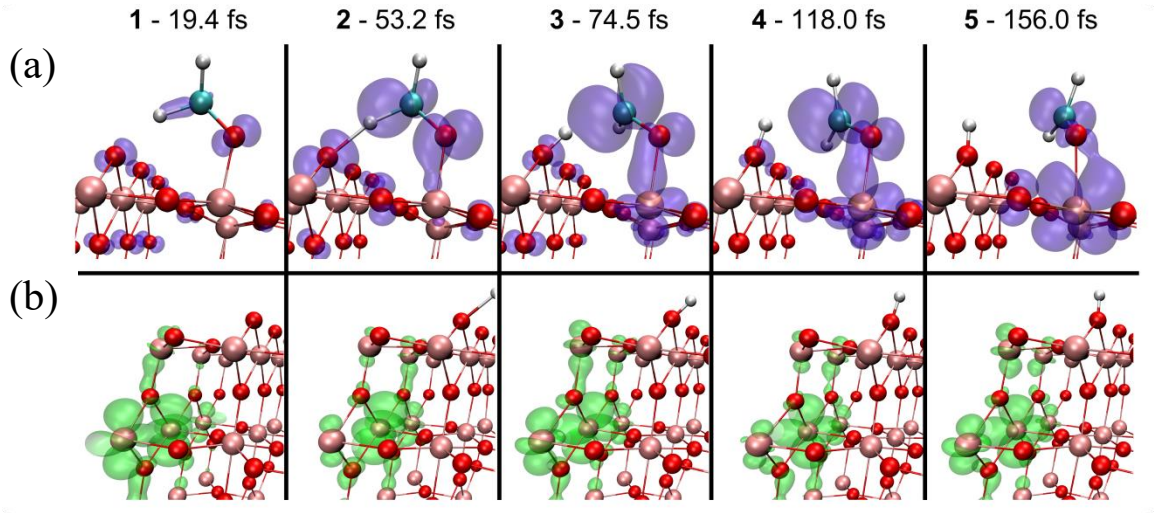

Figure 4-4. Snapshots of the "hole" and "electron" densities along the dynamics trajectory. (a) The "hole" state; blue isosurface (isosurface value 0.002 a.u.) depicts spin-up magnetization density. Prior to the proton abstraction (frames 1-2) spin-up magnetization density corresponds to depletion of electron density, while after the proton abstraction (frames 3-5) to the excess of it. (b) The "electron" state; green isosurface depicts spin-down magnetization density that corresponds to the excited electron. It forms a stable polaron and remains inactive during the simulation.

Remarkably, the snapshots of "hole" evolution in real time (Figure 4-4) contain a wealth of information about the intermediate steps of the reaction. This information is embodied in the shape changes of the hole orbital, which can be translated into a qualitative interpretation of the mechanism using Lewis structures, which we present below. The "electron" is transferred to the titanium atoms in the second subsurface layer with the formation of a polaron (that is, an electron trapped in a lattice distortion), where it stays throughout the simulation. This behavior is consistent with previous work characterizing titanium *d* states as efficient electron traps.[12, 21, 34]

First, we note that after photo-excitation the hole state is delocalized throughout the slab on the *p*-type lone pairs of oxygen atoms, and partly on the n_σ lone pair of methoxy. Contrary to the commonly held view, the hole is not "trapped" on any of the surface sites, but is drawn to the molecule in the course of the reaction. A favorable coplanar arrangement of the C–H_t bond (H_t denotes the abstracted hydrogen atom) and the oxygen n_σ lone pair facilitates the transfer of electron density from the σ-bond to the methoxy O atom through a hyperconjugation mechanism. As a result of this process, the hole density spreads to the C–H_t bond in ~20 fs (Figure 4-5).

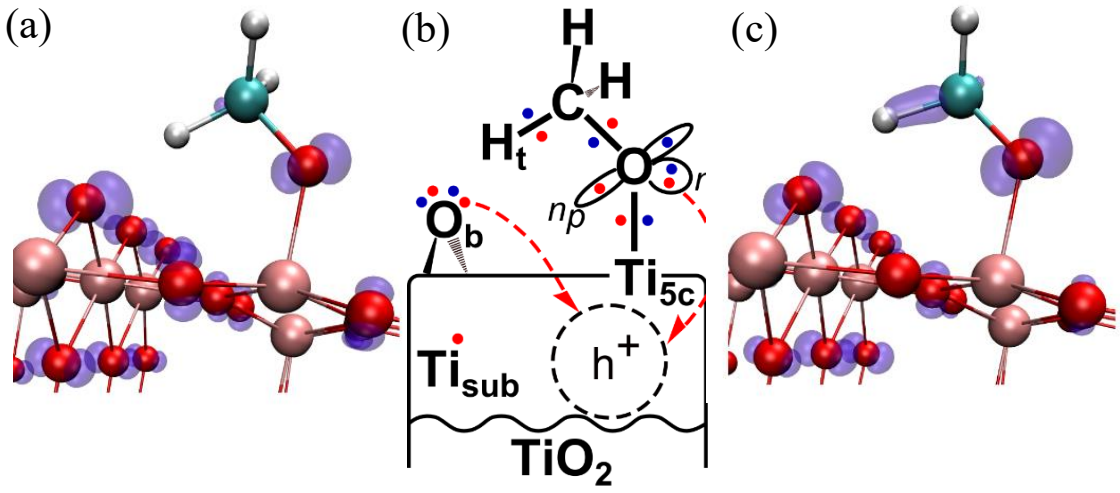

Figure 4-5. The first fragment of the trajectory.

The snapshots of the hole wavefunction at the beginning of the trajectory (a) and at 19.4 fs (c). The onset of hole migration from the slab to the adsorbate is shown with electron pushing scheme in the Lewis structure (b).

Due to the *s*-rich character of the n_σ lone pair, positioning the hole there is less favorable compared to the n_p lone pair. During the interval of 20 to 50 fs the hole on methoxy O atom rotates to become almost normal to the C–O–Ti plane. At the same time, C–H_t bond rotates around the C–O axis to maintain the stabilizing hyperconjugation with the hole lobe on methoxy

O atom. The pronounced directional character of the *p*-orbital results in a stronger overlap between n_p and σ_{CH_t} compared to that between n_σ and σ_{CH_t}. Therefore, electron donation from σ_{CH_t} increases simultaneously with O hole lobe rotation. Additional evidence for electron transfer from σ_{CH_t} is provided by 0.07 Å increase of the amplitude of C–H stretching oscillations, which is consistent with the increase of the hole lobe size between those atoms and the weakening of the bond. Stabilizing hyperconjugation makes methoxy an energetically preferred location for the hole compared to oxygen lone pairs in titania, resulting in a complete one-electron oxidation of the adsorbate by 50 fs (Figure 4-6):

$$\mathrm{CH_3O}^-_{ads} + h^+_{slab} = \mathrm{CH_3O}^{\cdot}_{ads} \qquad (4.4.1)$$

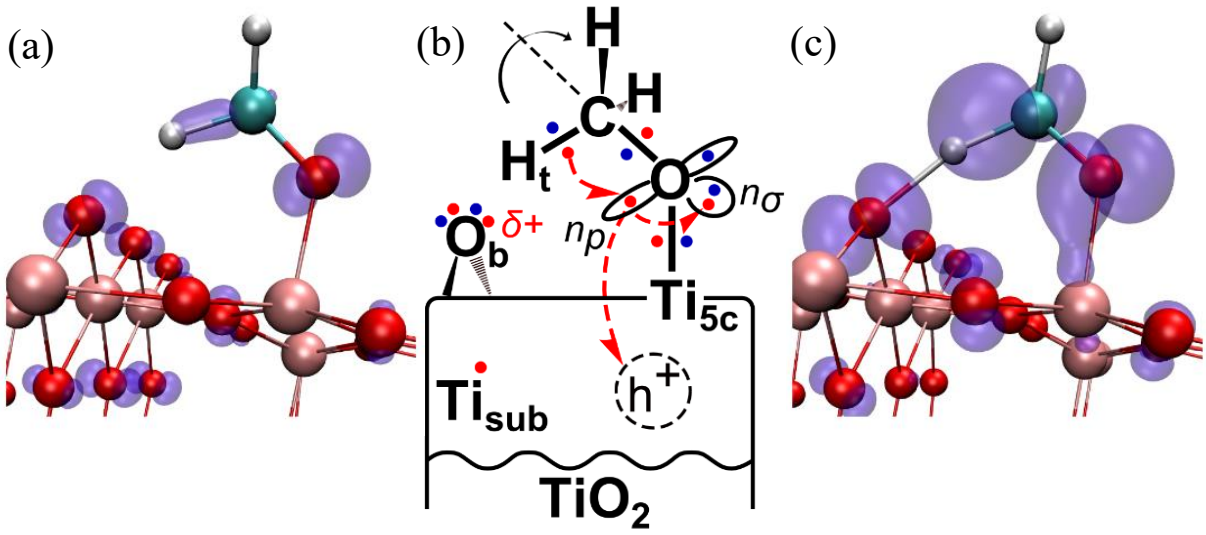

Figure 4-6. Hole migration to the adsorbed methoxy species.

The snapshots of the hole wavefunction at 19.4 fs (a) and at 53.2 fs (c). The complete electron transfer from methoxy to the titania slab, resulting in complete hole quenching, is shown with electron pushing scheme in the Lewis structure (b).

The depletion of electron density from the C–H$_t$ bond manifests itself as an increase of the hole lobe in that region and weakens the σ_{C-H_t} bond sufficiently to allow its pronouncedly

heterolytic cleavage. The adjacent bridging oxygen serves as Brønsted base for proton abstraction in this process (Figure 4-7):

$$CH_3O^{\cdot}_{ads} + O_b = CH_2O^{\cdot -}_{ads} + OH^{+}_{b} \quad (4.4.2)$$

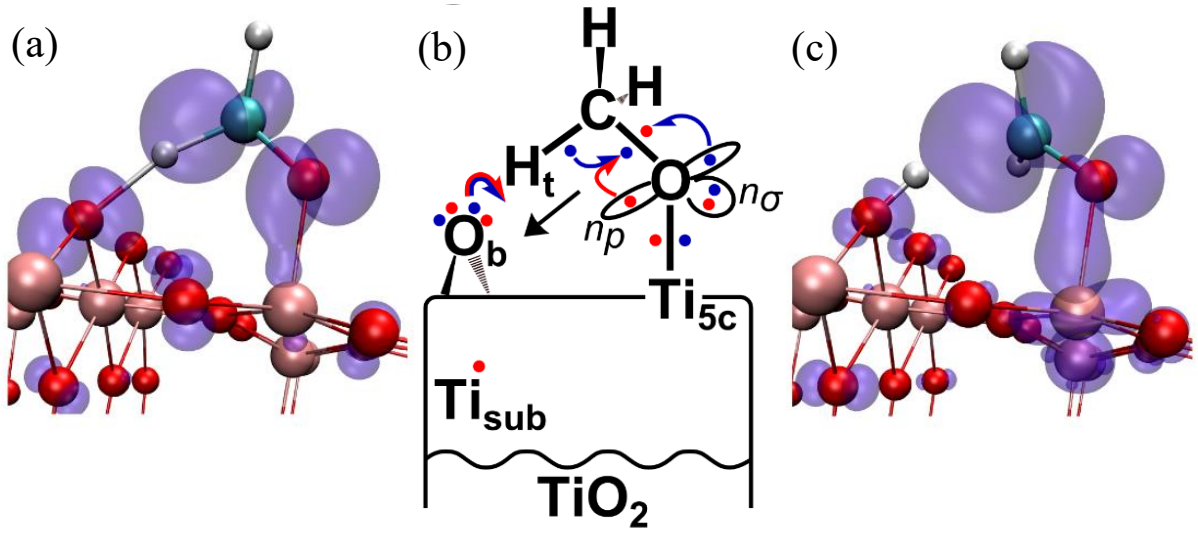

Figure 4-7. The cleavage of The C–H bond.

The snapshots of the hole wavefunction at 53.2 fs (a) and at 74.5 fs (c). The reorganization of formaldehyde anion-radical π system and simultaneous hydrogen abstraction are shown in the Lewis structure (b). Note the increase in the size of the back lobe on carbon atom and the shape of the hole orbital in (c) strongly resembling π^*.

Simultaneously with H abstraction, the hole on the C atom acquires a substantial back lobe, corresponding to the increase of *p*-type character as opposed to sp^3-type. At the same time the remaining CH_2O fragment acquires a planar geometry and the C–O bond shortens from 1.45 Å (bond length typical for single bond in alcohols) to 1.35 Å, intermediate between a single and a double bond, indicating a bond order of 1.5. The C–H_t bond cleavage induces electronic structure rearrangements which can be represented as:

$$(sp^3_C)^{1+\delta} \rightarrow (p_C)^{1+\delta} \quad (4.4.3)$$

$$(p_C)^{1+\delta}(n_{p,O})^{2-\delta} \rightarrow (\pi_{CH_2O})^2(\pi^*_{CH_2O})^1 \quad (4.4.4)$$

where δ stands for the fraction of the hole charge on the O n_p lone pair. Interpretation of the hole state observed at 75 fs as a singly occupied formaldehyde antibonding π^* orbital is further supported by its larger amplitude on the C atom, consistent with lower electronegativity of C compared to that of O. The presence of 3 electrons in the π system of formaldehyde explains the intermediate C–O bond length. We note that the hole from this moment on represents the presence of excess spin-up electron in the previously unoccupied antibonding state.

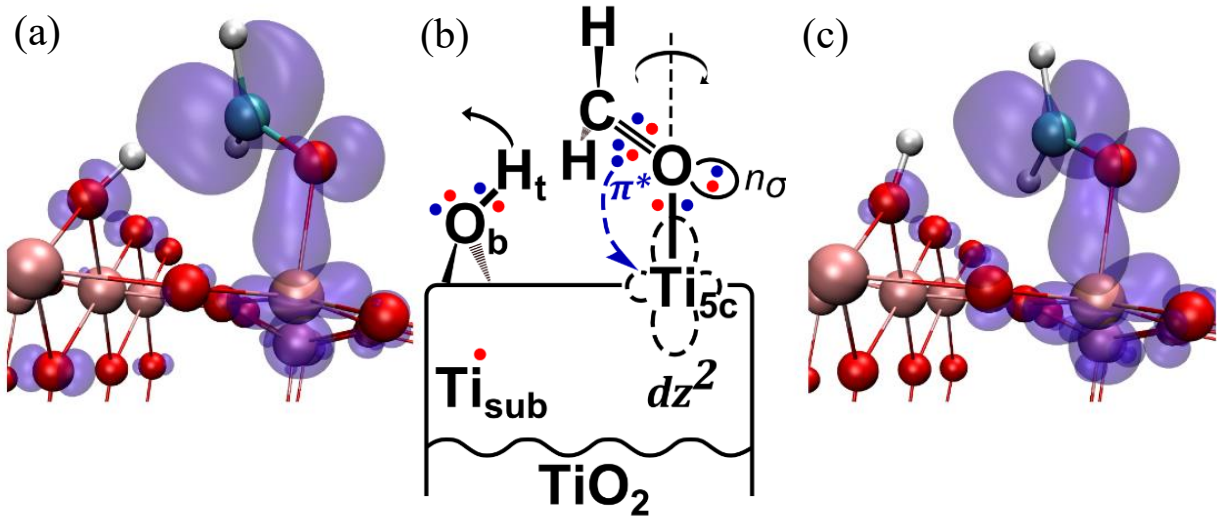

Figure 4-8. Formaldehyde anion-radical intermediate.

The snapshots of the hole wavefunction at 74.5 fs (a) and at 118 fs (c). The separation of methoxy C and transferred H_t atoms, as well as the symmetry-hindered transfer of extra electron from the intermediate to surface Ti_{5c} are shown in the Lewis structure (b).

The settling of bond oscillations in the newly formed O–H bond takes place during a period of 40 fs after cleavage of the C–H_t bond, leading to the complete decoupling of the O_b–H and CH_2O fragments (Figure 4-8). In the meantime, the π^* state of formaldehyde overlaps with the d_{z^2} orbital of the surface Ti atom, due to the tendency of the anion-radical to transfer the

electron from the antibonding orbital to the nonbonding level of the surface Ti atom. The π^*–d_{z^2} overlap is hindered by different symmetry with respect to the C–O–Ti plane. In order to achieve more favorable π type overlap with the Ti d_{zx} orbital, formaldehyde rotates around the Ti–O bond and aligns with the [001] direction of rows of under-coordinated O and Ti atoms on the surface. This rapid rotation takes place between points 115 and 135 fs; the angle between the C–O bond and the [001] direction decreases by 40^0 in that time interval.

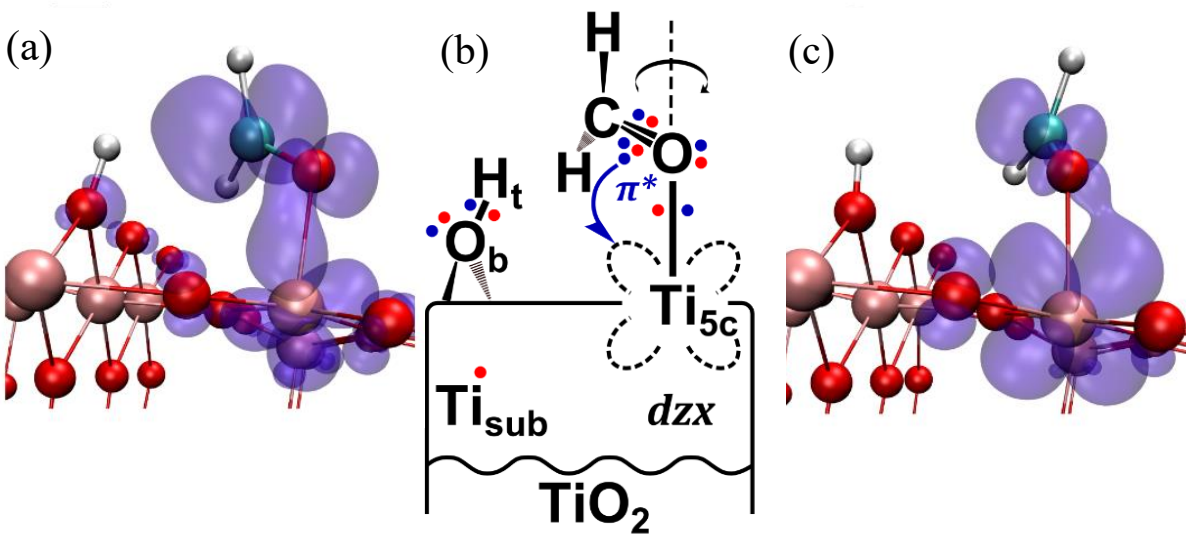

Figure 4-9. The formation of a polaron.

The snapshots of the hole wavefunction at 118 fs (a) and at 156 fs (c). The transfer of an electron from the antibonding level of the intermediate to titania with formation of a polaron is shown in the Lewis structure (b).

Comparison of the hole state snapshots (Figure 4-9) shows almost complete electron transfer from the formaldehyde π^* orbital to the Ti d_{zx} on the surface. Consistent with that, the C–O average distance decreases to 1.23 Å, indicating an increase of the bond order between those species from 1.5 to 2. Overall, the final stage of the process can be written as

$$CH_2O^{\cdot-}_{ads} = CH_2O_{ads} + e^-_{pol} \qquad (4.4.5)$$

where e_{pol}^{-} indicates an electron trapped at the surface Ti atom with formation of a polaron.

Additional insights can be obtained by studying the process energetics (Figure 4-10). On the excited-state energy landscape the system moves downhill in energy, and no substantial energy barriers are encountered. The initial excitation energy of 2.5 eV is expended in part (1.2 eV) toward driving the endothermic methoxy-to-formaldehyde transformation, and the rest is dissipated through ionic motion. In the end of the excited-state trajectory the electronic structure becomes qualitatively similar to that of the ground state of the products, namely, a neutral formaldehyde molecule and two bridging hydroxyls with the extra electrons localized as polarons on Ti atoms. Therefore, we can consider the adsorbate oxidation reaction as a pathway for excitation quenching.

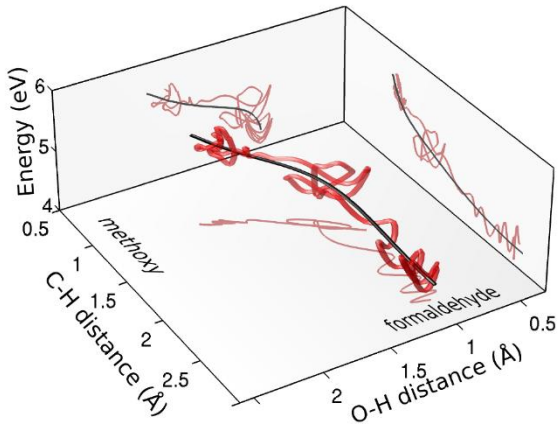

Figure 4-10. The excited-state pathway of methoxy C-H bond cleavage.

The electronic energy of the system after excitation along the reaction trajectory (red). The black line is a guide to the eye. In the course of the reaction the electronic system energy decreases by 1.2 eV.

Figure 4-11. The C–H acidity mechanism of methoxy photo-oxidation on titania.

(a) The overall trajectory of the reaction; blue denotes the ground-state molecular dynamics, red – excited-state. Numbers correspond to the specific points in time: 1 – 19.4 fs, 2 – 53.2 fs, 3 – 74.5 fs, 4 – 118.0 fs, 5 – 156 fs. Darker red corresponds to slow atomic motion, brighter red (between points 2 and 3) highlights the rapidly traversed part of the trajectory.

(b) The snapshots of the hole wavefunction in the specified time moments, see text for discussion.

(c) Interatomic distances: C–H bond being cleaved, O–H being formed, and methoxy C–O. Based on the electronic structure analysis, the intermediate (highlighted in yellow) can be identified as formaldehyde anion-radical. Note the increase in C–O bond oscillation frequency with increase in bond order and the harmonic character of newly formed O–H bond vibrations.

(d) Electron pushing schemes depicting the rearrangement of electronic structure leading from the previous time frame to the current one; reaction pathway from the oxidation of methoxy, through the C–H bond cleavage with the formation of anion-radical and, finally, to the formation of neutral formaldehyde and polaron on titania can be seen in its entirety.

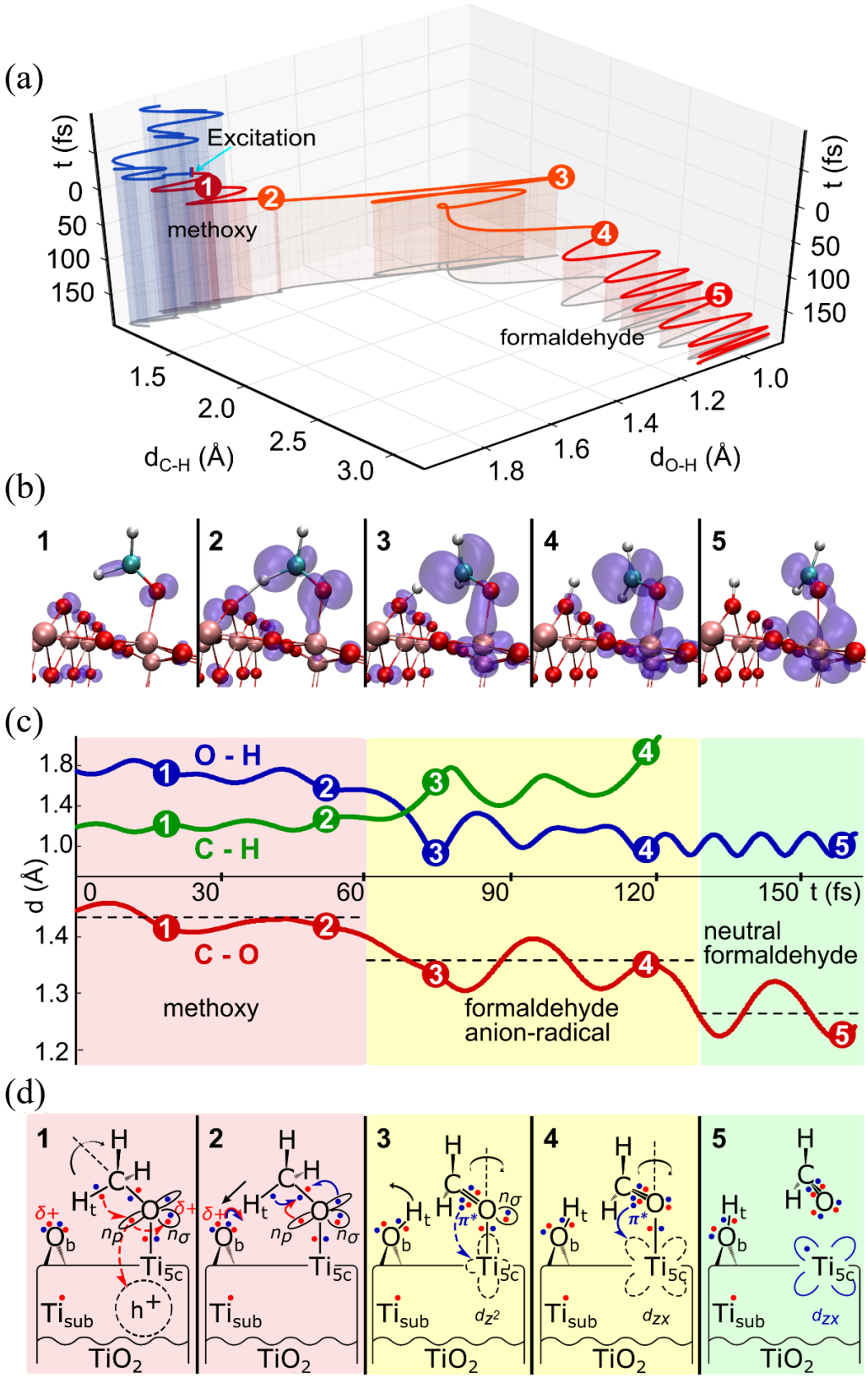

Figure 4-11 (continued).

Our analysis enables us to put forward a generalization of the C–H acidity mechanism (Figure 4-11) – a set of design principles for efficient photo-oxidation reactions. First is the strong coupling of the reactants and reactive intermediates to the semiconductor surface – in our case, methoxy and the formaldehyde anion. Second is the ability of the semiconductor surface to accommodate the reaction products – for the rutile TiO_2 (110) surface, it is the presence of efficient electron trapping sites on surface titanium atoms and Brønsted basic sites for proton abstraction on bridging oxygen atoms. Finally, the adsorbate should have non-hindered geometric and electronic structure rearrangement options, which make the adsorbed molecule a preferred location for the hole compare to the photo-catalyst. In this case this role was fulfilled by the hyperconjugation between the C–H bond and the oxygen lone pairs, as well as the sterically available rotations of different atomic groups in the adsorbate which sustained and enhanced that stabilizing effect. The above-mentioned properties will make the reaction pathway an efficient and fast channel for photo-generated carrier relaxation.

4.5. References

1. Fujishima, A.; Zhang, X.; Tryk, D. A., TiO_2 Photocatalysis and Related Surface Phenomena. *Surface Science Reports* **2008**, *63*, 515-582.

2. Kudo, A.; Miseki, Y., Heterogeneous Photocatalyst Materials for Water Splitting. *Chemical Society Reviews* **2009**, *38*, 253-278.

3. Linsebigler, A. L.; Lu, G. Q.; Yates, J. T., Photocatalysis on TiO_2 Surfaces - Principles, Mechanisms, and Selected Results. *Chemical reviews* **1995**, *95*, 735-758.

4. Henderson, M. A., A Surface Science Perspective on TiO_2 Photocatalysis. *Surface Science Reports* **2011**, *66*, 185-297.

5. Henderson, M. A.; Lyubinetsky, I., Molecular-Level Insights into Photocatalysis from Scanning Probe Microscopy Studies on TiO_2 (110). *Chemical reviews* **2013**, *113*, 4428-4455.

6. Diebold, U., The Surface Science of Titanium Dioxide. *Surface Science Reports* **2003**, *48*, 53-229.

7. Fujishima, A.; Honda, K., Electrochemical Photolysis of Water at a Semiconductor Electrode. *Nature* **1972**, *238*, 37-38.

8. Kiejna, A.; Pabisiak, T.; Gao, S. W., The Energetics and Structure of Rutile TiO_2 (110). *J. Phys.-Condes. Matter* **2006**, *18*, 4207-4217.

9. Pang, C. L.; Lindsay, R.; Thornton, G., Chemical Reactions on Rutile TiO_2 (110). *Chemical Society Reviews* **2008**, *37*, 2328-2353.

10. Mostofi, A. A.; Yates, J. R.; Lee, Y. S.; Souza, I.; Vanderbilt, D.; Marzari, N., Wannier90: A Tool for Obtaining Maximally-Localised Wannier Functions. *Comput Phys Commun* **2008**, *178*, 685-699.

11. Thompson, T. L.; Yates, J. T., Surface Science Studies of the Photoactivation of TiO_2 - New Photochemical Processes. *Chemical Reviews* **2006**, *106*, 4428-4453.

12. Deskins, N. A.; Rousseau, R.; Dupuis, M., Localized Electronic States from Surface Hydroxyls and Polarons in TiO_2 (110). *Journal of Physical Chemistry C* **2009**, *113*, 14583-14586.

13. Chretien, S.; Metiu, H., Electronic Structure of Partially Reduced Rutile TiO_2 (110) Surface: Where Are the Unpaired Electrons Located? *Journal of Physical Chemistry C* **2011**, *115*, 4696-4705.

14. Deskins, N. A.; Rousseau, R.; Dupuis, M., Distribution of Ti^{3+} Surface Sites in Reduced TiO_2. *Journal of Physical Chemistry C* **2011**, *115*, 7562-7572.

15. Di Valentin, C.; Pacchioni, G.; Selloni, A., Electronic Structure of Defect States in Hydroxylated and Reduced Rutile TiO_2 (110) Surfaces. *Physical Review Letters* **2006**, *97*, 166803.

16. Janotti, A.; Varley, J. B.; Rinke, P.; Umezawa, N.; Kresse, G.; Van de Walle, C. G., Hybrid Functional Studies of the Oxygen Vacancy in TiO_2. *Physical Review B* **2010**, *81*, 085212.

17. Yim, C. M.; Pang, C. L.; Thornton, G., Oxygen Vacancy Origin of the Surface Band-Gap State of TiO_2 (110). *Physical Review Letters* **2010**, *104*, 036806.

18. Di Valentin, C., Scanning Tunneling Microscopy Image Simulation of the Rutile (110) TiO_2 Surface with Hybrid Functionals and the Localized Basis Set Approach. *Journal of Chemical Physics* **2007**, *127*, 154705.

19. Pan, H.; Gu, B. H.; Zhang, Z. Y., Phase-Dependent Photocatalytic Ability of TiO_2: A First-Principles Study. *J Chem Theory Comput* **2009**, *5*, 3074-3078.

20. Hansen, J. O.; Huo, P.; Martinez, U.; Lira, E.; Wei, Y. Y.; Streber, R.; Laegsgaard, E.; Hammer, B.; Wendt, S.; Besenbacher, F., Direct Evidence for Ethanol Dissociation on Rutile TiO_2 (110). *Physical Review Letters* **2011**, *107*, 136102.

21. Deskins, N. A.; Rousseau, R.; Dupuis, M., Defining the Role of Excess Electrons in the Surface Chemistry of TiO_2. *Journal of Physical Chemistry C* **2010**, *114*, 5891-5897.

22. Wang, Z. T.; Deskins, N. A.; Henderson, M. A.; Lyubinetsky, I., Inhibitive Influence of Oxygen Vacancies for Photoactivity on TiO_2 (110). *Physical Review Letters* **2012**, *109*, 266103.

23. Sanchez, V. M.; Cojulun, J. A.; Scherlis, D. A., Dissociation Free Energy Profiles for Water and Methanol on TiO_2 Surfaces. *Journal of Physical Chemistry C* **2010**, *114*, 11522-11526.

24. Zhao, J.; Yang, J.; Petek, H., Theoretical Study of the Molecular and Electronic Structure of Methanol on a TiO_2 (110) Surface. *Physical Review B* **2009**, *80*, 235416-235416.

25. Phillips, K. R.; Jensen, S. C.; Baron, M.; Li, S.-C.; Friend, C. M., Sequential Photo-Oxidation of Methanol to Methyl Formate on TiO_2 (110). *Journal of the American Chemical Society* **2013**, *135*, 574-577.

26. Farfan-Arribas, E.; Madix, R. J., Different Binding Sites for Methanol Dehydrogenation and Deoxygenation on Stoichiometric and Defective TiO_2 (110) Surfaces. *Surface Science* **2003**, *544*, 241-260.

27. Farfan-Arribas, E.; Madix, R. J., Role of Defects in the Adsorption of Aliphatic Alcohols on the TiO_2 (110) Surface. *Journal of Physical Chemistry B* **2002**, *106*, 10680-10692.

28. Weinhold, F.; Landis, K., *Valency and Bonding: A Natural Bond Orbital Donor-Acceptor Perspective*; Cambridge University Press: New York, 2005, p 760.

29. Migani, A.; Mowbray, D. J.; Iacomino, A.; Zhao, J.; Petek, H.; Rubio, A., Level Alignment of a Prototypical Photocatalytic System: Methanol on TiO_2 (110). *Journal of the American Chemical Society* **2013**, *135*, 11429-11432.

30. Migani, A.; Mowbray, D. J.; Zhao, J.; Petek, H.; Rubio, A., Quasiparticle Level Alignment for Photocatalytic Interfaces. *J Chem Theory Comput* **2014**, *10*, 2103-2113.

31. Henkelman, G.; Jonsson, H., Improved Tangent Estimate in the Nudged Elastic Band Method for Finding Minimum Energy Paths and Saddle Points. *Journal of Chemical Physics* **2000**, *113*, 9978-9985.

32. Henkelman, G.; Uberuaga, B. P.; Jonsson, H., A Climbing Image Nudged Elastic Band Method for Finding Saddle Points and Minimum Energy Paths. *Journal of Chemical Physics* **2000**, *113*, 9901-9904.

33. Kresse, G.; Furthmuller, J., Efficient Iterative Schemes for Ab Initio Total-Energy Calculations Using a Plane-Wave Basis Set. *Physical Review B* **1996**, *54*, 11169-11186.

34. Kowalski, P. M.; Camellone, M. F.; Nair, N. N.; Meyer, B.; Marx, D., Charge Localization Dynamics Induced by Oxygen Vacancies on the TiO_2 (110) Surface. *Physical Review Letters* **2010**, *105*, 146405.

35. Kang, W.; Hybertsen, M. S., Quasiparticle and Optical Properties of Rutile and Anatase TiO_2. *Physical Review B* **2010**, *82*, 085203.

36. Calzado, C. J.; Hernandez, N. C.; Sanz, J. F., Effect of on-Site Coulomb Repulsion Term *U* on the Band-Gap States of the Reduced Rutile (110) TiO_2 Surface. *Physical Review B* **2008**, *77*, 045118.

37. Morgan, B. J.; Watson, G. W., A DFT+*U* Description of Oxygen Vacancies at the TiO_2 Rutile (110) Surface. *Surface Science* **2007**, *601*, 5034-5041.

38. Anisimov, V. I.; Zaanen, J.; Andersen, O. K., Band Theory and Mott Insulators – Hubbard-*U* Instead of Stoner-*I*. *Physical Review B* **1991**, *44*, 943-954.

39. Perdew, J. P.; Burke, K.; Ernzerhof, M., Generalized Gradient Approximation Made Simple. *Physical Review Letters* **1996**, *77*, 3865-3868.

40. Di Valentin, C.; Fittipaldi, D., Hole Scavenging by Organic Adsorbates on the TiO_2 Surface: A DFT Model Study. *Journal of Physical Chemistry Letters* **2013**, *4*, 1901-1906.

41. Zhou, C., et al., Site-Specific Photocatalytic Splitting of Methanol on TiO_2 (110). *Chemical Science* **2010**, *1*, 575-580.

42. Shen, M.; Acharya, D. P.; Dohnalek, Z.; Henderson, M. A., Importance of Diffusion in Methanol Photochemistry on TiO_2 (110). *Journal of Physical Chemistry C* **2012**, *116*, 25465-25469.

43. Shen, M.; Henderson, M. A., Identification of the Active Species in Photochemical Hole Scavenging Reactions of Methanol on TiO_2. *Journal of Physical Chemistry Letters* **2011**, *2*, 2707-2710.

44. Shen, Q., et al., Charge Transfer and Recombination at the Metal Oxide/$CH_3NH_3PbClI_2$/Spiro-OMeTAD Interfaces: Uncovering the Detailed Mechanism Behind High Efficiency Solar Cells. *Physical Chemistry Chemical Physics* **2014**, *16*, 19984-19992.

45. Guo, Q.; Xu, C.; Yang, W.; Ren, Z.; Ma, Z.; Dai, D.; Minton, T. K.; Yang, X., Methyl Formate Production on TiO_2 (110), Initiated by Methanol Photocatalysis at 400 nm. *Journal of Physical Chemistry C* **2013**, *117*, 5293-5300.

46. Jensen, S. C.; Friend, C. M., The Dynamic Roles of Interstitial and Surface Defects on Oxidation and Reduction Reactions on Titania. *Topics in Catalysis* **2013**, *56*, 1377-1388.

47. Yuan, Q.; Wu, Z. F.; Jin, Y. K.; Xu, L. S.; Xiong, F.; Ma, Y. S.; Huang, W. X., Photocatalytic Cross-Coupling of Methanol and Formaldehyde on a Rutile TiO_2 (110) Surface. *Journal of the American Chemical Society* **2013**, *135*, 5212-5219.

48. Guo, Q.; Xu, C.; Ren, Z.; Yang, W.; Ma, Z.; Dai, D.; Fan, H.; Minton, T. K.; Yang, X., Stepwise Photocatalytic Dissociation of Methanol and Water on TiO_2 (110). *Journal of the American Chemical Society* **2012**, *134*, 13366-13373.

49. Kolesov, G.; Granas, O.; Hoyt, R.; Vinichenko, D.; Kaxiras, E., Real-Time TD-DFT with Classical Ion Dynamics: Methodology and Applications. *J Chem Theory Comput* **2016**, *12*, 466-476.

PART II

MODELING THE DEFECT PROPERTIES OF SEMICONDUCTORS AND THE PERFORMANCE OF SEMICONDUCTOR-BASED SOLAR CELLS

We start this part by discussing in Chapter 5 an internally consistent scheme for computation of defect formation energies in the density functional theory simulations. The key advantages of the described approach are the absence of free parameters and the applicability to both bulk semiconductor surfaces and two-dimensional materials. In Chapter 6 we discuss an application of that scheme to studying the properties of sulfur vacancies in a material of substantial interest – MoS_2 monolayer. We analyze the defect-induced perturbations of electronic and geometric structure of the material and its behavior under strain in the context of "ripplocation" structure formation. Finally, in Chapter 7 we demonstrate a model for estimating the power conversion efficiency of *p-i-n* solar cells based on the electronic structure of the intrinsic absorber layer, and apply it to the important class of solar cells based on hybrid organic-inorganic absorbers with perovskite structure. Our findings allow us to demonstrate how the device performance can be linked to the chemical composition of the material and the alignment of its electronic structure to that of carrier-transporting materials.

CHAPTER 5. A CONSISTENT METHOD FOR COMPUTATION OF DEFECT FORMATION ENERGIES IN NON-HOMOGENEOUS DIELECTRIC ENVIRONMENTS

Defects can have a dramatic impact on the properties of semiconducting materials;[1-2] the most prominent effect is the *p*- and *n*-doping of semiconductors, when the trace amounts of the dopant dramatically enhance the conductivity. More than that, the introduction of defect levels in the band gap changes the optical properties of the material: absorption or luminescence, and facilitates the non-radiative recombination of carriers on the defect levels in the band gap.[3] An understanding of defect and impurities behavior is also important for rationalizing the structural properties of materials: defects modify the vibrational spectrum of the host, and are often involved in the diffusion processes during materials growth.[4] A key quantity describing defect behavior is its formation energy as a function of charge. In this chapter we discuss a novel method for computation of defect formation energies of charged defects. The material presented here is being prepared for publication: D. Vinichenko, M. G. Sensoy, C. M. Friend, E. Kaxiras. Accurate formation energies of charged defects in solids: a systematic approach.

5.1. Experimental methods for studying defect properties

A wide variety of experimental techniques exist for investigating the properties of defects.[5] Secondary ion mass spectrometry (SIMS) allows measurement of the depth profile of defect concentration inside the material, achieving sensitivity of 10^{13} cm^{-3} – however, identification of native defects is not possible with this method, unless isotope techniques are used.[6] Another method for investigating point defect concentrations is positron annihilation

spectroscopy[7] (PAS), which is typically limited to working with vacancies only. Optical techniques, such as photoluminescence (PL) spectroscopy, provide only information about the presence of the defects, but do not lend any insight into the nature of the PL peaks.[8-9] Extended x-ray absorption fine structure (EXAFS) can directly provide information regarding the local atomic structure around impurity, but only for heavy atoms.[10] Scanning tunneling microscopy[11] (STM) and a variable-bias extension thereof, scanning tunneling spectroscopy[12] (STS) can be a powerful tool for revealing atomic and electronic structure of defects on surfaces and, in some cases, even for subsurface defects.[13-15] While the STM data can be related to computational simulations, a careful analysis of STS data is oftentimes required.[16] Finally, the most powerful tool for studying defects is the electron paramagnetic resonance[17-18] (EPR) – it can provide information about local environment and concentrations of the defect. Since the key output of the method are the hyperfine parameters and *g* tensors of unpaired electrons, the applications are most often limited to paramagnetic defects. Additional insight can be derived by comparing the method's output to computational simulations of the magnetic parameters,[19] which however, involves some methodological intricacies associated with the implementation of gauge-invariant theory.[20-21] However, even using EPR it is hard to establish the most important quantity determining the abundance of a given defect or impurity – its formation energy; it is due to comparative difficulty in getting the information about microscopic properties of the defects from the experimental data, the computational methods based on quantum theory of matter have been developed and used quite widely.

5.2. First-principles approaches to defect simulation

Several approaches to computation of defect electronic structure and formation energy were developed. First group of methods is based on the Green's functions formalism;[22-26]

potentially they are capable of delivering accurate results, even for hyperfine couplings relevant in the context of EPR, but the implementation of those methods is very involved, which limits the applicability. Moreover, atomic relaxations typically associated with defect formation, are very hard to include. Another group of methods models the defect immersed a finite cluster of several shells of the host material.[27-28] Such approach allows to take advantage of powerful methods of quantum chemistry for accurate treatment of electron correlation: configuration interaction (CI), complete active space self-consistent field (CASSCF), or coupled clusters(CC).[29] On the other hand, such models require termination of the dangling bonds on the outer boundary of the model cluster, and the results of the simulation are typically sensitive to such finite-size effects. Variations of the method with embedding the cluster into a phenomenologically treated host material – so-called embedded cluster models, also exist.[30] Moreover, unfavorable scaling of the quantum-chemical approaches with the model system size puts very strict limitation on the size of the cluster under consideration, adding an extra layer of uncertainty to the problem.

An alternative approach to the problem is based on the density functional theory-based (DFT) simulations with utilization of periodic boundary conditions – the so-called supercell method.[31-32] In this method, the defect is introduced in one unit cell of the host material, and this cell is then surrounded by a shell of stoichiometric host cells. Moreover, defects in the surface regions can be simulated by creating surface slabs separated by a layer of vacuum. The resulting rectangular supercell is the model for an isolated defect, for which the electronic structure and corresponding total energy can be obtained by applying the DFT formalism, as implemented in the modern efficient plane-wave codes.[33-34] Additionally, electrons can be added to or removed from the model in order to simulate the charge state of the defect in question. The formation

energy E_f^q for an impurity in a charge state q can be calculated in the Zhang-Northrup formalism[35] as:

$$E_f^q = E_{DFT}^{def,q} - E_{DFT}^{st} - \sum_i \mu_i n_i + q(E_{VBM} + E_F) \quad (5.2.1)$$

where $E_{DFT}^{def,q}$ is the DFT total energy of the supercell with a defect, E_{DFT}^{st} is the DFT total energy of the stoichiometric supercell, μ_i are the chemical potentials of the species added or removed to create the defects under appropriate thermodynamic conditions, n_i are the stoichiometric coefficients for those species ($n_i > 0$ for added atoms, $n_i < 0$ for atoms being removed), E_{VBM} is the absolute position of the valence band maximum, and E_F is the Fermi level position of the material with respect to valence band maximum.

The term in the brackets in equation 5.2.1, relevant for the charged state of the defects, corresponds to the chemical potential of the electrons in the host material; it is separated into two terms: E_{VBM} and E_F for the ease of comparison of different sets of results. This is important for carrying out comparisons of formation energies of different charge states for a given defect – for that purpose the formation energies are plotted as a function of parameter E_F – Fermi level within band gap; the intersection points of lines corresponding to different charge states are the thermodynamic transition levels. They can be compared with the data from experimental techniques, such as deep level transient spectroscopy, which is capable of inferring thermal ionization energies.[36] Such analysis allows to figure out which charge state of the defect is stable depending on the Fermi level position in the material; in the case when all present defects are known, the Fermi level in the material can be calculated[1] through a self-consistent thermodynamic calculation. There are several limitations and typical sources of errors associated with the supercell method, which we consider in further discussion.

5.3. Main sources of errors in the supercell method

The main complications resulting from utilization of supercell method based on DFT arise from two main sources: approximations of the density functional theory itself, and the artificial interactions of the defects in the adjacent supercells.[1] We consider those groups of problems in turn.

5.3.1. Errors resulting from the use of approximate functionals

In the context of the supercell simulations for defects, the limitations of density functional theory manifest themselves in two important ways: first, the electronic structure of the host material can be described inaccurately enough to induce substantial errors in the geometrical structure of host material and, consequently, the relaxations associated with the defect; second, the potentially incorrect positions of the edges of valence and conduction band can result in incorrect determination of the chemical potential of electrons in the material and wrong positions of transition levels.[37] Moreover, errors in computation of the band gap can lead to qualitatively incorrect classification of defects as shallow or deep. Those pitfalls have one common source – incorrect description of electron self-interaction, mentioned in chapter 2. In particular, the fundamental band gap E_g^f of the material can be defined as the difference between the ionization energy $I = E_{el}(N-1) - E_{el}(N)$ and electron affinity $A = E_{el}(N+1) - E_{el}(N)$, where $E_{el}(N)$ is the electronic total energy as a function of number of electrons N; $E_g^f = I - A$. For the exact exchange-correlation functional the fundamental band gap can be estimated from derivatives of electronic total energy:

$$E_g^{DFT} = \lim_{\delta \to 0} \left. \frac{\partial E_{el}}{\partial N} \right|_{N+\delta} - \lim_{\delta \to 0} \left. \frac{\partial E_{el}}{\partial N} \right|_{N-\delta} \tag{5.3.1}$$

for the total number of electrons approaching N from above and below, respectively. In the case of local or semilocal functionals this expression is written in terms of the eigenvalues corresponding to the highest occupied and lowest unoccupied orbitals, respectively: $E_g^{LDA/GGA} = \epsilon_{LUMO}^{LDA/GGA}(N) - \epsilon_{HOMO}^{LDA/GGA}(N)$. The main issue with this expression is the fact that the true exchange-correlation functional should experience a discontinuous change in the derivative once the number of electrons crosses N; for the approximate semilocal functionals, however, due to self-interaction (incomplete cancellation of Coulomb interaction of an electron with its own charge density), this is not the case – so the derivative of the exchange-correlation functional taken for $\delta > 0$ is not equal to the Kohn-Sham eigenvalue $\epsilon_{LUMO}^{LDA/GGA}$ calculated for the system of N electrons – therefore, the band gap of the material obtained in the semilocal approximation is incorrect.

The methods for coping with this band gap underestimation problem developed by scientific community can be divided into two broad categories: in the first are the empirical extrapolation approaches, and in the second are the methods for improving the electronic structure description. The extrapolation approaches are most often used in cases when the simplified semilocal functional suffices for the description of the geometry and total energy of the material, but fails for the band gap width. The simplest of such methods is a scissor operator for stretching the band gap in order to match its experimental value, leaving the defect levels with respect to the valence band intact.[38-39] More involved techniques move the levels of the donor defects along with the band gap, while leaving the positions of acceptor defects intact;[40] even more sophisticated technique involves projection of the defect level onto valence and conduction bands of the material and shifting its position in proportion to the weight of the conduction band character.[41]

The methods of the second group involve fixing the underlying description of the electronic structure of the system in question. Since the band gap issues are the most prominent for the transition metal oxides, the most popular method for electronic structure correction is the earlier described LDA+U.[42-43] More sophisticated approaches, for example, GW,[44] can be applied, but they typically have a host of issues of their own, including various flavors of GW,[45] as well as prohibitive computational cost, hence their use is less widespread. A compromise option are the hybrid DFT[46-47] functionals, which allow to address the self-interaction error without adjustable parameters (like DFT+U) and at a lower computational cost than GW.

Overall, the electronic structure component of the problem should be addressed on compound-by-compound basis, and as long as the band gap is correctly described and the defect levels are correctly placed within the band gap, this aspect can be successfully decoupled from the errors inherent to the supercell method itself, which we discuss next.[48]

5.3.2. Errors due to the interactions of a defect with its periodic images

The first issue to address, perhaps not surprisingly, is the direct overlap of the wavefunctions corresponding to the defects in the neighboring cells. This problem manifests mostly in the cases of shallow defects, which give rise to delocalized states resonant with either conduction or valence bands; even from simple tight-binding-type considerations it is clear that such overlap will give rise to non-zero matrix element of the system Hamiltonian between those orbitals, and in turn the dispersion of the defect-induced band throughout the Brillouin zone will arise. Particularly dire is the situation when the band is degenerate and only partially occupied[2, 49] – for example, in the case of doubly charged vacancy in diamond, where only two of the three degenerate orbitals are occupied. Defect band dispersion in such situation will lead to the full occupation of defect band for part of the k-points, and non-occupied orbitals in the other part of

the Brillouin zone, which, in turn, will adversely affect the properties derived from weighted summation over k-points. It should be noted, however, that the importance of this source of errors is overestimated, especially the degeneracy part – in real simulations including ionic relaxations the degeneracy of the bands will be lifted anyway through the Jahn-Teller effect; moreover, such splitting will be larger than the artificial dispersion of the defect band for any reasonable sizes of the supercells involved. For example, even without ionic relaxation the dispersion effect is virtually non-existent for supercells of diamond involving more than 64 atoms[1] – by all modern measures an extremely modest supercell size, which will not be used in actual material simulations due to insufficiency of its size for containing the defect-induced relaxations. Another issue related to the wavefunction overlap is the magnetic interaction of the defects. For semiconductors, interaction between magnetic defects can be described by a Heisenberg-type Hamiltonian with the coupling coefficients decaying as inverse cube of the inter-defect separation – not surprisingly, since magnetism involves exchange interactions, which scale as differential overlap between orbitals involved. It should be noted, however, that the situation can be substantially different for metals – for some of them the magnetoelastic coupling is very strong, so the change of magnetic state can induce substantial geometric rearrangements.[50]

The next major source of uncertainty is the elastic interaction between defects. Creation of a defect or impurity in the host lattice creates a perturbation in the electronic structure – bonds can be cleaved upon vacancy formation, they can be altered by introduction of atoms with varying electronegativity; finally, even the introduction of an atom of large radius or interstitial will create physical strain in the system. Such perturbation will exert forces on atoms and create a displacement field; the associated strain field is characterized by dipole elastic tensor,[51] so the

elastic interactions between point defects have character of dipole-dipole interaction and the elastic energy of a periodic array of defects scales as inverse cube of the supercell size. Therefore, for any reasonable supercell size the associated elastic interactions can be considered to be absent if the simulation supercell is large enough to contain the defect-induced ionic relaxations and to have "bulk-like" region of unperturbed host material around the defect – the displacements are inversely proportional to square of distance from the defect, so if the relaxations are contained. One more feature of the geometry distortion by defects is the change is the normal modes of the lattice – however, such changes are typically smaller than the other sources of errors in the simulation – such as electrostatic interaction between charged defects, to which we turn next.

5.3.3. Electrostatic energy in the supercell method

Apart from the issue mentioned above, computation of the formation energy of charged defects involves a new component – electrostatic energy of the defect charge distribution. The issue is that utilization of the periodic boundary conditions in the supercell method imposes a constraint on the total charge of the simulation cell – it has to be neutral, otherwise the electrostatic energy of an infinite array of such supercells would diverge due to long-range nature of the Coulomb interaction.[52] In practice this condition is enforced in the computational simulation by setting the average value of the electrostatic potential to zero, which is equivalent to having zero net charge. This situation gives rise to the spurious energy contributions from the interaction of the defect charge with the neutralizing background and the periodic images thereof. The magnitude of those interactions for point charge defects was shown to scale as an inverse of the supercell size,[51, 53] with additional terms resulting from higher-order moments of the charge distribution scaling as inverse cube of the supercell size. It can be shown that the

additional term in the electrostatic potential resulting from the introduction of the uniform neutralizing background has parabolic shape, and the value at the charge origin is the same as the correction term mentioned before.[54] The simplest of corrections accounting for electrostatic interaction is the Makov-Payne correction, amounting to a difference between electrostatic energy of a point charge under open boundary conditions, and the Madelung sum for its energy under periodic boundary conditions – however, in practical applications such correction was proven hard to reliably use;[55-56] one of the reasons is that the expression for the correction energy has the macroscopic dielectric constant in the denominator – but the supercell method deals with the material on a microscopic scale and therefore the bulk limit might not be applicable. Therefore, variations on the topic were developed, where the true formation energy of an isolated defect would be calculated for a series of supercells with the same shape and progressively increasing size,[57-59] and then the resulting energy would be fit to a scaling law with inverse size of the supercell, treating the dielectric constant of the material as a parameter of the model; variants of the scheme accounting for anisotropic dielectric tensor have been also implemented.[60]

More advanced methods for correcting the electrostatic component of the energy recognize the difference between the actual non-trivial shape of the charge associated with the defect, $|\phi_d(\vec{r})|^2$, and a point charge. The simplest solution is to model the defect charge with a Gaussian fit – this allows to explicitly calculate incorrect component of the total energy – electrostatic self-energy of the Gaussian model under periodic boundary conditions E_{PBC}^{mod} and the proper electrostatic energy term – energy of an isolated Gaussian charge E_{iso}^{mod}.[61-63] The benefit of using Gaussian charge is that the expressions for electrostatic energy in both cases are readily available:

$$E_{iso}^{mod} = \frac{q^2}{2\sigma\epsilon\sqrt{\pi}} \tag{5.3.2}$$

$$E_{PBC}^{mod} = \frac{2\pi}{\epsilon\Omega}\sum_{\vec{G}\neq 0}\frac{\left[\rho_{mod}(\vec{G})\right]^2}{G^2} \tag{5.3.3}$$

where q is the total charge of the defect wavefunction, ρ_{mod} is the Gaussian fit for it, $\rho_{mod}(\vec{G})$ are the Fourier transform components for the Gaussian, σ is the width of Gaussian fit, ϵ is the dielectric constant of the material and Ω is the volume of the supercell. The expressions for the model potential induced by the Gaussian model charge V_{PBC}^{mod} are also readily available. This potential is supposed to represent the response of the host material to introduction of a charge in it – mainly the dielectric screening response –and the same potential field V_{resp} can be obtained by subtracting the electrostatic potentials resulting from the DFT calculations for the stoichiometric host material, V_{DFT}^{st}, and the charged defect, $V_{DFT}^{def,q}$: $V_{resp} = V_{DFT}^{st} - V_{DFT}^{def,q}$. The Gaussian model does a reasonable job of describing the long-range behavior of the screening potential V_{resp}, but the short-range part has some discrepancy between V_{PBC}^{mod} and V_{resp} due to a much more complex structure of the actual defect wavefunction, including angular nodal features, compared to a simple isotropic Gaussian model. Since the computations are done under periodic boundary conditions, the cell averages for both V_{PBC}^{mod} and V_{resp} are zero, and therefore the discrepancy in the short-range results in a constant shift in the values of those potentials in the long-range region. This additional terms is known as "potential alignment", ΔV, and it is typically expressed as:

$$\Delta V = V_{PBC}|_{far} - \left(V_{DFT}^{st} - V_{DFT}^{def,q}\right)|_{far} \tag{5.3.4}$$

With all the terms discussed earlier, the full expression for the electrostatic correction E_{corr} can be written in terms of the model Gaussian energies for open E_{iso}^{mod} and periodic E_{PBC}^{mod} boundary conditions:

$$E_{corr} = E_{iso}^{mod} - E_{PBC}^{mod} + \Delta V \tag{5.3.5}$$

It is important to emphasize that the presence of ΔV term is solely due to usage of a crude model for the defect-induced charge $|\phi_d(\vec{r})|^2$; it was shown that the potential alignment term is unnecessary if the electrostatic energy part of the problem is properly described.[48, 64] Another motivation to remove the potential alignment term is the fact that it becomes increasingly hard to define when the relaxation of ionic positions in the material are included in the model, since motion of atoms changes the electrostatic potential substantially, and it becomes practically impossible to carry out the alignment with "far" bulk-like region in the expression for ΔV.

The models discussed so far only deal with the case of the defects in the bulk semiconductors – however, for the case of defects in the surface regions of the materials, new complications arise. In all schemes considered before, the dielectric constant of the material was treated as a constant parameter, either obtained from DFT simulations, or taken to be equal to the experimental. However, when the defects on surfaces are simulated in the supercell method, a layer of vacuum is introduced in the direction normal to the surface; the thickness of the vacuum layer is typically chosen large enough so that the periodic images of the surface slabs do not interact. In such case the dielectric constant of the material turns into dielectric *profile*. Qualitatively, the situation is depicted in Figure 5-1: the output of DFT calculation for charged defect has the defect immersed into the compensating homogeneous background, and enclosed in a finite simulation cell, whereas the goal is to simulate the formation energy of an isolated defect on the boundary of vacuum and host material without any interaction with either periodic

replicas or neutralizing charge. Not surprisingly, the electrostatic energy of such model has spurious and strong dependence on artificial parameters of the simulation domain, such as vacuum layer thickness.[65] In this work, we describe the implementation of the correction for electrostatic energy and propose an internally consistent method for computation of formation energy of defects in complex dielectric environments.

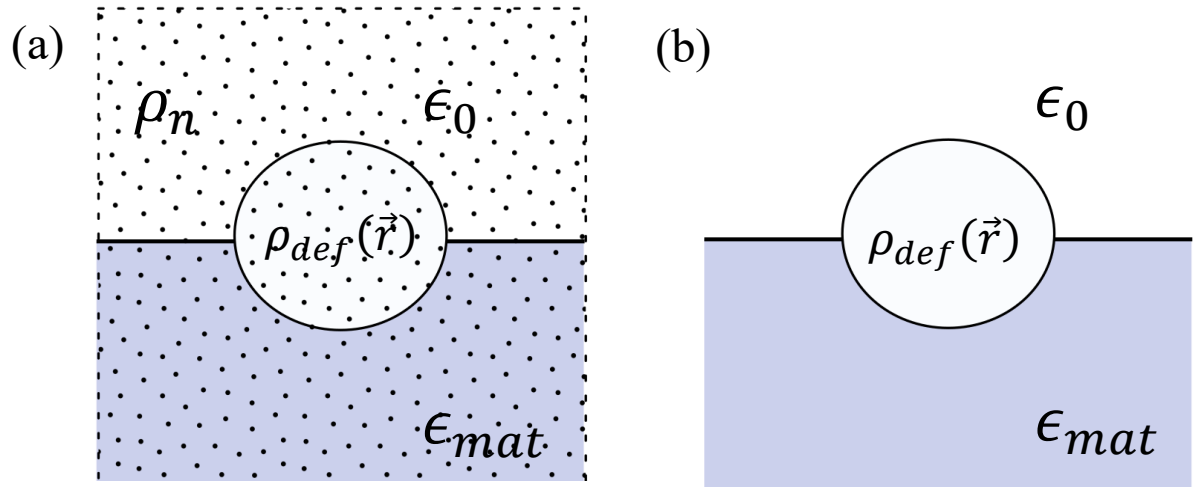

Figure 5-1. Schematic representation of a charged defect on the surface of host material.

(a) Model representing the output of DFT simulation for the surface slab containing the defect charge $\rho_{def}(\vec{r})$ on the boundary of the host material (blue shade) with dielectric constant ϵ_{mat} and vacuum with permittivity ϵ_0. Dots indicate the neutralizing homogeneous background ρ_n, and the dashed lines denote the boundary of the slab simulation cell.

(b) The target dielectric environment – an isolated defect charge on the boundary, with infinite layer of host material below it and infinite vacuum thickness above.

5.4. Computation of the electrostatic energy correction

As was discussed above, in the equation 5.3.5 for the electrostatic correction the key component is the difference between the open boundary conditions energy and periodic

boundary conditions energy, $E_{iso}^{mod} - E_{per}^{mod}$. Here we describe first the practical aspects of implementating the calculation of energy for periodic boundary conditions, followed by the discussion of issues regarding open boundary conditions computation.

5.4.1. Electrostatic energy under periodic boundary conditions

The law governing the electrostatic potential $V(\vec{r})$ can be expressed in the Poisson equation:

$$\epsilon_0 \nabla[\epsilon(z)\nabla V(\vec{r})] = -\rho_d(\vec{r}) \tag{5.4.1}$$

where $\rho_d(\vec{r})$ is the unscreened charge distribution associated with the defect level $\phi_d(\vec{r})$ in the band gap, $\rho_d(\vec{r}) = e|\phi_d(\vec{r})|^2$; ϵ_0 is the dielectric permittivity of vacuum; $\epsilon(z)$ is the dielectric profile. After inferring the electrostatic potential, it is trivial to calculate the electrostatic energy under periodic boundary conditions by integrating over the simulation supercell volume:

$$E_{PBC}^{mod} = \frac{1}{2}\int_{\Omega} V(\vec{r})\rho_d(\vec{r})d\vec{r} \tag{5.4.2}$$

Here we assume that the simulation supercell is set up in such way that the z direction is normal to the surface, and the size of the cell is (L_x, L_y, L_z), so that $x \in [0, L_x], y \in [0, L_y], z \in [0, L_z]$. With those assumption, the dielectric profile is piecewise-constant in z direction, with the value of dielectric constant 1.0 for vacuum and the corresponding value for the material in question; following previous work[65] we additionally introduce a smooth joining between the interfaces, described by an error function. We note that in principle the material can be anisotropic, that is, instead of a single dielectric constant characterized by up to three different diagonal components of the dielectric tensor: $\{\epsilon_{xx}, \epsilon_{yy}, \epsilon_{zz}\}$. When simulating this material in surface slab geometry, instead of a single dielectric profile all its three components will have dependence on the z coordinate: $\{\epsilon_{xx}(z), \epsilon_{yy}(z), \epsilon_{zz}(z)\}$. In such case all three dielectric profiles

will have the same interface positions, defined by the positions of the host material atoms. Our scheme can accommodate such case as well, as we will demonstrate below. We will describe in detail the procedure for obtaining dielectric profile in subsequent sections; for now, we assume it is known and is parameterized as described above.

Since the problem is formulated with periodic boundary conditions, it naturally lends itself to solution in the Fourier space due to availability of the fast Fourier transform numerical routines, heavily used in modern plane-wave codes for computation of the electrostatic potential. In Fourier space, the derivatives of the functions correspond to multiplication by the corresponding wavevector:

$$\hat{F}\left[\frac{\partial f}{\partial x}\right] \rightarrow iG_x\hat{F}[f] = iG_x\hat{f} \tag{5.4.3}$$

where $\hat{F}$ is the Fourier transform operator, f is the function in question, $\hat{f}$ is its Fourier transform; and the product of two functions in real space corresponds to a convolution in Fourier space:

$$\hat{F}[f \cdot g] \rightarrow \hat{F}[f] * \hat{F}[g] = \hat{f} * \hat{g} = \int \hat{f}\left(\vec{G}' - \vec{G}\right)\hat{g}\left(\vec{G}'\right)d\vec{G}' \tag{5.4.4}$$

Using those expressions, we can transform the Poisson equation in the following way:

$$\epsilon_0\nabla[\epsilon(z)\nabla V(\vec{r})] = \epsilon_0\left[\epsilon(z)\cdot\Delta V(\vec{r}) + \frac{\partial\epsilon(z)}{\partial z}\cdot\frac{\partial V(\vec{r})}{\partial z}\right] = -\rho_d(\vec{r}) \tag{5.4.5}$$

where the second term includes derivative only with respect to z dimension, since the dielectric profile is a function of z only. In Fourier space the Poisson equation will look like

$$\epsilon_0\left[\hat{\epsilon}(G_z) * |G|^2\hat{V}\left(\vec{G}\right) + G_z\hat{\epsilon}(G_z) * G_z\hat{V}\left(\vec{G}\right)\right] = \hat{\rho}_d\left(\vec{G}\right) \tag{5.4.6}$$

where $\hat{\epsilon}$ is Fourier transform of the dielectric profile, $\hat{V}$ – Fourier transform of the potential, $\hat{\rho}_d$ – Fourier transform of the defect charge, and the negative sign in the right-hand

side has cancelled out with the negative sign coming from i^2 resulting from second derivative in the Fourier space. In the following discussion we assume atomic system of units, with $\epsilon_0 = \frac{1}{4\pi}$, $e = 1,\ \hbar = 1, m_e = 1$ for simplification of the expressions involved. It is important to note that due to the fact that the dielectric profile varies only in z direction, the convolutions in equation 5.4.6 are one-dimensional, along z direction only; this simplifies the problem substantially, since with such assumptions only the z components of the charge Fourier transform $\hat{\rho}_d(G_x, G_y, G_z{}')$ and potential Fourier transform $\hat{V}(G_x, G_y, G_z{}')$ are coupled to each other for every value of G_x and G_y – a much more tractable problem compared to full-blown three-dimensional convolution. In order to connect to the actual numerical description of the problem, we introduce the discretized version of those equations in the next section.

5.4.2. Discretization of the Poisson equation

In actual computational applications the quantities described above: charge density of the defect $\rho_d(\vec{r})$ and the corresponding potential $V(\vec{r})$ are represented on a discrete mesh of size (N_x, N_y, N_z), and corresponding mesh spacings $\Delta x = L_x/N_x$, with analogous expressions for y, z. According to the basic theory of discrete Fourier transforms, for such discrete mesh in real space the corresponding Fourier-space representation would be a mesh of wavevectors $\{\vec{G}^{ijk}\} = (G_x^i, G_y^j, G_z^k)$ with an equal number of points to the real-space one. The spacing on the wavevector mesh is related to the real-space mesh through:

$$\Delta G_x = \frac{2\pi}{L_x};\ \Delta G_y = \frac{2\pi}{L_y};\ \Delta G_z = \frac{2\pi}{L_z} \tag{5.4.7}$$

and the values for wavevectors are conventionally taken to be

$$\left\{G_x^i\right\} = -\frac{\pi(N_x - 1)}{L_x} \ldots \frac{\pi N_x}{L_x} \tag{5.4.8}$$

with similar expressions for G_y^j and G_z^k. Like the real-space cell under periodic boundary conditions, the wavevector mesh is also invariant to translations by a full lattice vector, $\frac{2\pi N_x}{L_x} = \frac{2\pi}{\Delta x}$, rendering the discrete Fourier transform periodic in the reciprocal space. The spacing in the real-space mesh, Δx, defines the minimal wavelength which can be resolved on such mesh through Nyquist condition, $\lambda_{min} = 2\Delta x$, since the wave has to be sampled at least twice on its wavelength to determine its phase. Correspondingly, the maximal wavevector for this wavelength is $G_{max}^x = \frac{2\pi}{\lambda_{min}} = \frac{\pi N_x}{L_x}$. Waves with wavevector components larger than that cannot be resolved on such grid and will in discrete representation look identical to a wave in which the corresponding component is smaller by G_{max}^x – hence the periodicity in wavevector space. Analogously, the smallest wavevector on the mesh, $G_x = \frac{2\pi}{L_x}$ corresponds to the largest possible wavelength which would fit the cell, $\lambda_{max} = L_x$. It should be noted that attempting to represent the waves with wavelengths incommensurate with the mesh lengths will result in artificial effects, when the discrete Fourier transform is blurred; however, since for our application of interest the quantities like ρ_d are enforced to be commensurate with the lattice by Bloch theorem, we do not encounter such effects in this work.

With the definition of the mesh in the Fourier space, the integral in the convolutions in equation 5.4.6 is reduced to a sum:

$$\int \hat{f}\left(\vec{G}' - \vec{G}\right)\hat{g}\left(\vec{G}'\right)d\vec{G}' \to \sum_{G'} \hat{f}\left(\vec{G}' - \vec{G}\right)\hat{g}\left(\vec{G}'\right) \tag{5.4.9}$$

and then the discretized form of the Poisson equation can be simplified as following:

$$\hat{\epsilon}(G_z) * |G|^2 \hat{V}(\vec{G}) + G_z \hat{\epsilon}(G_z) * G_z \hat{V}(\vec{G})$$

$$= \sum_{G_z'} \hat{\epsilon}(G_z - G_z') G_z'^2 \hat{V}(G_x, G_y, G_z')$$

$$+ \sum_{G_z'} \hat{\epsilon}(G_z - G_z')(G_x^2 + G_y^2) \hat{V}(G_x, G_y, G_z') \tag{5.4.10}$$

$$+ \sum_{G_z'} \hat{\epsilon}(G_z - G_z')(G_z - G_z') \hat{V}(G_x, G_y, G_z')\, G_z'$$

$$= \hat{\rho}_d(G_x, G_y, G_z)$$

After making some rearrangements, one can arrive at a simplified version:

$$\sum_{G_z'} \hat{\epsilon}(G_z - G_z')(G_x^2 + G_y^2 + G_z G_z') \hat{V}(G_x, G_y, G_z') = \hat{\rho}_d(G_x, G_y, G_z) \tag{5.4.11}$$

which in the discrete representation will read as:

$$\sum_l \epsilon_{k-l+1} \left(\left(G_x^i\right)^2 + \left(G_y^j\right)^2 + G_z^k G_z^l \right) V_{ijl} = \rho_{ijk} \tag{5.4.12}$$

where we have introduced shorthand notation: $\epsilon_{k-l} = \hat{\epsilon}(G_z^k - G_z^l)$, $V_{ijl} = \hat{V}(G_x^i, G_y^j, G_z^l)$, and $\rho_{ijk} = \hat{\rho}_d(G_x^i, G_y^j, G_z^l)$. It is clear that the presence of a non-trivial dielectric profile in z direction results in coupling between components of V_{ijl} and ρ_{ijk} for $k, l = 1 \ldots N_z$. Since the coefficients along x and y directions are not coupled, the problem is factorized into $N_x N_y$ systems of linear equations defined by matrices $\mathbf{M}^{ij}$ with matrix elements M_{kl}^{ij}

$$M_{kl}^{ij} = \epsilon_{k-l+1} \left(\left(G_x^i\right)^2 + \left(G_y^j\right)^2 + G_z^k G_z^l \right) \tag{5.4.13}$$

Such expression can be readily represented in terms of the circulant matrix. For a vector of Fourier components of the dielectric profile $\epsilon_k, k = 1 \ldots N_z$ the circulant is a $N_z * N_z$ matrix

$$\hat{\mathbf{C}}[\epsilon] = \begin{pmatrix} \epsilon_1 & \epsilon_{N_z} & \cdots & \epsilon_3 & \epsilon_2 \\ \epsilon_2 & \epsilon_1 & & \epsilon_4 & \epsilon_3 \\ \vdots & & \ddots & \vdots & \\ \epsilon_{N_z-1} & \epsilon_{N_z-2} & & \epsilon_1 & \epsilon_{N_z} \\ \epsilon_{N_z} & \epsilon_{N_z-1} & \cdots & \epsilon_2 & \epsilon_1 \end{pmatrix} \tag{5.4.14}$$

It is interesting to note that the condition number of the circulant matrix is the ratio of the largest Fourier transform component of the original vector generating the circulant to the smallest one. In our case, since the vector ϵ_k is a Fourier transform itself, the condition number of the circulant will be the ratio of largest value of real-space dielectric profile to the smallest one. Given the framework for the dielectric profile we use in this work, this ratio ends up being equal to the dielectric constant of the material in question – so there are no concerns in our particular case. However, this is helpful to keep in mind when developing the test cases for the code involving circulants.

With such definition, the matrix elements M_{kl}^{ij} of the matrices $\mathbf{M}^{ij}$ defining the problem in question can be expressed through circulant formed from the vector of Fourier components of the dielectric profile:

$$M_{kl}^{ij} = \hat{C}_{kl}[\epsilon]\left(\left(G_x^i\right)^2 + \left(G_y^j\right)^2\right) + \hat{C}_{kl}[\epsilon](G_z^k G_z^l) \tag{5.4.15}$$

It is worth noting that since i, j are parameters enumerating the wavevectors in uncoupled x, y dimensions, the first term in the sum is just the circulant multiplied by a real non-negative number. The second term in the sum is more involved – it is a Hadamard product of circulant with matrix $\mathbf{G}$ with matrix elements defined by $G_{kl} = G_z^k G_z^l$. That is, $\mathbf{G}$ is an outer product of the vector containing all allowed values of the z-component of the wavevector on a given mesh. In most modern software libraries, the enumeration of wavevectors inside the $\{G_x^i\}$ set is implemented in the following way: the first half of the set is the wavevectors from $G_x^1 = 0$ to

$G_x^{N_x/2+1} = \frac{\pi N_x}{L_x}$, and the second half of the set are the negative wavevectors in ascending order, from $G_x^{N_x/2+2} = -\frac{\pi(N_x-1)}{L_x}$ to $G_x^{N_x} = -\frac{\pi}{L_x}$. With that notation, the outer product matrix $\mathbf{G}$ has zero matrix elements along the first row and first column, having rank of $N_z - 1$. Therefore, the Hadamard product $\hat{\mathbf{C}}[\epsilon]\mathbf{G}$ also is rank-deficit. For this reason, for the case $i = 1, j = 1$, when the components $G_x^1 = 0, G_y^1 = 0$, so is the first term in equation 5.4.15, and the matrix $\mathbf{M}^{11}$ is rank-deficient. The component at the head of this matrix establishes the relation between the average value of the charge over the simulation cell, ρ_{111}, and the cell average of the electrostatic potential under periodic boundary conditions, V_{111} (there are no other elements involved, since first row of $\mathbf{M}^{11}$ is comprised of zero elements as well). This situation is due to the fact that the average value of electrostatic potential is undefined for charged systems. An analogue of this effect can be seen in the simple version of the Poisson equation, with constant dielectric profile, $\epsilon\Delta V = -\rho$. In the discrete representation the Laplacian-type matrix corresponding to the divergence operator Δ has full rank under open boundary conditions. Imposing periodic boundary conditions leads to addition of the elements in the "upper-right" and "lower-left" corners of the operator matrix, turning a Laplacian into a circulant with an incomplete rank and a one-dimensional kernel. This manifests itself in that the solution of the Poisson equation – electrostatic potential is determined only up to a constant. Given the problem of having an infinite array of charges, the customary choice made is setting the average value of electrostatic potential to zero. By the same logic, any value can be set for the matrix element M_{11}^{11} – it does not affect the shape of the resulting potential, and only sets the average value. This, however, does not affect the final result, since the average value of the electrostatic potential is set to zero after carrying out the computation. Therefore, in practice the matrices $\mathbf{M}^{ij}$ have full rank, and the system of linear equations can always be solved correctly. For the condition numbers of these

matrices, we find the distribution from ~1 to ~1000, and the condition numbers tend to increase with increasing the finesse of the real-space mesh due to the presence of $\mathbf{G}$ term. We find that for typically used values of the mesh spacing in our situation – around 0.2 Bohr – the condition numbers rarely exceed 1000, and the solution of equation (5.4.15) is numerically stable using the standard LU factorization method. For more sensitive cases, more stable linear system solvers based on the singular value decomposition (SVD) algorithm can be used.

The scheme described here can be easily extended to the case of the host material with anisotropic dielectric tensor, when instead of one dielectric profile $\epsilon(z)$ the problem will have three profiles corresponding to the components of the dielectric tensor, $\{\epsilon_{xx}(z), \epsilon_{yy}(z), \epsilon_{zz}(z)\}$. In this case the expression for the Poisson equation (5.4.10) will take the following form:

$$\begin{aligned}
&\{\hat{\epsilon}(G_z)\} * |G|^2 \hat{V}(\vec{G}) + G_z \hat{\epsilon}_{zz}(G_z) * G_z \hat{V}(\vec{G}) \\
&\quad = \sum_{G_z'} \hat{\epsilon}_{zz}(G_z - G_z') G_z'^2 \hat{V}(G_x, G_y, G_z') \\
&\quad + \sum_{G_z'} \hat{\epsilon}_{xx}(G_z - G_z') G_x^2 \hat{V}(G_x, G_y, G_z') \\
&\quad + \sum_{G_z'} \hat{\epsilon}_{yy}(G_z - G_z') G_y^2 \hat{V}(G_x, G_y, G_z') \\
&\quad + \sum_{G_z'} \hat{\epsilon}_{zz}(G_z - G_z')(G_z - G_z') \hat{V}(G_x, G_y, G_z') G_z' \\
&\quad = \hat{\rho}_d(G_x, G_y, G_z)
\end{aligned} \tag{5.4.16}$$

Correspondingly, after discretization the expressions for matrices $\mathbf{M}^{ij}$ can be written in terms of circulant matrices $\hat{\mathbf{C}}[\epsilon_{xx}], \hat{\mathbf{C}}[\epsilon_{yy}], \hat{\mathbf{C}}[\epsilon_{zz}]$ generated from the discrete Fourier transforms of $\{\epsilon_{xx}(z), \epsilon_{yy}(z), \epsilon_{zz}(z)\}$, respectively:

$$M_{kl}^{ij} = \hat{C}_{kl}[\epsilon_{xx}]\left(G_x^i\right)^2 + \hat{C}_{kl}[\epsilon_{yy}]\left(G_y^j\right)^2 + \hat{C}_{kl}[\epsilon_{zz}](G_z^k G_z^l) \quad (5.4.17)$$

The same considerations regarding the condition numbers apply: the head element of the first matrix, M_{11}^{11} is set to arbitrary positive value, e.g. 1, and in the Fourier transform of the electrostatic potential obtained after solving all linear systems the head component V_{111}, corresponding to the cell-averaged value of the electrostatic potential, is set to zero. We turn now to the question of implementation of the described scheme and questions of computational efficiency.

5.4.3. Parallel implementation of the Poisson equation solver

As we have discussed before, for the case of discretized representation on a grid of size (N_x, N_y, N_z) the Poisson equation (5.4.6) reduces to $N_x \times N_y$ systems of linear equations defined by matrices $\mathbf{M}^{ij}$. Such structure of the problem naturally lends itself to parallel implementation. We have implemented the computation code in Python language, version 2.7; for both parallel and serial versions. For the parallelization we have used the message passing interface (MPI) protocol, as implemented in the mpi4py[66] library. In the implementation, we have sought to achieve parallelization seamless for the user, so that the code can be run with minimal changes on both desktop and laptop machines with only standard Python interpreters, as well as large supercomputing facilities with the mpi4py library installed. We first describe the computational algorithm we have used in the serial case and its computational complexity, and then move on to parallel implementation and its efficiency.

The procedure for computation is as following:

1. Read in from file the charge density $\rho_d(\vec{r}_{ijk})$ on a discretized grid – this is typically a file generated by a DFT computational package. Alternatively, for modeling and testing purposes, a Gaussian charge distribution of specified width and amplitude can be

generated. Obtain the dielectric profile parameters – positions of the material-vacuum boundaries, material dielectric constant, width of the smoothing error function.

2. Carry out the Fourier transform of the charge field to obtain its reciprocal-space representation $\hat{\rho}_d(\vec{G}_{ijk}) = \rho_{ijk}$. The Fourier transform is carried out using the fast Fourier transform (FFT) algorithm, as implemented in SciPy library.[67]

3. For each $i = 1 \dots N_x, j = 1 \dots N_y$ form the corresponding matrix $\mathbf{M}^{ij}$, equation 5.4.15; for $i, j = 1$ set the head component $M_{11}^{11} = 1.0$. Then solve the linear systems for all i, j in order to obtain the Fourier components of the potential, V_{ijk}. Solution of the linear systems can be carried out using either LU factorization or singular value decomposition (SVD); both of those algorithms are implemented in the SciPy library.[67] We find that LU algorithm is sufficiently stable for the case at hand, and use it in the following discussion.

4. Set the cell average value of the potential to zero, $V_{111} = 0$, carry out inverse Fourier transform.

5. Compute the electrostatic energy for the model under periodic boundary conditions, according to equation 5.4.2.

There are two major components of the problem that require non-trivial computational effort: the fast Fourier transforms and the solution of linear systems $\mathbf{M}^{ij}$. The algorithm most often used these days for Fourier transforms – the fast Fourier transform, has computational complexity of $O(N \ln N)$ for 1D transform; the 3D Fourier transform amounts to carrying out $N_x \times N_y$ 1D FFT's in z dimension with computational intensity $O(N_x N_y N_z \ln N_z)$, followed by $N_x \times N_z$ 1D FFT's in y dimension with computational intensity $O(N_x N_z N_y \ln N_y)$, and $N_z \times N_y$ 1D FFT's in x dimension with computational intensity $O(N_x N_y N_z \ln N_z)$. Overall the complexity of 3D FFT adds up to $O(N_x N_y N_z \ln(N_x N_y N_z))$. The solution of linear systems using

the LU factorization has the same computational complexity as matrix multiplication, $O(N^{2.8})$, where the reduction from normally-expected power of 3 to 2.8 is achieved by using Strassen algorithm.[68] The savings in the algorithm come from representation of the problem at hand as a series of 2x2 matrix multiplications, which can be done in 7 operations instead of 8. Since there are $N_x \times N_y$ linear systems to solve of size $N_z \times N_z$, the overall complexity of the linear systems solution is $O\left(N_x N_y N_z^{2.8}\right)$.

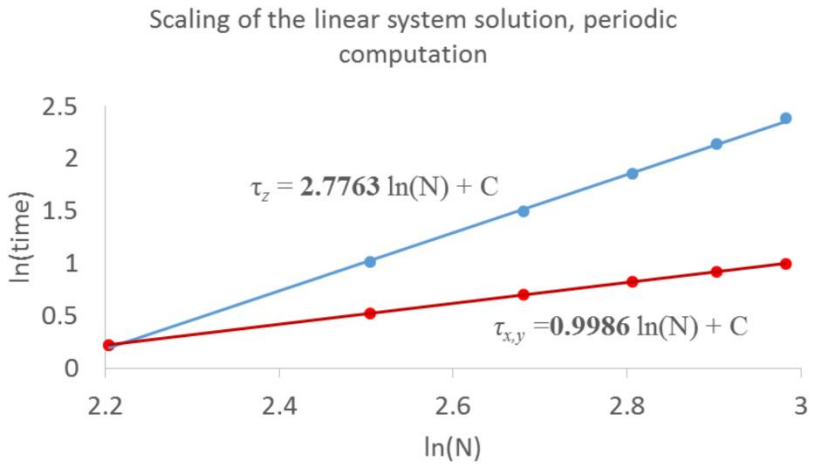

Figure 5-2. Computational complexity of solving the linear systems $\mathbf{M}^{ij}$.
Linear scaling with the size of the mesh in x and y direction is shown with red curve, and complexity of linear systems solution is shown with blue curve. The slope of the curve for scaling the number of points in z dimension is 2.8, as expected from using Strassen algorithm for matrix multiplication.

We see that for a typical problem size of interest, which is several hundred mesh points in each direction, the linear system part of the problem is the major component of the workload. We test the scaling on the code for the case of a Gaussian charge with width 1 Bohr in a simulation cell of $10 \times 10 \times 10$ Bohr as a function of mesh size, as shown in Figure 5-2 using the double logarithmic coordinates. In such coordinates, $\tau = CN_x N_y N_z^{2.8}$ turns into $\ln(\tau) = \ln N_x + \ln N_y +$

$2.8 \ln N_z + C_1$. We demonstrate the dependence of computational time on each of those parameters changes separately. For reference, for a mesh size of $N_x = N_y = N_z = 100$, the time for FFT is about 1 second, whereas the linear system time is about 70 seconds, and with increasing the system size the difference becomes even more pronounced.

Given the considerations of computational complexity of the parts of the problem discussed above, as well as the goal of having uniform design of the code for serial and parallel execution, we adopt the following approach for parallel implementation: we treat all the MPI processes on equal footing, and we impose the invariant of all MPI processes having all the main arrays involved in the computation: input charge $\rho_d(\vec{r}_{ijk})$, its Fourier transform $\hat{\rho}_d(\vec{G}_{ijk})$, Fourier transform of the potential $V_{ijk} \equiv \hat{V}(\vec{G}_{ijk})$ obtained from solving linear systems (equation 5.4.15), and the resulting real-space potential $V(\vec{r}_{ijk})$. In order to achieve that, the input is read in by each process independently, then Fourier transform is carried out serially each process. We choose to not parallelize the FFT routine since in our experience the time for doing the computation is comparable or even less than the time for doing the MPI collective communication, such as Scatterv. This situation holds for typical system sizes of up to ~500 points in each direction, which is about the size of the systems which we deal with (indeed, in our experience with plane-wave codes, such as QuantumEspresso,[33] where the FFT is implemented in parallel, the time for Fourier transforms is dominated by the MPI collective communication calls). Then, after each process carries out the FFT locally, the workload for solving linear systems is shared between processes. The right-hand side of equation 5.4.15, FT of the potential $\hat{\rho}_d(\vec{G}_{ijk})$ is partitioned along the x axis; each of the processes gets assigned a slice of the array with full size in y, z dimensions and reduced size in x. Then each process solves the $N_x^{loc} \times N_y$ linear systems to get its local fragment of the potential Fourier transform, $\hat{V}_{loc}(\vec{G}_{ijk})$.

The full array with potential FT is created on each process after a collective communication MPI call, AllGatherv (the vectorized version of the call is used to accommodate the case of the mesh size in x dimension not divisible by the number of processes). After that, the inverse FFT is carried out locally on each process, and as the result every process has the full solution of the Poisson equation, $V(\vec{r}_{ijk})$, in memory. The potential bottleneck for such scheme could be the amount of memory available to each process, since it has to be sufficient to hold the entire mesh. However, for practically used mesh sizes the amount of 2-4 GB per process is sufficient, so we have not experienced any problems with memory while running the code on modern computational clusters. We also note that in the case of serial execution the MPI calls are not executed, however exactly the same computational kernel used for solving part of the problem's linear systems is called, but with the entire right-hand side, and all $N_x \times N_y$ linear systems are solved locally.

With such scheme, the program consists of two parts: serial, with Fourier transforms, and parallel – the solution of linear systems and MPI calls. As such, the efficiency of parallelization over number of processors n_{proc} compared to serial execution can be expressed in terms of speedup $S(n_{proc})$ determined by Amdahl's law:

$$S(n_{proc}) = \frac{\tau_s + \tau_p}{\tau_s + \tau_p(n_{proc})} \tag{5.4.18}$$

where τ_s is the execution time of the serial part of the program (FFTs), τ_p is the execution time of the parallelizable part of the problem (linear systems) when running in serial, and $\tau_p(n_{proc})$ is the time spent on executing the parallel part of the program when running on n_{proc} (ideally, it should be τ_p/n_{proc}). We test the performance of the code to ideal behavior for up to 24 processors (about the number one would use for a real simulation), as shown in Figure 5-3. The

code scaling closely follows the limit imposed by Amdahl's law. We should note that even for small system sizes the parallel part is two orders of magnitude more time-consuming than the serial, and the difference increases fast with system size – so the code can be parallelized to up to hundreds of processors.

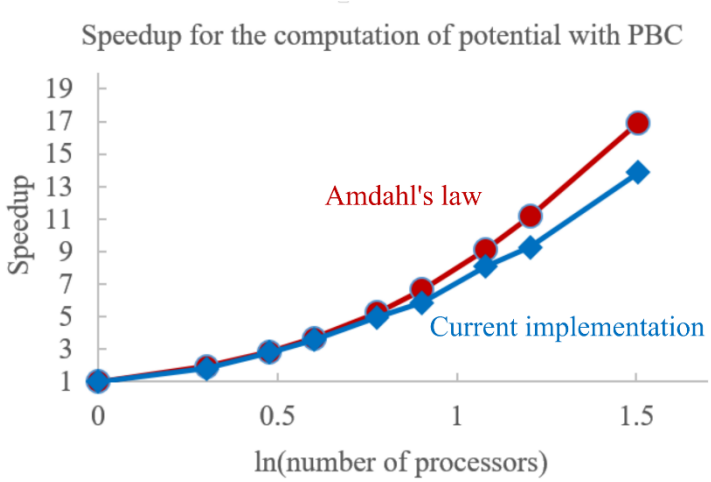

Figure 5-3. Comparison of the current implementation with Amdahl's law (see discussion in the text).

5.4.4. Sensitivity of the computational scheme

For a given size of the simulation supercell (determined by the setup of the DFT computation), the key technical parameter determining the quality of the results is the mesh spacing. The main output of the simulation is the electrostatic energy obtained from the potential solved for above. We test the convergence of energy for the case of Gaussian charge distribution of width $\sigma = 1$ Bohr for the cases of charge in vacuum, charge at the boundary of vacuum and dielectric material with dielectric constant $\epsilon = 5.0$, and in the material, 6 Bohr away from the boundary. We find that convergence of electrostatic energy to within 0.03 eV is achieved for mesh spacing $\Delta x = \sigma/2$, and the values of electrostatic potential converge to within 0.01 eV for mesh spacing of $\Delta x = \sigma/5$. It should be noted that for the charge distributions encountered in

real computations the characteristic size is over 2 Bohr, so the mesh spacing sufficient for satisfactory code performance is about 0.3 – 0.4 Bohr, which is lower than the values resulting from the DFT calculations well-converged with respect to that parameter. This fact enables downsampling the output of the first-principle codes by a factor of 2 to 3 in order to save the computational effort.

We next consider the sensitivity of the electrostatic energy for a model Gaussian on the boundary of a dielectric material and vacuum (with charge position offset by 4 Bohr into the material) – the arrangement presented on Figure 5-4. The model is supposed to emulate a NaCl slab with thickness 21.2 Bohr (twice the lattice parameter) with a Cl vacancy-induced charge near the upper boundary.

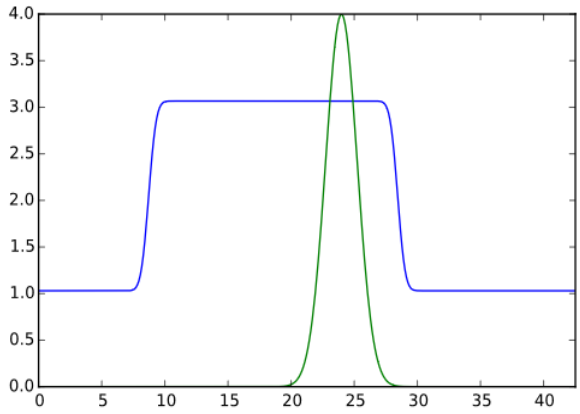

Figure 5-4. Model setup for testing sensitivity to parameter variation.

The model is a mesh with size $10.6 \times 10.6 \times 42.4$ Bohr. The dielectric constant $\epsilon_{mat} = 3.1$, with the vacuum and material joined by error function with width parameter 0.7 Bohr. Gaussian charge of width 2 Bohr is situated 4 Bohr away from the boundary (green line is the xy-average of the model charge).

The key parameters in this model are: 1. separation between the center of the Gaussian charge and the closest material-vacuum boundary, 2. value of the dielectric constant of the model profile in the material region, 3. width of the error functions joining the material and vacuum region of the dielectric profile. We seek to understand by how much those parameters need to be

changed in order to change the value of the electrostatic energy by 0.1 eV – typical accuracy of the modern plane-wave DFT codes for simulation cells larger than 100 atoms. We find that the most sensitive parameter is the distance between the Gaussian charge center and the closest material boundary – changes of 0.15 Bohr in this value change the electrostatic energy substantially. Next most important parameter is the value of dielectric constant – for material dielectric constant $\epsilon_{mat} = 3.1$ change of the constant by +/-0.3 is sufficient to change the model energy by 0.1 eV. We note, however, that such low value of ϵ_{mat} is an extreme case, and the values of dielectric constant typical for real materials (10 and higher) have higher tolerance to errors. The least important parameter is the width of the error function – we find that one need to increase it by a factor of over 2, up to 2.0 Bohr, compared to the original value of 0.7 Bohr, in order to change the electrostatic energy substantially. Such results demonstrate the dangers associated with using Gaussian fits for modeling the real charge distributions – we will get to that point once again when discussing our approach to the problem in the next section.

5.5. Computation of electrostatic energy for an isolated charge

The electrostatic potential for isolated charge is also governed by the Poisson equation (5.4.1) with the boundary conditions for the potential to decay to zero at infinity, $\lim_{\vec{r}\to\infty} V(\vec{r}) = 0$. This makes the explicit solution of Poisson equation by the discretization of Laplacian operator on a grid not tractable. Instead, a substantial number of modern approaches to electrostatic problem under open boundary conditions, like the fast multipole method (FMM)[69] are based on direct summation of potential induced by discretized charge elements, with some techniques utilized for improving efficiency.[70] In our case, however, inhomogeneous dielectric profile complicates the problem, so we have to resort to direct summation technique for potential computation, as described in the following section.

5.5.1. Implementation of direct summation

The approach we implement here is based on the image charge method. The key idea is that for a discrete representation of the defect charge on the boundary of two dielectric media the potential induced by the point charge elements on both sides of the dielectric boundary can be calculated analytically.[71] The situation is illustrated in Figure 5-5, which shows a schematic view of the slice of the charge distribution along xz plane. The thick black line denotes the boundary between two media with dielectric constants of ϵ_1 and ϵ_2, respectively. First of all, we note that for the potential computation an auxiliary grid is introduced, since the $1/r$ Coulomb potential is singular. We set this auxiliary grid to be shifted by a vector $\left(\frac{\Delta x}{2}, \frac{\Delta y}{2}, \frac{\Delta z}{2}\right)$ compared to the charge mesh, and compute the potential values on it. The relative position of charge and mesh grids is shown in Figure 5-5, with charge points denoted by white dots and potential mesh points are red. For each charge point the contributions to all points on the potential grid are computed; there are two types of potential expression depending on the positions of charge and potential mesh points relative to the interface. For points on the same side of the interface, the potential is induced by the charge itself: $1/\epsilon_1 r_1$, with the dielectric constant corresponding to the material in that part of the simulation domain. Another contribution is so-called "image" charge, which induces potential with effective dielectric constant $\frac{\epsilon_1 - \epsilon_2}{\epsilon_1 + \epsilon_2}$. In the limit of charge in vacuum near the metal surface, the effective dielectric constant is -1, which corresponds to well-known limit of image charge of the equal magnitude and the opposite sign. The lateral positions of the image charge are the same as of the original charge element, and the z coordinate is obtained by applying a mirror reflection operation in the plane separating the media. For the points on the opposite side

of the interface, only the original charge element has contribution with effective dielectric constant $\frac{2}{\epsilon_1+\epsilon_2}$.

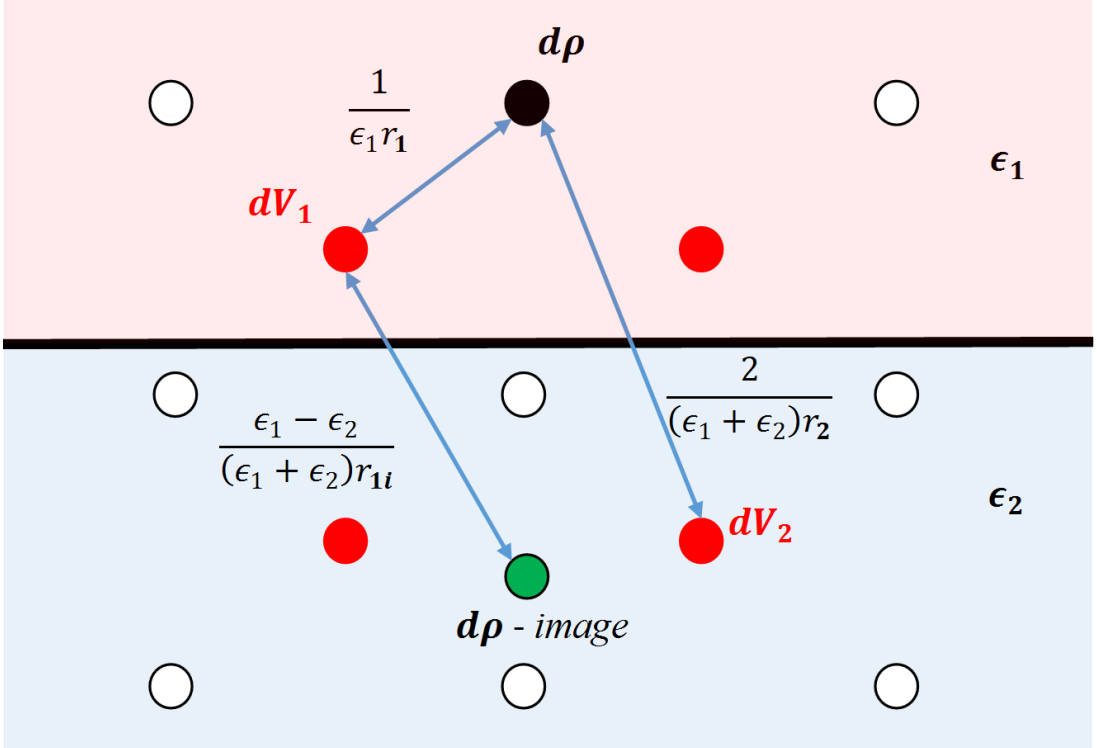

Figure 5-5. Illustration of computation under open boundary conditions.

The points on the mesh for discretizing charge are shown as hollow white circles, and points of the potential mesh are shown as red circles. Computation of the potential induced by a given charge point is demonstrated on an example of a point shown by black circle (and the corresponding image charge is shown by a green circle). For the potential points on the same side, both direct and image terms contribute, whereas for points on the opposite side of the boundary only a direct term is computed.

After obtaining the potential values on the offset grid the values of the potential are interpolated back on the original charge mesh. From the perspective of computational intensity, this method is very demanding: $O\left(N_x^2 N_y^2 N_z^2\right)$, and a much smaller contribution for interpolation. However, it lends itself naturally to parallelization, since the interpolation time is about 3 orders

of magnitude smaller than the potential computation time. In parallelization, we follow the same strategy as described for the periodic computation: availability of important arrays (charge $\rho(\vec{r}_{ijk})$, offset potential $V^o(\vec{r}_{ijk})$, and potential interpolated back to the original grid $V(\vec{r}_{ijk})$) on all processes and unification of interface for serial and parallel implementation.

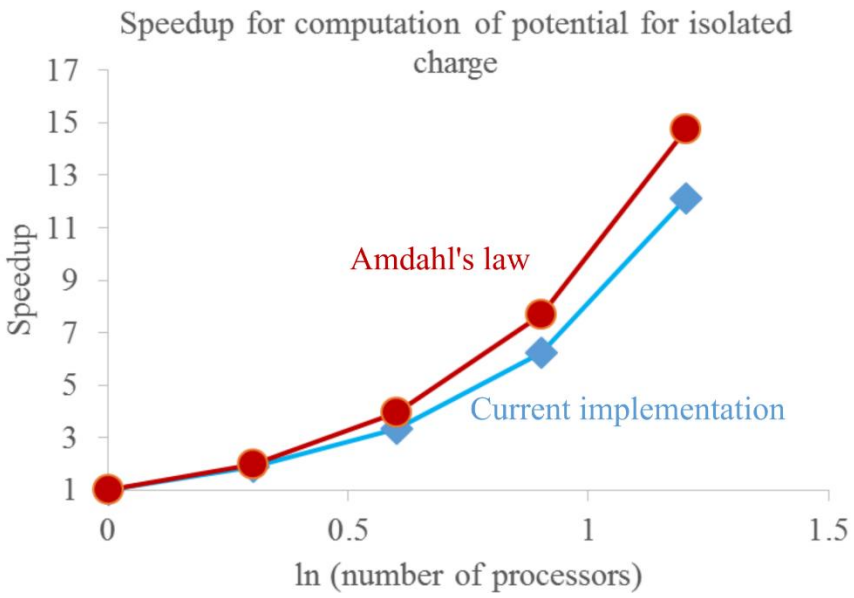

Figure 5-6. Comparison of the current implementation of the computation of the potential for isolated charge with the Amdahl's law (see discussion in the text).

With that in mind, we implement the computation in the following way: first, the charge distribution is read in by all processes. Then, the size of the offset mesh $V^o(\vec{r}_{ijk})$ is evaluated, and it is partitioned among the processes along x direction. After that, each process uses its locally available full charge distribution for calculation of the local slice of the offset potential mesh; the direct summation can be easily vectorized using the Boolean arrays implemented in NumPy. Finally, the local parts of $V^o_{loc}(\vec{r}_{ijk})$ are assembled into full array on each process using the AllGatherv MPI command, and then the interpolation to the original charge grid is done

locally on each machine. The implemented code shows favorable scaling, closely following the idealized Amdahl's law limit (Figure 5-6).

The electrostatic energy for computation described above converges to within 0.05 eV when the mesh spacing is half of the Gaussian width, $\Delta x = \sigma/2$. We hypothesize that the fast convergence is due to the fact that energy computation according to equation 5.4.2 is a numerical integration, which is an error-cancelling procedure (as opposed to error-amplifying numerical derivative). We do not explicitly use the potential for isolated charge in this work, so we do not test its convergence.

Having implemented computation of both periodic and open boundary conditions potential, we can evaluate the importance of the implemented correction and demonstrate its correctness. We consider the case of a Gaussian charge with width 1 Bohr in a simulation cell of 10x10x20 Bohr near the interface of two dielectric media: vacuum and material with dielectric constant 4.0. We then compute the energy with periodic boundary conditions and open boundary conditions, according to the methods described above, upon varying the position of the charge center relative to the boundary. The results are presented in Figure 5-7: the region modeling material is on the left, with negative distance values, and vacuum region is on the right. When the charge position is far from the interface, its energy can be computed by adding up two components: first is the electrostatic self-energy of the Gaussian charge distribution:

$$E_G = \frac{q^2}{2\sigma\epsilon\sqrt{\pi}} \tag{5.5.1}$$

where $q = \int \rho(\vec{r})d\vec{r}$ is the total charge of the distribution, σ is the width of the distribution, and ϵ is the dielectric constant of the medium in which the charge is immersed. Second is the energy

of interaction of the Gaussian with the image charge generated by it, treating the model Gaussian as a point charge $E_{img} = \frac{q^2(\epsilon_1-\epsilon_2)}{2(\epsilon_1+\epsilon_2)\epsilon_1 d}$, where d is the distance between the charge and its image.

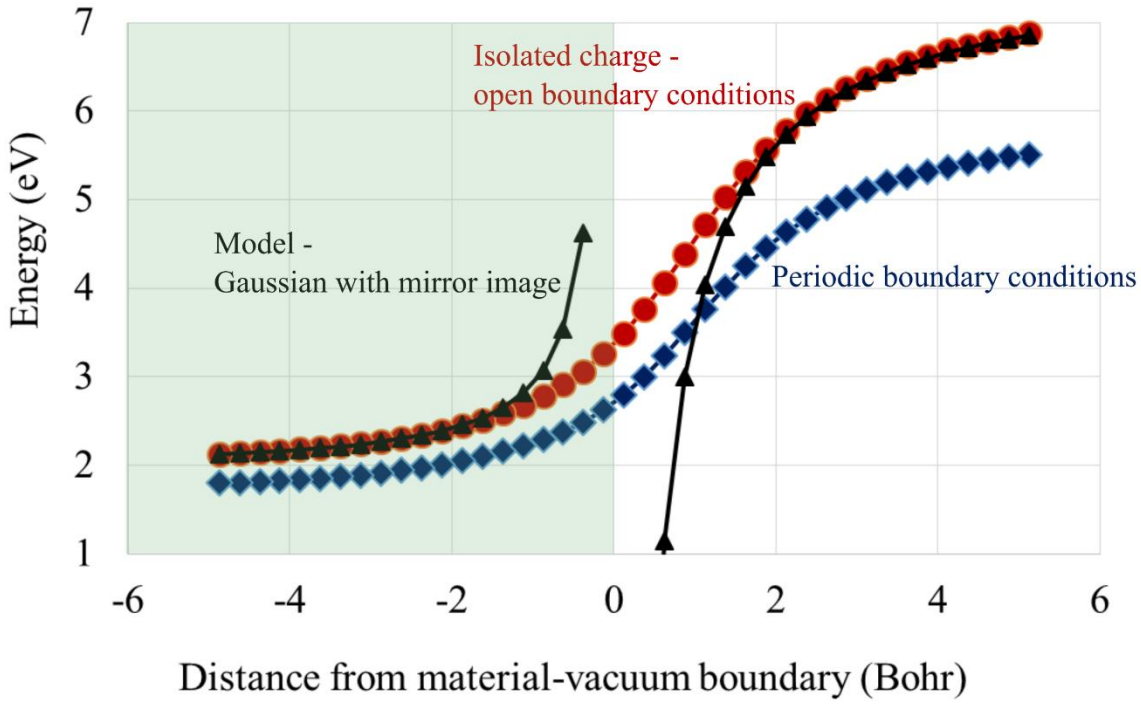

Figure 5-7. Electrostatic energy of Gaussian charge distribution as a function of its position relative to the dielectric media boundary.

Dielectric material with $\epsilon = 4.0$ is on the left, vacuum region – on the right. Red curve denotes the energy computed under open boundary conditions for isolated charge; black – self-energy of a Gaussian in the corresponding medium with energy of interaction with its mage charge; blue – energy of the charge distribution under periodic boundary conditions – this would be the electrostatic component of the total energy from charged defect DFT computation.

This energy is plotted in black curve in Figure 5-7; it demonstrates the attractive interaction of charge with its mirror image in the higher-dielectric constant medium, and repulsive interaction with the image in low-dielectric constant medium. However, this model energy diverges as the charge approaches the interface. On the contrary, energy values computed with the method for

open boundary conditions (red curve), smoothly transition from one limit to another; analogously, the energy under periodic boundary conditions is continuous across the interface. It is important to note that the difference between the periodic boundary conditions energy and open boundary conditions energy can reach more than 1.5 eV in the vacuum region, and about 0.4 eV deep in the material, with about 1 eV on the boundary. This illustrates the magnitude of potential errors in the computation of defect formation energies. It should be noted, however, that the energy difference will decrease with increasing the dielectric constant, increasing width of the charge distribution and increasing the size of the simulation cell – so in real simulations the energy difference might not be as dramatic, but still can be a source of substantial random error.

5.5.2. Computation of energy for open boundary conditions by extrapolation

The approach to computation of the electrostatic energy for isolated charge described above has two major drawbacks: first is the computational intensity – the method scales as the 6^{th} power of mesh linear size; second is the simplicity of the model for dielectric media boundary – a single plane is a very crude approximation to the actual dielectric interface, and in reality the dielectric profile is actually a smooth function changing on the scale comparable to the typical scale of the defect charge distribution itself. Moreover, such boundary model accommodates only one interface, thereby effectively excluding 2D materials from consideration. Therefore, we investigate another approach to computing the electrostatic energy under open boundary conditions through extrapolation of the periodic boundary conditions energy to infinite cell size. This method is inspired by the "scaling relationships" discussed in earlier methodology work,[59] where it was shown that the error in electrostatic energy scales as inverse of the supercell size.

Such method was mentioned in the literature before,[65] but it has several important caveats which we discuss here for the first time. While explicit construction of infinite model supercell is not feasible, the isolated energy can be recovered by carrying out a series of model electrostatic calculations for increasingly large supercells scaled by a factor α compared to the original size, and then fitting the resulting energies to a straight line as a function of $1/\alpha$; the limit of $1/\alpha \to 0$ is the electrostatic energy of an isolated charge.

It is apparent that in the extrapolation method the key variable is the maximal extrapolation scale, since it determines how close does the straight line in energy vs $1/\alpha$ coordinates get to 0 corresponding to the isolated limit. In Figure 5-8 we demonstrate the situation for a Gaussian charge of width 1.0 Bohr in vacuum in a cell of 12x12x12 Bohr. The extrapolation procedure is carried out by computing the electrostatic energies for that charge for a number of scales up to 7 (84x84x84 Bohr), and fitting the resulting energies to a straight line. The result matches closely the true electrostatic self-energy of isolated Gaussian charge distribution in vacuum $E_{Gauss} = q^2/2\epsilon\sigma\sqrt{\pi} = 0.2821$ Ha $= 7.672$ eV. We investigate the error in the extrapolation as a function of the maximal scale. The results for Gaussian charge in vacuum demonstrate that for maximal scale of 3 times the original cell size the error is about 0.05 eV, scale 5 – 0.03 eV and scale 7 – 0.02 eV. Bearing in mind that the computational workload scales as linear dimension of the cell to the power of 4.8, the additional accuracy comes at a very steep resource requirement.

One should keep in mind that the electrostatic energies and, correspondingly, differences between them scale inversely with the dielectric constant of the system, so the calculations in vacuum represent an upper bound on the error estimates in our case. Therefore, in practice it should be sufficient to set the maximal scale to 4 or 5, depending on the available resources and

the size of the supercell in question. It is also important to note that the electrostatic correction decreases with the increase in the size of the original supercell, so for larger systems the error is dominated by random errors in the $E(\alpha)$ values, which depend on the sampling of the defect charge state, especially for anisotropic wavefunctions and reach ~0.03 eV from our experience.

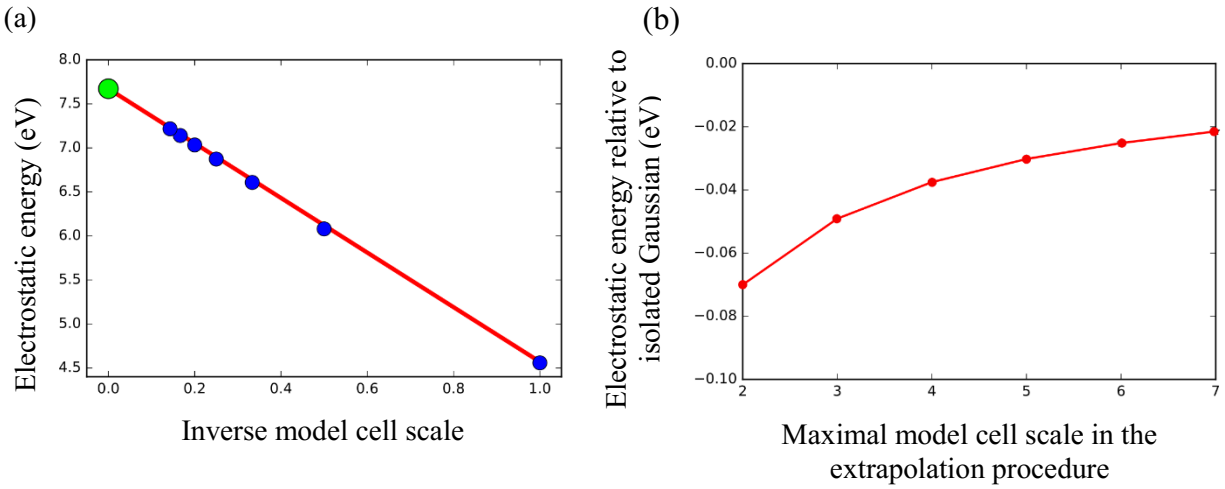

Figure 5-8. Extrapolation of periodic boundary conditions electrostatic energy.

(a) Blue circles – energies for Gaussian charge in vacuum for relative cell sizes 1 to 7 (compared to original $12 \times 12 \times 12$ Bohr, see text); red line – linear least squares fit to inverse dependence of energy on inverse supercell size; green dot – energy for isolated Gaussian charge in vacuum.

(b) Difference between energy of isolated charge and the approximate value obtained from extrapolation to maximal scale denoted on x axis. See text for discussion.

Another important component of the problem is the initial shape of the cell containing the charge. We have found that even for a Gaussian charge in vacuum any deviation of the original cell shape from cubic will result in random errors which are very large (up to 5 eV for starting shape of $24 \times 24 \times 12$ Bohr, Figure 5-9). This is a critical point never present in the discussion of the extrapolation to date. Only with cubic shape of the original supercell does the extrapolated energy converge to the proper limit – one can see that on the figure below for initial cell sizes of

12, 18, and 24 Bohr. This has pretty important implications for simulation of real materials, as we will show later.

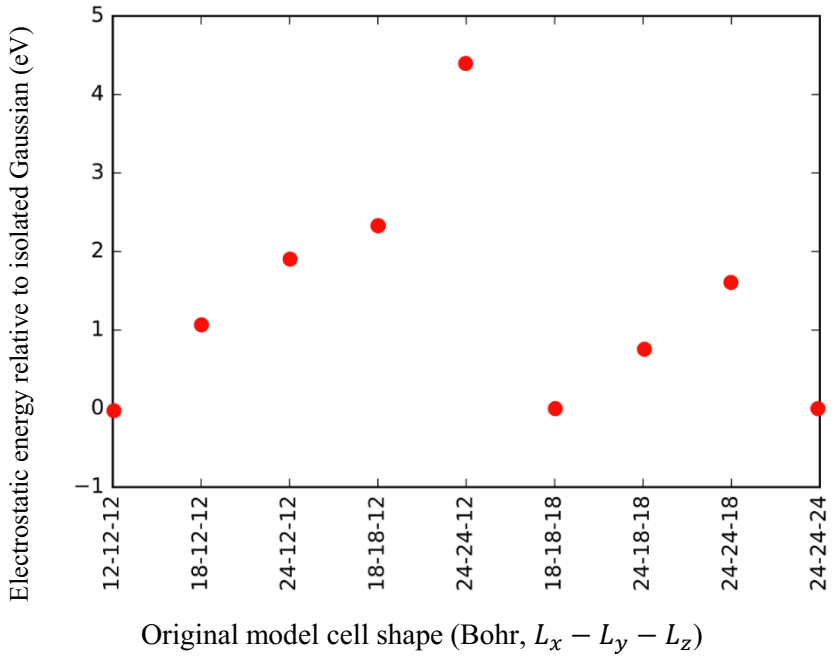

Figure 5-9. Dependence of the extrapolated energy of Gaussian charge in vacuum on the model cell shape.

Only cubic shape of the original cell results in correct extrapolated energy. The same condition holds for other cases, like charge on the boundary of dielectric media.

The next aspect of the problem is the question of applicability of such approach to computation of electrostatic energy for materials in real dielectric environments. Additional complication here is the ambiguity in relative position of the charge and the dielectric profile boundaries. This situation is demonstrated in Figure 5-10(a) below for the case imitating the charge distribution resulting from Cl vacancy in the NaCl slab. The outer boundary of the dielectric profile extends from the nuclei positions into the vacuum by the distance roughly equal to atomic radius of the comprising atoms (in this case we use the offset equal to 2.5 Bohr),

whereas the charge distribution associated with the defect is typically centered around the former position of the atomic nucleus of the removed atom. In our test we've shown that the relative positions of the charge center and dielectric interface boundary have substantial impact on the resulting electrostatic energy. Therefore, upon scaling the system, we pad the mesh with the charge distribution with zeros on all sides, placing the old charge in the center of the scaled cell.

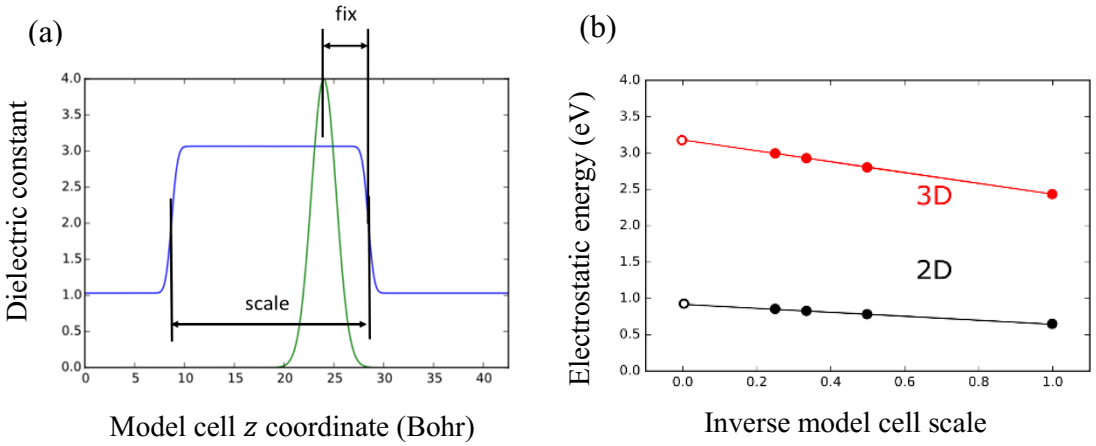

Figure 5-10. Extrapolation with nontrivial dielectric profiles.

(a) Model of a charge at the interface of material slab. Upon extrapolation, the distance to the nearest interface is kept fixed, whereas the width of model dielectric profile is increased in proportion to the scaling factor used. For simulation of dielectric profiles of 2D materials the width of the material is fixed as well.

(b) Extrapolation of periodic energies for cases of scaling the model material thickness ("3D") and keeping it constant ("2D"). 3D case is the model described above, 2D is Gaussian charge at the interface of a layer of dielctric with 8 Bohr thickness, in-plane component of dielectric constant $\epsilon_{\parallel} = 15.0$, out-of-plane $\epsilon_{\perp} = 6.0$.

For the case of semiconductor surface regions, we calculate the position of the dielectric profile boundary closest to the charge by setting the offset to be the same as in the original cell. The second boundary position is calculated by requiring the thickness of the material to increase

by the same amount as the supercell size. For the case of 2D materials we keep the material thickness the same. In both cases the extrapolation results in a linear dependence of energy of the scaled supercell on the inverse scale, which allows extrapolation to infinite supercell size.

5.6. Computation of the model dielectric profile

So far, we were assuming that a dielectric profile is given; in this section we will consider available methods for its computation. They are typically based on investigating the response of the electronic density of the model supercell to applied electric field. Since the applied external electric potential also has to obey periodic boundary conditions, then along with the constant electric field in part of the supercell with atom there has to be a part with oppositely directed electric field, such that where the potential goes back to the original value. Overall, this corresponds to so-called "saw-tooth" potential shape, implemented in many DFT codes.[33] There are several ways to extract the dielectric response of the system. One of them is based on analysis of the change in exchange-correlation potential with application of electric field;[72] the change in the self-consistent Hartree potential in the model cell upon application of electric field is calculated, and then the resulting slope of this difference term is divided by the slope of the sawtooth potential:

$$\epsilon(z) = \frac{\overline{V^H}_{el}(z) - \overline{V^H}_{host}(z)}{\partial V_{saw}(z)/\partial z} \tag{5.6.1}$$

where $\overline{V^H}_{host}(z)$ is the xy-averaged value of the host cell Hartree potential without applied electric field, $\overline{V^H}_{el}(z)$ is the same quantity after applying the sawtooth potential $V_{saw}(z)$. We find however, that this method is ripe with noise-amplifying procedures: subtraction, numerical derivative, and division; as a result, the dielectric profile is very noisy.

Therefore, we use another approach which takes advantage of classical electrostatics.[73] It is known that the electric field in the dielectric materials causes a change in electronic density. In the following discussion we assume that the electric field is applied in the z direction, and we consider xy -plane averages of the necessary quantities. So the electric field $\vec{E} = (0, 0, E_z)$ results in density displacement in z direction, $\Delta\rho(z)$. By definition, this density displacement is connected to a quantity called polarization, $\vec{P}$, through continuity-type equation, $\nabla\vec{P} = \Delta\rho(z)$. Again, due to applied electric field being in the z direction, the polarization vector is pointing the same way, and the divergence turns into derivative with respect to z: $\frac{dP_z}{dz} = \Delta\rho(z)$. Again, by definition, there is no polarization in the vacuum region at z=0, so with that boundary condition the polarization can be found explicitly by integration: $P_z(z) = \int_0^z \Delta\rho(z')dz'$. NB that without the vacuum layer this would lead to uncertainty in polarization definition, so the dielectric response in bulk semiconductors is one of the core problems in modern condensed-matter physics. Next, we know that in case of dielectric materials the conserved quantity upon traversing borders is the perpendicular component of the electric displacement vector, $D_\perp^1 = D_\perp^2$. The electric displacement vector in the dielectric material is given by $D(z) = E(z) + 4\pi P(z)$, and from continuity equation we know that throughout the dielectric layer the displacement should be equal to the electric field in vacuum: $E_{vac}(z) = E_0 = D_{vac} = D_{diel}(z) = E_{diel}(z) + 4\pi P(z)$. From the above, knowing the polarization expression, we can get the electric field inside the material, $E_{diel}(z)$. Finally, it is well-known that the definition of electric susceptibility is the ratio of polarization to electric field in the dielectric: $\chi(z) = 4\pi P(z)/E(z)$, and the dielectric profile is trivially related to susceptibility: $\epsilon(z) = 1 + \chi(z)$. The final expression for the xy-averaged value of the dielectric profile is:

$$\epsilon(z) = 1 + \frac{4\pi \int_0^z \Delta\rho(z')dz'}{E_0 - \int_0^z \Delta\rho(z')dz'} \tag{5.6.2}$$

The necessary $\Delta\rho(z)$ is obtained as difference between total electron densities of the calculation with and without applied sawtooth potential.

While this method provides a convenient framework for getting dielectric profiles, there are several uncertainties associated with it. First, the density functional theory itself has approximations built in the exchange-correlation kernel. Next, the results show a spurious dependence on the magnitude of applied electric field, as shown on Figure 5-11.

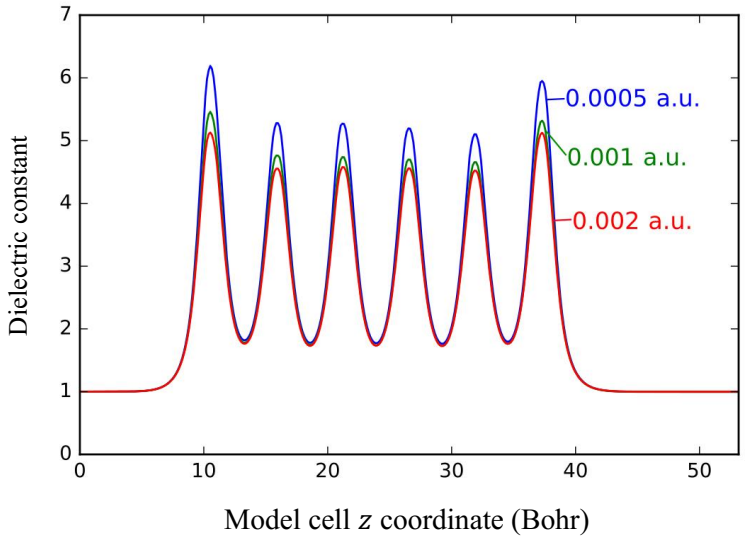

Figure 5-11. Computation of dielectric profile from applied sawtooth potential. The resulting dielectric profile amplitude can differ by over 1.0 depending on the field strength used in the simulation.

Moreover, there is need for averaging the oscillations due to presence of atomic nuclei, and the averaging methods, typically based on a convolution with some sort of smoothing kernel,[74] result in shifting of the media interface. Finally, this method ceases to be applicable when the ionic relaxations are included, because then the ionic displacements lead to too substantial rearrangements of the electronic density and to discontinuities in the dielectric profile. This could be remedied by carrying out a separate calculation for estimating the ionic contribution to screening, based on a charge partitioning scheme, but such calculations are

typically implemented for bulk systems only and would not allow to infer position and shape of dielectric boundary interface.

Overall, the number of issues pointed out throughout the chapter so far: sensitivity of electrostatic energy to position of model Gaussian charge, ambiguity in determining the dielectric profile from first principles, and presence of the legacy potential alignment term in the modern correction scheme, lead us to propose an internally consistent scheme for computation of defect formation energies, capitalizing on the electrostatic computation methodology described above. We describe this approach and its applications in the next section.

5.7. Internally consistent electrostatic correction scheme

The corrected formation energy of a charged defect, $E_f(q)$, is obtained after applying post-processing correction to the total energy of the supercell with a charged defect obtained from DFT, $E_{DFT}^{def}(q)$:

$$E_f(q) = E_{DFT}^{def}(q) - E_{DFT}^{st} + \sum_i \mu_i n_i + q(E_{VBM} + E_F) + E_{corr} \tag{5.7.1}$$

where E_{DFT}^{st} is the DFT total energy of the stoichiometric slab, μ_i are the chemical potentials of the species added or removed to create the defects under appropriate thermodynamic conditions, n_i are the stoichiometric coefficients for those species, E_{VBM} is the position of the valence band maximum for stoichiometric host material, E_F is the Fermi level position of the material with respect to valence band maximum, and E_{corr} is the correction energy. This term is equal to the difference of electrostatic energies of the defect charge distribution under open boundary condition, E_{iso}, and periodic boundary conditions, E_{PBC}. As discussed above, computation of those energies relies on two main inputs: defect charge distribution and the shape of the dielectric profile. The key difference of our approach is that we use the exact charge distribution

from the DFT computation, and fine-tune the dielectric profile features to eliminate the potential alignment term, as explained further in the text. We consider the case of positively charged Cl vacancy, V_{Cl}^{+} on the NaCl (100) surface to facilitate comparison to previous work.[65] For DFT computations in this section we use QuantumEspresso package.[33] For simulation of NaCl surfaces we used $2 \times 2 \times 3$ supercell with $4 \times 4 \times 1$ k-point sampling grid, cutoffs for wavefunctions expansion in terms of plane waves of 30 Ry, and for density – 300 Ry. For MoS_2 (to be considered next) we use 6×6 supercell which we reshape to conform to rectangular shape, with Γ point sampling for Brillouin zone, and 50 Ry cutoff for the wavefunctions, 500 Ry for charge density.

Instead of using a Gaussian distribution for the defect-related charge, we use the actual $|\phi(\vec{r})|^2$ obtained from the DFT calculation. There are several reasons for doing this: first, we find that very often the corresponding defect wavefunctions are highly anisotropic and have several lobes (Figure 5-12), so a smooth Gaussian model is an inappropriate description; second, the complex shape of the wavefunction leads to a substantial ambiguity in locating the center of the Gaussian and our model calculations reveal that the shift of the charge center in the direction perpendicular to the surface by 0.15 Bohr (well within the ambiguity involved) results in changing the electrostatic model energy by 0.1 eV. Similarly, the anisotropic shape of the wavefunction results in a poorly determined Gaussian fit width and the uncertainty in this parameter can lead to differences of up to 0.25 eV in the model energy. Moreover, for the multi-lobe defect wavefunctions, like those of related to forming a sulfur vacancy in MoS_2, the over-estimation in the width of the Gaussian can lead to "spilling over" of the model charge from the simulation cell, which is the case when the cell dimensions are smaller than 8 standard deviations of the Gaussian fit ($\pm 4\sigma$ is required to contain 99.99% of the charge). This is important, since

we find that losing more than 0.1% of the charge results in errors in electrostatic energy on the order of 0.1 eV. A Gaussian model is historically used due to the availability of analytical expressions for the electrostatic energy and fast convergence of the model electrostatic energy with respect to the discretized mesh size. In our parallel implementation of the potential computation this is not an important factor and we can explicitly use the defect wavefunction in the Poisson equation.

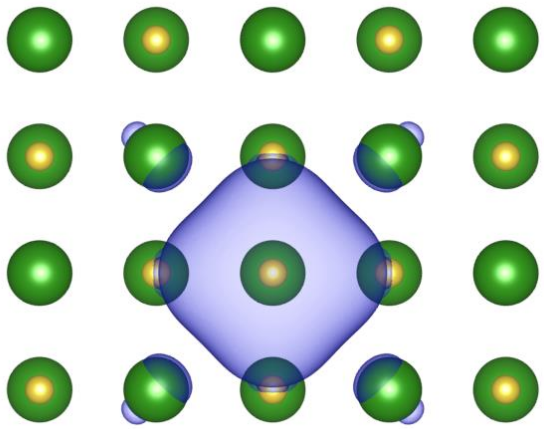

Figure 5-12. Example of a defect-induced wavefunction.

Top view of NaCl (100) surface with V_{Cl}^{+}. Green – Cl ions, yellow – Na ions. Blue cloud represents the $\rho_d(\vec{r}) = |\phi_d(\vec{r})|^2$ for the defect level in the bandgap. One can note asymmetry in the charge distribution shape and additional lobes on Cl atoms around the vacancy.

The other input to our method is the shape of the dielectric profile. For this, we use a model of two constant dielectric regions joined by error functions at the materials interfaces, with the parameters defining the profile being: the material dielectric constant in each region (for vacuum it is 1.0 by definition) and the positions of the interfaces. As we have discussed before, computation of the dielectric profile based on DFT is flawed with limitations, such as the commonly employed semi-local exchange-correlation functionals, and inapplicability when the ionic relaxations are included. Moreover, since we are interested in microscopic screening properties, the bulk dielectric constant might not be the right data to use. Instead, we model the dielectric profile approximately starting from the experimental value for material dielectric

constant and using the average of atomic radii of the surface atoms to get the profile boundaries. We then fine-tune those parameters in order to achieve alignment to the DFT potential. As we've mentioned before, the entire notion of potential alignment was shown to define the mismatch of the potentials induced by the unscreened defect charge in the model calculation and in the actual DFT computation.[65] This term is typically expressed as

$$\Delta V = V_{PBC}|_{far} - \left(V_{DFT}^{st} - V_{DFT}^{def}(q)\right)|_{far} \tag{5.7.2}$$

where V_{DFT}^{st} is the electrostatic potential for the stoichiometric slab, $V_{DFT}^{def}(q)$ is the potential for the slab with a charged defect and subscript "*far*" denotes the vacuum region of the supercell farthest from the defect. The potential alignment term arises due to the approximations made in the electrostatic model. Since we are using the exact wavefunction of the defect, we adjust the dielectric profile parameters in a way that properly aligns the model potential and the DFT potential difference in the vacuum region of the simulation supercell far from the defect.

The fitting procedure is greatly facilitated by the fact that the model electrostatic potential has qualitatively different dependence on the dielectric constant and the positions of the profile boundaries, as shown in Figure 5-13: varying the value of the dielectric constant changes the amplitude of the features on the model potential and the slope in the alignment region. For NaCl we choose the value of 2.8 which results in flat ΔV, as shown on the right panel of Figure 5-13. Variation of the profile boundary position results in a rigid shift of the potential in vacuum. Overall, by adjusting those parameters one can find a combination resulting in a flat line about zero for ΔV.

Finally, we note that the inaccuracies associated with sampling the defect-induced charge lead to random errors of about 0.03 eV in the values of E_{PBC}. These errors do not converge fast with finer mesh sampling, so there is no need to specifically increase the sampling and plane

wave expansion cutoff in the DFT calculations. Moreover, we find that downsampling of the output wavefunction by a factor of 2 or 3 (to the mesh size of about 0.3-0.4 Bohr) changes the E_{PBC} by about 10^{-3} eV – which can be used to choose computational parameters optimally to reduce the cost of the calculations.

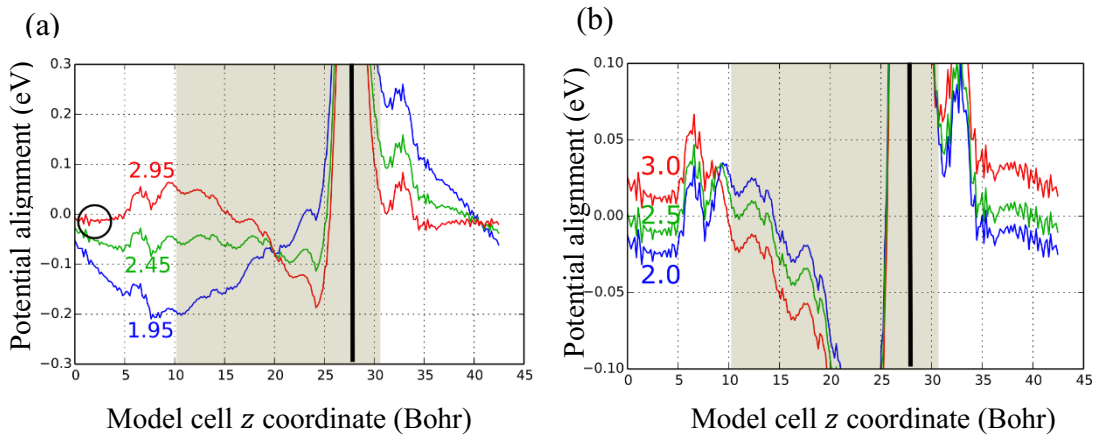

Figure 5-13. Potential alignment as a function of model profile parameters.

(a) Potential alignment as a function of the dielectric constant of host material in the model dielectric profile. Potential alignment region is denoted by a black circle; position of the slab is denoted by shading, position of the defect – by a thick black line.

(b) Potential alignment as a function of offset of the boundaries of the model dielectric profile compared to ionic positions of slab outer layer atoms.

It is worth reiterating that the extrapolation is valid only for the model supercell of strictly cubic shape; extrapolation from cells of different shape result in vastly different and incorrect E_{iso} values. Accordingly, when simulating real materials, the defect wavefunction has to be trimmed to a cubic shape for use in the extrapolation procedure. Specifically, upon scaling the system, we pad the trimmed charge distribution with zeros on all sides, placing it in the center of the scaled cell (Figure 5-10(a)). For the case of semiconductor surface regions, we calculate the position of the dielectric profile boundary closest to the charge by setting the offset

to be the same as in the original cell. The second boundary position is calculated by scaling the thickness of the material proportionally to the supercell size.

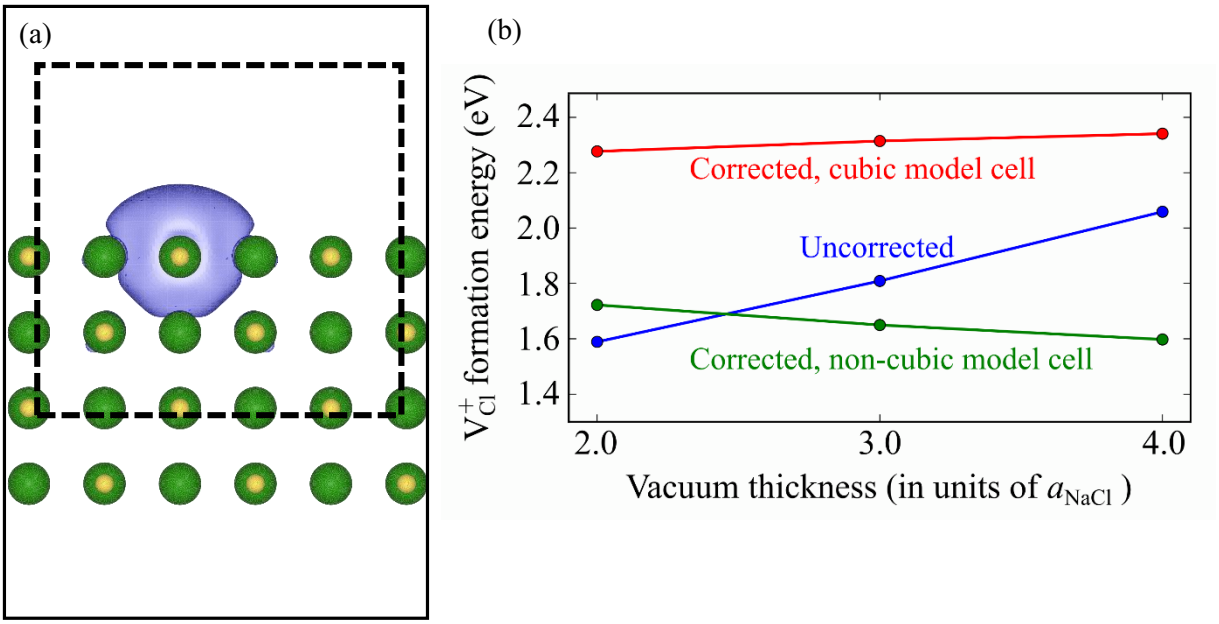

Figure 5-14. Cl vacancy formation on NaCl (100) surface.

(a) NaCl (100) surface with V_{Cl}^{+}. Green – Cl ions, yellow – Na ions; vertical direction is perpendicular to the surface. Blue cloud represents the $\rho_d(\vec{r}) = |\phi_d(\vec{r})|^2$ for the defect level in the bandgap. Solid line is DFT simulation cell boundary. Dashed line denotes the trimmed cubic part of the defect charge used in extrapolation procedure.

(b) Formation energy of V_{Cl}^{+} as a function of vacuum size (in units of $a_{NaCl} = 10.6$ Bohr). Uncorrected (blue) and corrected energies with non-cubic model cell (green) used for extrapolation show large variance; usage of the correct extrapolation procedure results in formation energy values consistent within 0.06 eV (red).

We investigate the performance of the correction scheme by calculating formation energies of V_{Cl}^{+} on the NaCl (100) surface for several supercells with varying vacuum thickness and lateral dimensions. The results are shown in Figure 5-14 for the case of varying vacuum thickness. The variance in the uncorrected energies (blue line) is as substantial as the variance in

corrected energies with extrapolation from the wavefunction charge distributions of non-cubic shape (green line). Only correction with the proper extrapolation procedure results in consistent formation energies within 0.06 eV, independent of the supercell shape (red line). Analogously, the dependence on the lateral size of the cell is eliminated.

The above scheme can be successfully used for 2D materials as demonstrate with the example of the V_S^- defect in a MoS_2 monolayer. The only change needed is the method of scaling the model profile in computing the E_{iso}: in this case, the positions of both profile boundaries are fixed relative to the charge, which results in keeping the material thickness constant throughout the extrapolation procedure.

An important feature of low-dimensional systems is that the actual values of the diagonal elements of the dielectric tensor do not affect the model potential as much as the positions of the boundaries of the dielectric profile. As shown in Figure 5-15, the values of the model potential in the alignment region are very close. We use the values of $\epsilon_\perp = 6.0$ for the out-of-plane component and $\epsilon_\parallel = 15.0$ for the in-plane component; we find the optimal position of the profile boundaries to be at an offset of 2.7 Bohr outwards from the S atoms. Application of our correction scheme results in elimination of the dependence of the vacancy formation energy on the vacuum layer thickness, the corrected formation energies being consistent to within 0.1 eV. In such manner, we avoid the spurious methods for estimating the dielectric profile by renormalizing charge density.[76]

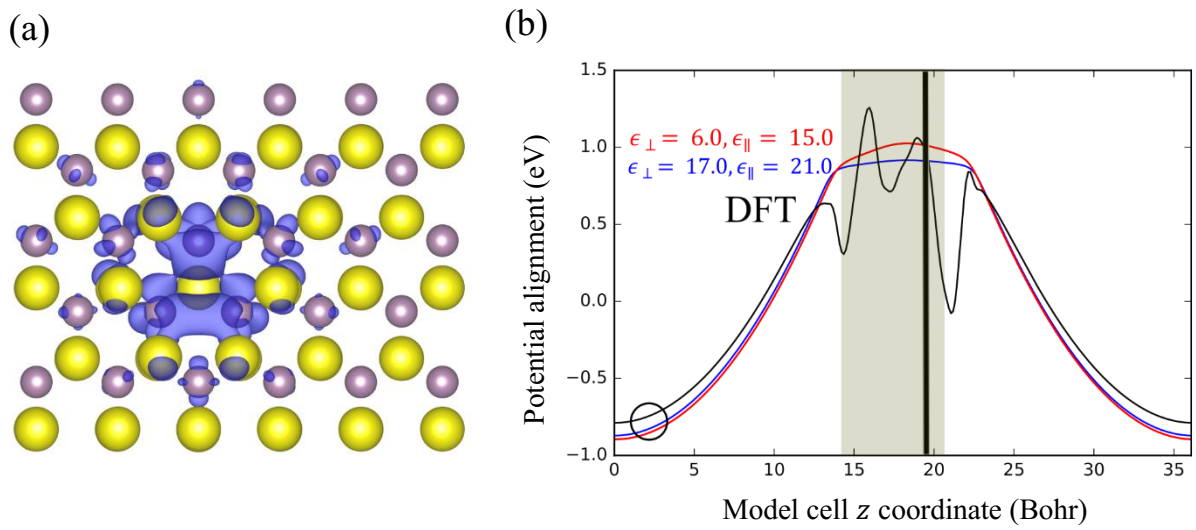

Figure 5-15. S vacancy formation in MoS_2 monolayer.

(a) Top view of MoS_2 monolayer with V_S^-. Yellow – S atoms, purple – Mo. The complex multi-lobe structure of the charge distribution, $\rho_d(\vec{r}) = |\phi(\vec{r})|^2$ is apparent.

(b) Potential alignment of DFT potentials for V_S^- (black) and model potentials for a variety of choices for in-plane and out-of-plane components of the dielectric tensor. The material atom positions are denoted by the shading, the defect position – by a black vertical line, the alignment region – by a black circle. The evident mismatch in the alignment region can be fixed by moving the positions of the dielectric profile boundaries outwards.

Overall, we have demonstrated an internally consistent scheme for computation of charged defect formation energies in systems with complex dielectric profiles. The algorithm is the following:

1. Construct the stoichiometric and defected slabs, obtain $|\phi(\vec{r})|^2$, the defect charge density, and the absolute level of the VBM and electrostatic potentials V_{DFT}^{st}, $V_{DFT}^{def}(q)$.

2. Fine-tune the parameters of the model dielectric profile, that is, the values of the dielectric constant and the positions of interfaces in order to achieve alignment between the model V_{PBC} and $V_{DFT}^{st} - V_{DFT}^{def}(q)$; calculate the corresponding E_{PBC}.

3. Trim $|\phi(\vec{r})|^2$ to a cubic shape, change the dielectric boundary positions accordingly, calculate E_{PBC} for a series of scaled model cells; obtain E_{iso} through extrapolation to infinite cell size.

4. Add the correction $E_{corr} = E_{iso} - E_{PBC}$ to the defect formation energy.

We note that the electrostatic correction described here is best suited for applications to 2D materials or semiconductors with low (< 10) dielectric constant. In materials with stronger screening the value of the electrostatic correction is small; at the same time, introduction of charged defects into the supercell results in substantial rearrangements of atoms, which are hard to contain in a supercell, even of a size as large as 1000 atoms. This leads to large errors due to elastic energy contributions, which become the dominant term among errors associated with the supercell method for such materials (an example of such case is TiO_2).

5.8. References

1. Freysoldt, C.; Grabowski, B.; Hickel, T.; Neugebauer, J.; Kresse, G.; Janotti, A.; Van de Walle, C. G., First-Principles Calculations for Point Defects in Solids. *Reviews of Modern Physics* **2014**, *86*, 253-305.

2. Van de Walle, C. G.; Neugebauer, J., First-Principles Calculations for Defects and Impurities: Applications to Iii-Nitrides. *Journal of Applied Physics* **2004**, *95*, 3851-3879.

3. Nieminen, R. M., Issues in First-Principles Calculations for Defects in Semiconductors and Oxides. *Modelling and Simulation in Materials Science and Engineering* **2009**, *17*, 084001.

4. Van de Walle, C. G.; Janotti, A., Advances in Electronic Structure Methods for Defects and Impurities in Solids. *Physica Status Solidi B-Basic Solid State Physics* **2011**, *248*, 19-27.

5. Alkauskas, A.; McCluskey, M.; Van de Walle, C. G., Tutorial: Defects in Semiconductors - Combining Experiment and Theory. *Journal of Applied Physics* **2016**, *119*, 181101.

6. Bracht, H.; Haller, E. E.; Clark-Phelps, R., Silicon Self-Diffusion in Isotope Heterostructures. *Physical Review Letters* **1998**, *81*, 393-396.

7. Puska, M. J.; Nieminen, R. M., Theory of Positrons in Solids and on Solid-Surfaces. *Reviews of Modern Physics* **1994**, *66*, 841-897.

8. Gilliland, G. D., Photoluminescence Spectroscopy of Crystalline Semiconductors. *Materials Science & Engineering R-Reports* **1997**, *18*, 99-399.

9. Hourahine, B.; Jones, R.; Safonov, A. N.; Oberg, S.; Briddon, P. R.; Estreicher, S. K., Identification of the Hexavacancy in Silicon with the B-80(4) Optical Center. *Physical Review B* **2000**, *61*, 12594-12597.

10. Rehr, J. J.; Albers, R. C., Theoretical Approaches to X-Ray Absorption Fine Structure. *Reviews of Modern Physics* **2000**, *72*, 621-654.

11. Tersoff, J.; Hamann, D. R., Theory of the Scanning Tunneling Microscope. *Physical Review B* **1985**, *31*, 805-813.

12. Feenstra, R. M., Scanning Tunneling Spectroscopy. *Surface Science* **1994**, *299*, 965-979.

13. Feenstra, R. M., Tunneling Spectroscopy of the (110)-Surface of Direct-Gap III-V Semiconductors. *Physical Review B* **1994**, *50*, 4561-4570.

14. Feenstra, R. M., Cross-Sectional Scanning-Tunneling-Microscopy of III-V Semiconductor Structures. *Semicond Sci Tech* **1994**, *9*, 2157-2168.

15. Garleff, J. K.; Wijnheijmer, A. P.; Koenraad, P. M., Challenges in Cross-Sectional Scanning Tunneling Microscopy on Semiconductors. *Semicond Sci Tech* **2011**, *26*, 064001.

16. Grandidier, B.; de la Broise, X.; Stievenard, D.; Delerue, C.; Lannoo, M.; Stellmacher, M.; Bourgoin, J., Defect-Assisted Tunneling Current: A Revised Interpretation of Scanning Tunneling Spectroscopy Measurements. *Applied Physics Letters* **2000**, *76*, 3142-3144.

17. Watkins, G. D.; Chow, K. H., Self-Interstitials in Semiconductors: What We Are Learning from Interstitial Zn in ZnSe. *Physica B-Condensed Matter* **1999**, *273-4*, 7-14.

18. Watkins, G. D.; Corbett, J. W., Defects in Irradiated Silicon - Electron Paramagnetic Resonance + Electron-Nuclear Double Resonance of Si-E Center. *Phys Rev* **1964**, *134*, 1359-1377.

19. Van De Walle, C. G.; Blochl, P. E., 1st-Principles Calculations of Hyperfine Parameters. *Physical Review B* **1993**, *47*, 4244-4255.

20. Pickard, C. J.; Mauri, F., All-Electron Magnetic Response with Pseudopotentials: Nmr Chemical Shifts. *Physical Review B* **2001**, *63*, 245101.

21. Pickard, C. J.; Mauri, F., First-Principles Theory of the EPR *g*-Tensor in Solids: Defects in Quartz. *Physical Review Letters* **2002**, *88*, 086403.

22. Baraff, G. A.; Schluter, M., New Self-Consistent Approach to the Electronic-Structure of Localized Defects in Solids. *Physical Review B* **1979**, *19*, 4965-4979.

23. Baraff, G. A.; Schluter, M., Self-Consistent Greens Function Calculation of Ideal Si Vacancy. *Physical Review Letters* **1978**, *41*, 892-895.

24. Beeler, F.; Andersen, O. K.; Scheffler, M., Theoretical Evidence for Low-Spin Ground-States of Early Interstitial and Late Substitutional 3d Transition-Metal Ions in Silicon. *Physical Review Letters* **1985**, *55*, 1498-1501.

25. Gerstmann, U., Ab Initio Green's Function Calculation of Hyperfine Interactions for Shallow Defects in Semiconductors. *Physica Status Solidi B-Basic Solid State Physics* **2011**, *248*, 1319-1336.

26. Car, R.; Kelly, P. J.; Oshiyama, A.; Pantelides, S. T., Microscopic Theory of Atomic Diffusion Mechanisms in Silicon. *Physica B & C* **1984**, *127*, 401-407.

27. Deak, P.; Miro, J.; Gali, A.; Udvardi, L.; Overhof, H., The Spin State of the Neutral Silicon Vacancy in 3c-SiC. *Applied Physics Letters* **1999**, *75*, 2103-2105.

28. Pacchioni, G., Ab Initio Theory of Point Defects in Oxide Materials: Structure, Properties, Chemical Reactivity. *Solid State Sci* **2000**, *2*, 161-179.

29. Jensen, F., *Introduction to Computational Chemistry*; John Wiley & Sons: Chichester, 2007.

30. Deak, P., Choosing Models for Solids. *Phys Status Solidi B* **2000**, *217*, 9-21.

31. Lambrecht, W. R. L., Which Electronic Structure Method for the Study of Defects: A Commentary. *Physica Status Solidi B-Basic Solid State Physics* **2011**, *248*, 1547-1558.

32. Louie, S. G.; Schluter, M.; Chelikowsky, J. R.; Cohen, M. L., Self-Consistent Electronic States for Reconstructed Si Vacancy Models. *Physical Review B* **1976**, *13*, 1654-1663.

33. Giannozzi, P., et al., Quantum Espresso: A Modular and Open-Source Software Project for Quantum Simulations of Materials. *J. Phys.-Condes. Matter* **2009**, *21*, 395502.

34. Bates, S. P.; Gillan, M. J.; Kresse, G., Adsorption of Methanol on TiO_2 (110): A First-Principles Investigation. *Journal of Physical Chemistry B* **1998**, *102*, 2017-2026.

35. Zhang, S. B.; Northrup, J. E., Chemical-Potential Dependence of Defect Formation Energies in GaAs - Application to Ga Self-Diffusion. *Physical Review Letters* **1991**, *67*, 2339-2342.

36. Lang, D. V., Deep-Level Transient Spectroscopy - New Method to Characterize Traps in Semiconductors. *Journal of Applied Physics* **1974**, *45*, 3023-3032.

37. Lany, S.; Zunger, A., Assessment of Correction Methods for the Band-Gap Problem and for Finite-Size Effects in Supercell Defect Calculations: Case Studies for ZnO and GaAs. *Physical Review B* **2008**, *78*, 235104.

38. Baraff, G. A.; Schluter, M., Migration of Interstitials in Silicon. *Physical Review B* **1984**, *30*, 3460-3469.

39. Zhang, S. B.; Wei, S. H.; Zunger, A., Intrinsic N-Type Versus P-Type Doping Asymmetry and the Defect Physics of ZnO. *Physical Review B* **2001**, *63,* 075205.

40. Zhang, S. B., The Microscopic Origin of the Doping Limits in Semiconductors and Wide-Gap Materials and Recent Developments in Overcoming These Limits: A Review. *J. Phys.-Condes. Matter* **2002**, *14*, R881-R903.

41. Boguslawski, P.; Briggs, E. L.; Bernholc, J., Native Defects in Gallium Nitride. *Physical Review B* **1995**, *51*, 17255-17258.

42. Lany, S.; Zunger, A., Polaronic Hole Localization and Multiple Hole Binding of Acceptors in Oxide Wide-Gap Semiconductors. *Physical Review B* **2009**, *80*, 085202-085202.

43. Lany, S.; Zunger, A., Intrinsic DX Centers in Ternary Chalcopyrite Semiconductors. *Physical Review Letters* **2008**, *100,* 016401.

44. Aryasetiawan, F.; Gunnarsson, O., The GW Method. *Rep. Prog. Phys.* **1998**, *61*, 237-312.

45. Shishkin, M.; Kresse, G., Self-Consistent GW Calculations for Semiconductors and Insulators. *Physical Review B* **2007**, *75*, 235102.

46. Heyd, J.; Scuseria, G. E., Efficient Hybrid Density Functional Calculations in Solids: Assessment of the Heyd-Scuseria-Ernzerhof Screened Coulomb Hybrid Functional. *Journal of Chemical Physics* **2004**, *121*, 1187-1192.

47. Heyd, J.; Scuseria, G. E.; Ernzerhof, M., Hybrid Functionals Based on a Screened Coulomb Potential. *Journal of Chemical Physics* **2003**, *118*, 8207-8215.

48. Chen, W.; Pasquarello, A., Correspondence of Defect Energy Levels in Hybrid Density Functional Theory and Many-Body Perturbation Theory. *Physical Review B* **2013**, *88*, 115104.

49. Schultz, P. A., Theory of Defect Levels and the "Band Gap Problem" in Silicon. *Physical Review Letters* **2006**, *96*, 246401.

50. Korzhavyi, P. A.; Abrikosov, I. A.; Johansson, B.; Ruban, A. V.; Skriver, H. L., First-Principles Calculations of the Vacancy Formation Energy in Transition and Noble Metals. *Physical Review B* **1999**, *59*, 11693-11703.

51. Leslie, M.; Gillan, M. J., The Energy and Elastic Dipole Tensor of Defects in Ionic-Crystals Calculated by the Supercell Method. *Journal of Physics C-Solid State Physics* **1985**, *18*, 973-982.

52. Kleinman, L., Comment on the Average Potential of a Wigner Solid. *Physical Review B* **1981**, *24*, 7412-7414.

53. Makov, G.; Payne, M. C., Periodic Boundary-Conditions in Ab-Initio Calculations. *Physical Review B* **1995**, *51*, 4014-4022.

54. Dabo, I.; Kozinsky, B.; Singh-Miller, N. E.; Marzari, N., Electrostatics in Periodic Boundary Conditions and Real-Space Corrections. *Physical Review B* **2008**, *77*, 115139.

55. Shim, J.; Lee, E. K.; Lee, Y. J.; Nieminen, R. M., Density-Functional Calculations of Defect Formation Energies Using Supercell Methods: Defects in Diamond. *Physical Review B* **2005**, *71*, 035206.

56. Wright, A. F.; Modine, N. A., Comparison of Two Methods for Circumventing the Coulomb Divergence in Supercell Calculations for Charged Point Defects. *Physical Review B* **2006**, *74*, 235209.

57. Castleton, C. W. M.; Hoglund, A.; Mirbt, S., Density Functional Theory Calculations of Defect Energies Using Supercells. *Modelling and Simulation in Materials Science and Engineering* **2009**, *17*, 084003.

58. Castleton, C. W. M.; Hoglund, A.; Mirbt, S., Managing the Supercell Approximation for Charged Defects in Semiconductors: Finite-Size Scaling, Charge Correction Factors, the Band-Gap Problem, and the Ab Initio Dielectric Constant. *Physical Review B* **2006**, *73*, 035215.

59. Hine, N. D. M.; Frensch, K.; Foulkes, W. M. C.; Finnis, M. W., Supercell Size Scaling of Density Functional Theory Formation Energies of Charged Defects. *Physical Review B* **2009**, *79*, 024112.

60. Murphy, S. T.; Hine, N. D. M., Anisotropic Charge Screening and Supercell Size Convergence of Defect Formation Energies. *Physical Review B* **2013**, *87*, 094111.

61. Freysoldt, C.; Neugebauer, J.; Van de Walle, C. G., Electrostatic Interactions between Charged Defects in Supercells. *Physica Status Solidi B-Basic Solid State Physics* **2011**, *248*, 1067-1076.

62. Freysoldt, C.; Neugebauer, J.; Van de Walle, C. G., Fully Ab Initio Finite-Size Corrections for Charged-Defect Supercell Calculations. *Physical Review Letters* **2009**, *102*, 016402.

63. Komsa, H.-P.; Rantala, T.; Pasquarello, A., Comparison between Various Finite-Size Supercell Correction Schemes for Charged Defect Calculations. *Physica B-Condensed Matter* **2012**, *407*, 3063-3067.

64. Kumagai, Y.; Oba, F., Electrostatics-Based Finite-Size Corrections for First-Principles Point Defect Calculations. *Physical Review B* **2014**, *89*, 195205.

65. Komsa, H.-P.; Pasquarello, A., Finite-Size Supercell Correction for Charged Defects at Surfaces and Interfaces. *Physical Review Letters* **2013**, *110*, 095505.

66. Dalcin, L.; Paz, R.; Storti, M., MPI for Python. *J Parallel Distr Com* **2005**, *65*, 1108-1115.

67. Millman, K. J.; Aivazis, M., Python for Scientists and Engineers. *Comput Sci Eng* **2011**, *13*, 9-12.

68. Strassen, V., Gaussian Elimination Is Not Optimal. *Numer Math* **1969**, *13*, 354-356.

69. Greengard, L.; Rokhlin, V., A Fast Algorithm for Particle Simulations. *Journal of Computational Physics* **1987**, *73*, 325-348.

70. Garcia-Risueno, P.; Alberdi-Rodriguez, J.; Oliveira, M. J. T.; Andrade, X.; Pippig, M.; Muguerza, J.; Arruabarrena, A.; Rubio, A., A Survey of the Parallel Performance and Accuracy of Poisson Solvers for Electronic Structure Calculations. *Journal of Computational Chemistry* **2014**, *35*, 427-444.

71. Taddei, M. M.; Mendes, T. N. C.; Farina, C., Subtleties in Energy Calculations in the Image Method. *Eur J Phys* **2009**, *30*, 965-972.

72. Resta, R.; Kunc, K., Self-Consistent Theory of Electronic States and Dielectric Response in Semiconductors. *Physical Review B* **1986**, *34*, 7146-7157.

73. Giustino, F.; Pasquarello, A., Theory of Atomic-Scale Dielectric Permittivity at Insulator Interfaces. *Physical Review B* **2005**, *71*, 144104.

74. Baldereschi, A.; Baroni, S.; Resta, R., Band Offsets in Lattice-Matched Heterojunctions - a Model and 1st-Principles Calculations for GaAs/AlAs. *Physical Review Letters* **1988**, *61*, 734-737.

75. Komsa, H. P.; Berseneva, N.; Krasheninnikov, A. V.; Nieminen, R. M., Charged Point Defects in the Flatland: Accurate Formation Energy Calculations in Two-Dimensional Materials. *Phys Rev X* **2014**, *4*, 031044.

CHAPTER 6. DEFECT PROPERTIES OF MONOLAYER MOLYBDENUM DISULFIDE

Here we apply the method for computation of the defect formation energies described in Chapter 5 to a system of practical relevance – monolayer MoS_2. We investigate the behavior of sulfur vacancy defects, the most abundant type of intrinsic defect in monolayer MoS_2, using first-principles calculations based on density functional theory. We consider the dependence of the isolated defect formation energy on the charge state and on uniaxial tensile strain up to 5%. We also consider the possibility of defect clustering by examining the formation energies of pairs of vacancies at various relative positions, and their dependence on charge state and strain. We find that there is no driving force for vacancy clustering, independent of strain in the material. The barrier for diffusion of S vacancies is larger than 1.9 eV in both charged and neutral states regardless of the presence of other nearby vacancies. We conclude that the formation of extended defects from S vacancies in both planar monolayer MoS_2 and in the ripplocation structure is hindered both thermodynamically and kinetically. The work presented here is being prepared for publication as M. G. Sensoy, D. Vinichenko, W. Chen, C. M. Friend, E. Kaxiras. On the behavior of isolated and paired sulfur vacancy defects in monolayer MoS_2: Implications for ripplocation formation; to which Dmitry Vinichenko and Mehmet Gokhan Sensoy contributed equally.

6.1. Introduction

The class of semiconducting, two dimensional (2D) materials referred to as transition metal dichalcogenides (TMDCs) has been the subject of intense research activity due to their interesting behavior which includes their PL properties,[1-2] interlayer excitons,[3] and valley

physics.[4] The prototypical material in this class, MoS_2, has been studied for potential applications in the context of catalysis,[5-6] in energy storage,[7] and as an electrode for hydrogen evolution.[8-13] Moreover, single layers of this compound have been proposed for use in field-effect transistors,[14-19] including bendable devices,[20] and integrated circuits.[21-22] In most of these applications, the defect properties of MoS_2 monolayers are of central importance. In this work, we use first-principles computations to study the sulfur vacancy in MoS_2, which is the intrinsic defect with the lowest formation energy.[23] We investigate the formation energy and diffusion barrier of this defect as a function of charge state and strain, as well as the collective behavior of vacancies, that is, the interactions between vacancy pairs at various distances. These aspects of their behavior are relevant in elucidating the nature of the recently proposed ripplocation structure.[24] Specifically, we apply these results to understand how tensile strain, which is associated with the presence of ripplocations in monolayer MoS_2 affects the formation energies and diffusion barriers for isolated and clustered vacancies, and what are the consequences for the electronic properties of the material

6.2. Methodology

Our calculations are based on the plane-wave, pseudopotential density functional theory (DFT) method as implemented in the Quantum Espresso code suite.[25] We use Γ-point Brillouin zone sampling, a cut-off energy of 50 Ry for the wavefunction plane-wave expansion and 500 Ry for the density, and a 6×6 (8×8) super-cell for the isolated defect (pairs) calculations with a 16 Å vacuum region separating the periodic images of layers. All defect configurations are relaxed to the point where the calculated forces on atoms do not exceed in magnitude 0.05 eV/Å. The defect formation energy $E_f(q)$ in the thermodynamic limit is obtained from

$$E_f(q) = E_{\mathrm{DFT}}^{def}(q) - E_{\mathrm{DFT}}^{st} + \mu_{\mathrm{S}} n_{\mathrm{S}} + q(E_{\mathrm{VBM}} + E_F) + E_{corr} \quad (6.2.1)$$

where $E_{\mathrm{DFT}}^{def}(q)$ is the DFT total energy of the layer containing the defects in the charge state q, E_{DFT}^{st} is the DFT total energy of the stoichiometric layer, μ_{S} the chemical potential of S, n_{S} the number of S vacancies in the simulation cell, E_{VBM} the valence band maximum energy, E_F the Fermi level with respect to the valence band maximum, and E_{corr} the electrostatic correction necessary for the charged states of the defect. For the chemical potential of S we use the DFT calculated energy of the S_8 molecule in gas phase. For the computation of E_{corr} we parameterize the dielectric profile of the model slab $\epsilon(z)$ as a piece-wise constant joined by a smoothing error function, with the dielectric constant of the material being $\epsilon_{\perp} = 6$, $\epsilon_{\parallel} = 15$ in the directions perpendicular and parallel to the layer, respectively,[26] and we take the effective layer thickness to be 6 Å. We obtain the defect charge distribution, $\rho_d(\vec{r})$, by summing the magnitude of the occupied defect states in the band gap. With this density, we solve for the defect-induced electrostatic potential under periodic boundary conditions, $V_{\mathrm{PBC}}(\vec{r})$, from the Poisson equation, $\nabla\,[\epsilon(z)\nabla V_{\mathrm{PBC}}(\vec{r})] = -\rho_d(\vec{r})$, and compute the corresponding defect charge electrostatic energy under periodic boundary conditions, $E_{\mathrm{PBC}} = \frac{1}{2}\int V_{\mathrm{PBC}}(\vec{r})\rho_d(\vec{r})d\vec{r}$. In order to obtain the electrostatic energy of the isolated charge under open boundary conditions, E_{iso}, we extrapolate E_{PBC} to the limit of an infinite model cell. The correction is then calculated as a difference of those quantities, $E_{corr} = E_{iso} - E_{\mathrm{PBC}}$. This approach and its limitations and advantages are discussed in more detail in a forthcoming article.[27]

6.3. Isolated vacancy

MoS_2 is a semiconductor, with band gap of 1.9 eV in monolayer[28-31] and 1.3 eV in bulk form.[32] The spin-orbit coupling, which is generally a prominent feature in TMDCs and especially

those formed by heavy atoms, does not have a substantial effect on the electronic structure of the MoS_2 monolayer.[33] Interestingly, the exciton binding energy is strongly enhanced in the monolayer relative to the bulk.[34-35] The Mo atoms are surrounded by trigonals prism of S atoms (Figure 6-1). Upon the creation of an isolated S vacancy the relaxation of neighboring atoms is limited to within 4 Å from the vacancy in both the neutral and charged $(q = -1)$ states (see Figure 6-2(a)). The displacements of atoms farther than 4 Å from the vacancy are smaller than 0.05 Å and are not shown. Moreover, the atomic displacements around the vacancy are 10% larger for the charged state compared to neutral state. In order to understand these features, we analyze in detail the electronic structure of the defect.

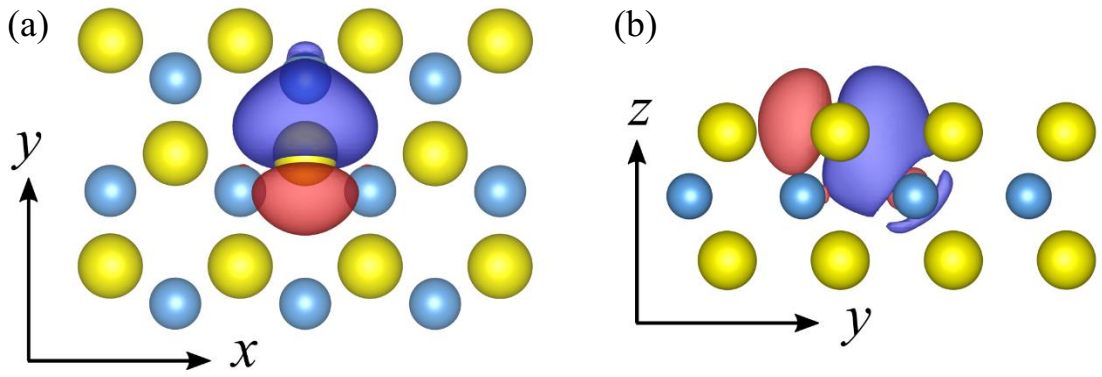

Figure 6-1. Wannier function corresponding to the σ-bond between Mo and S atoms, in (a) top and (b) side views. Mo atoms are cyan and S atoms yellow. Blue and red lobes are positive and negative the isosurfaces of the Wannier function.

In a localized representation of the electronic states based on Wannier orbitals,[36] the electronic structure of MoS_2 can be described as a combination of directed σ-bonds between Mo and S atoms, lone pairs on S atoms, pointing outwards, and non-bonding states localized on the Mo atoms. In this picture, the valence bands are spanned by the σ-bonds and the valence band edge is composed of non-bonding states on Mo.[37] Creation of a S vacancy introduces two

unoccupied defect levels in the band gap of the material and one fully occupied state below the valence band maximum. The localized orbitals of the states in the gap originate from the three severed Mo−S σ-bonds and, due to strong overlap between the d-orbitals on Mo atoms, they are delocalized between three Mo atoms adjacent to the vacancy. These localized states can potentially serve as electron traps and affect the exciton binding energy in monolayer MoS_2. Qualitatively, the electronic structure of defect states can be understood by considering a Huckel-type model for the hybrid orbitals of σ-bonds between Mo and S atoms. For three bonds in trigonal arrangement the electronic structure is analogous to the cyclopropene π-system: one occupied level is resonant with the valence band and two band gap levels are degenerate, with one nodal plane in each of them. The nodal planes are in the *xz* and *yz* planes, so the orbitals are labeled φ_x and φ_y, respectively, see Figure 6-2(b). Earlier work has shown that the S vacancy can act as an acceptor due to the presence of these empty levels.[23]

We investigate how the absolute position of those states in the gap changes under uniaxial strain in the *y* direction as defined in Figure 6-1, the direction that is more susceptible to stretching and is relevant for the possible formation of defects in the recently discovered "ripplocation" structure.[24] Overall, the band gap of the material decreases with applied strain, in line with previous computational results.[38] Applying strain in the *y* direction leads to removal of the degeneracy between φ_x and φ_y, with the energy of φ_y decreasing more than the energy of φ_x, see Figure 6-2 (c). We hypothesize that this is due to the presence of the nodal plane *xz* for φ_x, which is perpendicular to the strain direction, so the energy of this state is less affected by the decrease in the orbital overlap compared to φ_y.

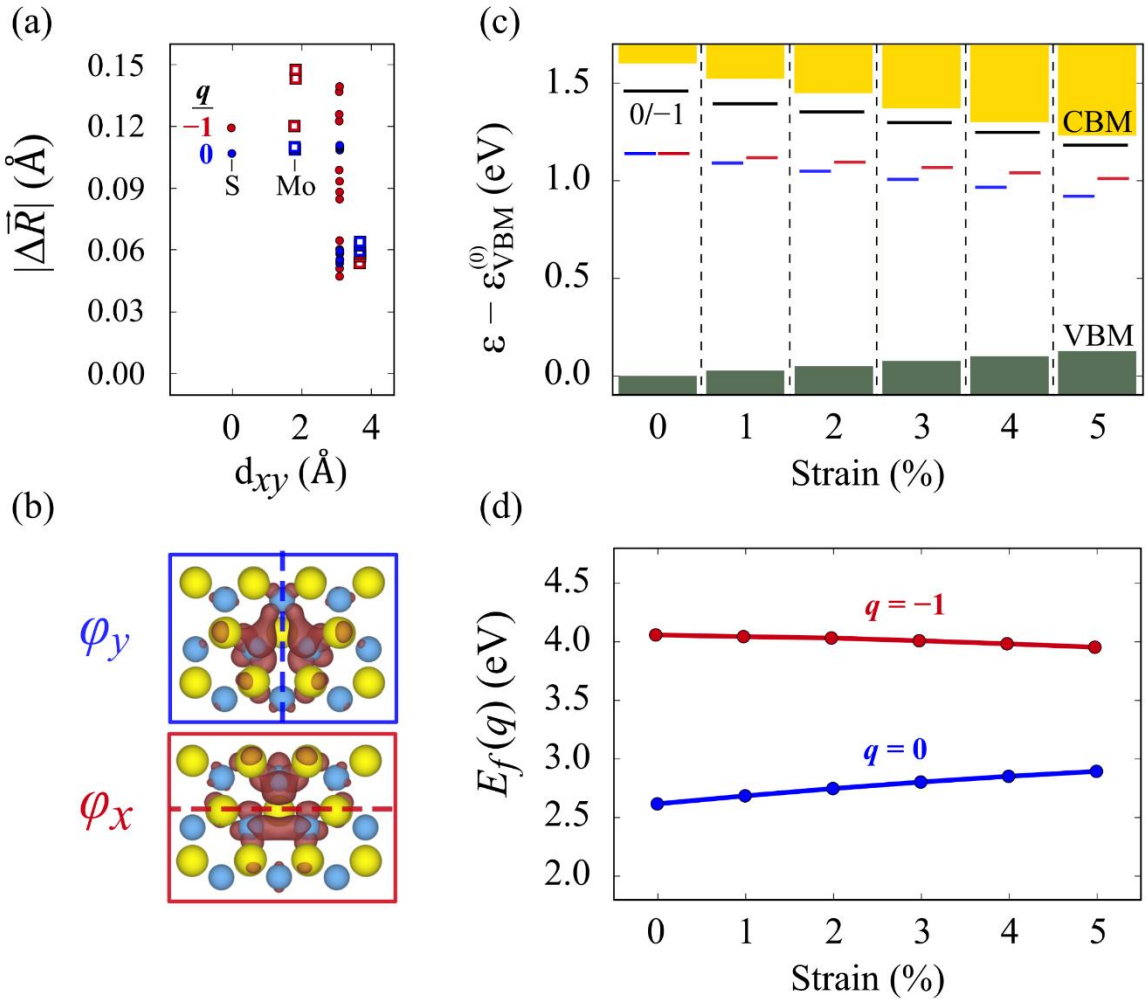

Figure 6-2. Properties of an isolated sulfur vacancy.

(a) Magnitude of atomic displacements around the isolated S vacancy in monolayer MoS_2, $|\Delta\vec{R}|$, as a function of distance d_{xy} between the atom and the vacancy in the *xy*-plane.

(b) Isosurfaces of the defect states $|\varphi_y(\vec{r})|^2$ and $|\varphi_x(\vec{r})|^2$.

(c) Energy levels of the defect states φ_y (blue), φ_x (red) and thermodynamic charge transition level (black) in the band gap associated with a single S vacancy: their position within the gap range as a function of strain, relative to the valance band maximum of the unstrained monolayer MoS_2, $\varepsilon^{(0)}_{VBM}$.

(d) Formation energy of isolated S vacancy as a function of strain.

Next, we calculate the formation energy of the isolated S vacancy in neutral ($q = 0$) and charged ($q = -1$) states as a function of uniaxial strain up to 5%, the breaking strain for monolayer MoS_2 being between 6 and 11%.[39] We do not consider the charge state $q = -2$ since it was shown to be unstable in previous work.[23] The results are presented in Figure 6-2(d). In the neutral state, for values of tensile strain up to 5% there is a clear trend, that is, an increase of the formation energy. For the charged state, the value of the formation energy shows a marginally decreasing dependence on applied strain. We suggest that this effect is due to occupation of the band gap defect levels φ_x, φ_y, which are localized on the Mo *d*-level and have antibonding character, as discussed above. Another factor could be the increase of the occupation of σ^* Mo−S antibonding orbitals surrounding the vacancy through the mechanism of geminal hyperconjugation, which facilitates the lattice relaxation for accommodating the defect, as evidenced by the larger relaxation of the lattice around the charged defect compared to the neutral vacancy, see Figure 6-2(a).

We have also considered the transition level, that is, the position of the Fermi level of the material in the gap at which $E_f(0) = E_f(-1)$, denoted by a black line in Figure 6-2(c), as a function of strain. We find that this level moves down in energy with increasing strain, indicating that the energy difference between charged and neutral defect states decreases, while the deep-acceptor character of the isolated S vacancy is preserved.

6.4. Vacancy pairs

We address next the possibility of clustering of S vacancies in the MoS_2 monolayer. To this end we calculate the formation energy of two S vacancies as a function of their spatial separation. We consider various configurations of the two vacancies: first on top of each other on

opposite sides of the sheet, labeled (0,0), and then, three different cases for vacancies on the same side of the sheet, for both *x* and *y* directions: immediately adjacent vacancies labeled ($0,1_x$) and ($0,1_y$), separated by one S atom labeled ($0,2_x$) and ($0,2_y$), and separated by two S atoms labeled ($0,3_x$) and ($0,3_y$). In Figure 6-3(a), we show the geometries of vacancy pairs along the *x* direction. As far as strain is concerned, we compared two cases: unstrained material and 5% tensile strain applied along the *y* direction. Formation energies along the *x* direction are shown in Figure 6-3(b). Formation energies along the *y* direction (not shown) are very similar to those along the *x* direction, indicating that the strain-induced anisotropy has negligible effect on vacancy cluster formation. For the neutral vacancies we see virtually no dependence of the formation energy on the relative position of the defects. We attribute this finding to the following effect: due to the two-dimensional nature of the material and its small dielectric constant[26] the relaxation of the atoms around the defect is limited and, as in the isolated vacancy case, the relaxations are localized within 4 Å from the defect, leading to a small value for the elastic component of the defect interaction. Under 5% strain, the neutral vacancy pair formation energy increases almost uniformly by about 0.6 eV for all relative positions of the vacancies, which is very close to twice the 0.3 eV increase in formation energy of the isolated vacancy, see Figure 6-2(d). From these results we conclude that the elastic interaction between neutral vacancies is negligible. For the case of two charged vacancies ($q = -2$) the formation energies of pairs of vacancies on the same side of the sheet (0, $n_{x/y}$), n = 1, 2, 3 are very close to each other; the formation energy under strain decreases slightly, similar to the isolated vacancy case, see Figure 6-2(d). For $q = -1$ there is a slight increase in pair formation energy with strain.

An interesting quantity is the difference of the pair formation energy from the energy of two isolated vacancies, $\Delta E_f(q)$, which is shown in Figure 6-3(c). For the neutral ($q = 0$) and

charged ($q = -1$) cases the pair formation energy is slightly (up to 0.1 eV) larger than the energy of two isolated vacancies. For $q = -2$, the difference is more pronounced, a fact that we attribute to the higher electrostatic energy of the defect-induced charge.

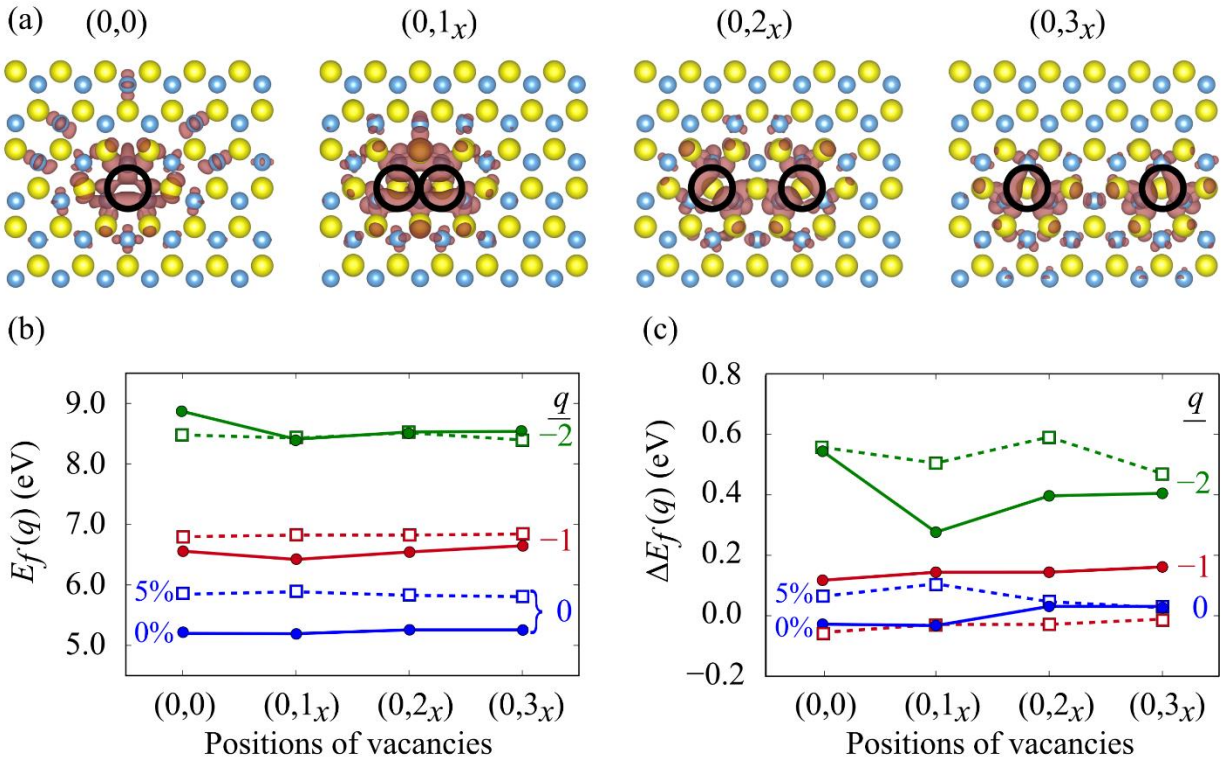

Figure 6-3. Clustering of sulfur vacancies.

(a) Defect wavefunctions for the pairs of S vacancies in the *x* direction for various relative positions; black circles highlight the vacancy positions.

(b) Formation energies $E_f(q)$ for vacancy pairs in charge states $q = 0, -1, -2$ along the *x* direction for 0% (solid line, filled circles) and 5% (dashed line, open squares) tensile strain along the *y* direction.

(c) Difference in formation energies of a pair of V_S relative to the energy of two isolated S vacancies, $\Delta E_f(q)$.

Moreover, the distances between the Mo atoms around the (0,0) vacancy for $q = -2$ are 0.05 Å larger than in the neutral state, consistent with the occupation of defect-induced

antibonding levels. In all charge states, the formation energies of pairs in configurations $(0,1_x)$, $(0,2_x)$, $(0,3_x)$ are very close to each other indicating no thermodynamic driving force for the clustering of S vacancies.

6.5. Diffusion energy barriers

We have also considered the possibility of diffusion of the S vacancies. We use the climbing image nudged elastic band method (CI-NEB)[40] for the computation of the activation energies for diffusion of the S vacancy between adjacent sites. We find that the barrier for isolated vacancy diffusion in the neutral ($q = 0$) state is 2.24 eV, a very large barrier for a thermally activated process. We investigated if the diffusion can be facilitated by other factors, such as applied strain, the presence of other vacancies nearby, or the charge state of the defect. We find that the barrier decreases with applied strain, to 1.95 eV for 5% uniaxial tensile strain, see Figure 6-4(a); this result can be rationalized by considering that the strained material can more easily accommodate the lattice relaxation related to defect diffusion. The presence of another vacancy nearby, that is, diffusion from position $(0,1_x)$ to $(0,2_x)$ in the notation of Figure 6-3(a), does not change the barrier appreciably, see Figure 6-4(b). Finally, the diffusion barrier in the negatively charged ($q = -1$) state is lowered by 0.3 eV compared to the neutral case, for both the isolated vacancy and a pair of vacancies. This result is in line with our previous discussion of charge-induced lattice softening caused by occupation of the antibonding levels in the band gap. Overall, we find that the diffusion of sulfur vacancies is a thermodynamically hindered process, as confirmed by recent experimental work indicating a vacancy jump frequency of 1 per 40 s.[41] Taking into account the absence of a driving force for clustering of the vacancies, we find that the formation of extended defects is an unlikely process barring strong external perturbation, like electron beam irradiation.[41-42]

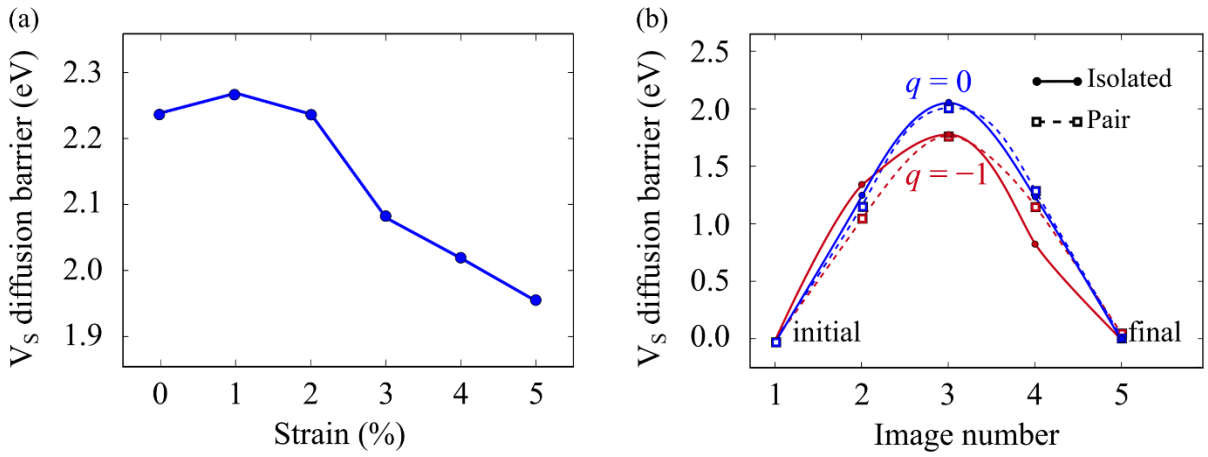

Figure 6-4. Energy barriers for S vacancy diffusion.

(a) Diffusion barrier as a function of strain applied along the y direction.

(b) Minimal energy path for the diffusion of an isolated vacancy (solid line) and the diffusion of a vacancy between the positions 1_x and 2_x in the presence of another vacancy at position 0 (dashed line) for the neutral (blue, $q = 0$) and charged (red, $q = -1$) states.

6.6. Summary

In summary we have reported a comprehensive examination of the properties of S vacancies, the most prevalent defects in monolayer MoS_2, using first-principles calculations. We find that the formation energy of the neutral S vacancy increases under uniaxial strain in the y direction, and for the negatively charged vacancy it decreases slightly; we attribute this finding to the fact that defect-induced states in the gap have antibonding character and their occupation leads to lattice softening. Our results also indicate that defect-induced lattice reorganization is very localized and there is no elastic interaction between two adjacent vacancies in either the neutral or the charged state. Accordingly, there is no thermodynamic driving force for the clustering of S vacancies and the barrier for the diffusion of vacancies is high, larger than 1.95 eV, in all cases considered.

It is interesting to examine these findings in the context of the recently reported "ripplocation" structure in exfoliated MoS_2 and its luminescence properties. Firstly, the large formation energy of isolated S vacancies and pair of vacancies, as well as the large barrier for the diffusion, strongly suggest that such defects are unlikely to form under thermodynamic equilibrium conditions. Secondly, the presence of tensile strain, as is likely the case in large deformations such as the ripplocation, does not change the formation energy or diffusion barrier of S vacancies substantially to alter their equilibrium properties. Another important factor in the ripplocation geometry can be the bending of the monolayer, which we will consider explicitly in future studies. Therefore, if S vacancies are present they will have to be introduced by external factors, like large forces during exfoliation. Moreover, if vacancies are present, their luminescence properties will be affected by the local strain.

We can estimate the energy of the photoluminescence peaks from our results as the energy difference between the lowest-energy defect state in the band gap, φ_y and the conduction band minimum. In the unstrained material this difference is 0.46 eV; in the material under strain this energy difference is 0.31 eV. In order to provide a better estimate for the photoluminescence peaks energy we rescale our DFT band gap of 1.6 eV to match the GW gap of MoS_2 monolayer of 2.8 eV,[35] as we consider the transition between the defect level and the conduction band without taking into account excitonic effects. Since the defect levels have Mo d-character, we scale the position of the defect level relative to the conduction band minimum in proportion to the band gap; after such scaling we infer that the approximate positions of defect-induced photoluminescence peaks are 0.81 eV for the unstrained material and 0.54 eV under 5% strain, representing a decrease of 0.27 eV after applying strain. We note that more accurate results can be obtained from time-dependent DFT simulations, which will be a subject of future work.

6.7. References

1. Cheiwchanchamnangij, T.; Lambrecht, W. R. L., Quasiparticle Band Structure Calculation of Monolayer, Bilayer, and Bulk MoS_2. *Physical Review B* **2012**, *85*, 205302.

2. Eda, G.; Yamaguchi, H.; Voiry, D.; Fujita, T.; Chen, M. W.; Chhowalla, M., Photoluminescence from Chemically Exfoliated MoS_2. *Nano Letters* **2011**, *11*, 5111-5116.

3. Rivera, P.; Seyler, K. L.; Yu, H. Y.; Schaibley, J. R.; Yan, J. Q.; Mandrus, D. G.; Yao, W.; Xu, X. D., Valley-Polarized Exciton Dynamics in a 2D Semiconductor Heterostructure. *Science* **2016**, *351*, 688-691.

4. Kuc, A.; Zibouche, N.; Heine, T., Influence of Quantum Confinement on the Electronic Structure of the Transition Metal Sulfide TS_2. *Physical Review B* **2011**, *83*, 245213.

5. Chhowalla, M.; Shin, H. S.; Eda, G.; Li, L. J.; Loh, K. P.; Zhang, H., The Chemistry of Two-Dimensional Layered Transition Metal Dichalcogenide Nanosheets. *Nat Chem* **2013**, *5*, 263-275.

6. Daage, M.; Chianelli, R. R., Structure-Function Relations in Molybdenum Sulfide Catalysts - the Rim-Edge Model. *J Catal* **1994**, *149*, 414-427.

7. Chang, K.; Chen, W. X., L-Cysteine-Assisted Synthesis of Layered MoS_2/Graphene Composites with Excellent Electrochemical Performances for Lithium Ion Batteries. *ACS Nano* **2011**, *5*, 4720-4728.

8. Greeley, J.; Jaramillo, T. F.; Bonde, J.; Chorkendorff, I. B.; Norskov, J. K., Computational High-Throughput Screening of Electrocatalytic Materials for Hydrogen Evolution. *Nat. Mater.* **2006**, *5*, 909-913.

9. Laursen, A. B.; Kegnaes, S.; Dahl, S.; Chorkendorff, I., Molybdenum Sulfides-Efficient and Viable Materials for Electro - and Photoelectrocatalytic Hydrogen Evolution. *Energy & Environmental Science* **2012**, *5*, 5577-5591.

10. Li, T. S.; Galli, G. L., Electronic Properties of MoS_2 Nanoparticles. *Journal of Physical Chemistry C* **2007**, *111*, 16192-16196.

11. Li, Y. G.; Wang, H. L.; Xie, L. M.; Liang, Y. Y.; Hong, G. S.; Dai, H. J., MoS_2 Nanoparticles Grown on Graphene: An Advanced Catalyst for the Hydrogen Evolution Reaction. *Journal of the American Chemical Society* **2011**, *133*, 7296-7299.

12. Merki, D.; Fierro, S.; Vrubel, H.; Hu, X. L., Amorphous Molybdenum Sulfide Films as Catalysts for Electrochemical Hydrogen Production in Water. *Chemical Science* **2011**, *2*, 1262-1267.

13. Zong, X.; Yan, H. J.; Wu, G. P.; Ma, G. J.; Wen, F. Y.; Wang, L.; Li, C., Enhancement of Photocatalytic H_2 Evolution on Cds by Loading MoS_2 as Cocatalyst under Visible Light Irradiation. *Journal of the American Chemical Society* **2008**, *130*, 7176-7177.

14. Radisavljevic, B.; Radenovic, A.; Brivio, J.; Giacometti, V.; Kis, A., Single-Layer MoS_2 Transistors. *Nat Nanotechnol* **2011**, *6*, 147-150.

15. Liu, L. T.; Kumar, S. B.; Ouyang, Y.; Guo, J., Performance Limits of Monolayer Transition Metal Dichalcogenide Transistors. *Ieee T Electron Dev* **2011**, *58*, 3042-3047.

16. Yoon, Y.; Ganapathi, K.; Salahuddin, S., How Good Can Monolayer MoS_2 Transistors Be? *Nano Letters* **2011**, *11*, 3768-3773.

17. Zhang, Y. J.; Ye, J. T.; Matsuhashi, Y.; Iwasa, Y., Ambipolar MoS_2 Thin Flake Transistors. *Nano Letters* **2012**, *12*, 1136-1140.

18. Kim, H.-S., et al., Lead Iodide Perovskite Sensitized All-Solid-State Submicron Thin Film Mesoscopic Solar Cell with Efficiency Exceeding 9%. *Sci Rep* **2012**, *2*, 591-591.

19. Lee, K.; Kim, H. Y.; Lotya, M.; Coleman, J. N.; Kim, G. T.; Duesberg, G. S., Electrical Characteristics of Molybdenum Disulfide Flakes Produced by Liquid Exfoliation. *Advanced Materials* **2011**, *23*, 4178-4182.

20. Pu, J.; Yomogida, Y.; Liu, K. K.; Li, L. J.; Iwasa, Y.; Takenobu, T., Highly Flexible Mos2 Thin-Film Transistors with Ion Gel Dielectrics. *Nano Letters* **2012**, *12*, 4013-4017.

21. Radisavljevic, B.; Whitwick, M. B.; Kis, A., Integrated Circuits and Logic Operations Based on Single-Layer MoS_2. *ACS Nano* **2011**, *5*, 9934-9938.

22. Wang, H.; Yu, L. L.; Lee, Y. H.; Shi, Y. M.; Hsu, A.; Chin, M. L.; Li, L. J.; Dubey, M.; Kong, J.; Palacios, T., Integrated Circuits Based on Bilayer MoS_2 Transistors. *Nano Letters* **2012**, *12*, 4674-4680.

23. Komsa, H. P.; Krasheninnikov, A. V., Native Defects in Bulk and Monolayer MoS_2 from First Principles. *Physical Review B* **2015**, *91*, 125304.

24. Kushima, A.; Qian, X. F.; Zhao, P.; Zhang, S. L.; Li, J., Ripplocations in Van Der Waals Layers. *Nano Letters* **2015**, *15*, 1302-1308.

25. Giannozzi, P., et al., Quantum Espresso: A Modular and Open-Source Software Project for Quantum Simulations of Materials. *J. Phys.-Condes. Matter* **2009**, *21*, 395502.

26. Wieting, T. J.; Verble, J. L., Infrared and Raman Studies of Long-Wavelength Optical Phonons in Hexagonal MoS_2. *Physical Review B* **1971**, *3*, 4286-4292.

27. Vinichenko, D.; Sensoy, M. G.; Friend, C. M.; Kaxiras, E., Accurate Formation Energies of Charge Defects in Solids: A Systematic Approach. *Submitted*.

28. Mak, K. F.; Lee, C.; Hone, J.; Shan, J.; Heinz, T. F., Atomically Thin MoS_2: A New Direct-Gap Semiconductor. *Physical Review Letters* **2010**, *105*, 136805.

29. Han, S. W., et al., Band-Gap Expansion in the Surface-Localized Electronic Structure of MoS_2 (0002). *Physical Review B* **2012**, *86*, 115105.

30. Splendiani, A.; Sun, L.; Zhang, Y. B.; Li, T. S.; Kim, J.; Chim, C. Y.; Galli, G.; Wang, F., Emerging Photoluminescence in Monolayer MoS_2. *Nano Letters* **2010**, *10*, 1271-1275.

31. Lee, C.; Yan, H.; Brus, L. E.; Heinz, T. F.; Hone, J.; Ryu, S., Anomalous Lattice Vibrations of Single- and Few-Layer MoS_2. *ACS Nano* **2010**, *4*, 2695-2700.

32. Beal, A. R.; Hughes, H. P., Kramers-Kronig Analysis of the Reflectivity Spectra of 2h-MoS_2, 2h-$MoSe_2$ and 2h-$MoTe_2$. *Journal of Physics C-Solid State Physics* **1979**, *12*, 881-890.

33. Zhu, Z. Y.; Cheng, Y. C.; Schwingenschlogl, U., Giant Spin-Orbit-Induced Spin Splitting in Two-Dimensional Transition-Metal Dichalcogenide Semiconductors. *Physical Review B* **2011**, *84*, 153402.

34. Komsa, H. P.; Krasheninnikov, A. V., Effects of Confinement and Environment on the Electronic Structure and Exciton Binding Energy of MoS_2 from First Principles. *Physical Review B* **2012**, *86*, 241201.

35. Ramasubramaniam, A., Large Excitonic Effects in Monolayers of Molybdenum and Tungsten Dichalcogenides. *Physical Review B* **2012**, *86*, 115409.

36. Marzari, N.; Mostofi, A. A.; Yates, J. R.; Souza, I.; Vanderbilt, D., Maximally Localized Wannier Functions: Theory and Applications. *Reviews of Modern Physics* **2012**, *84*, 1419-1475.

37. Fang, S. A.; Defo, R. K.; Shirodkar, S. N.; Lieu, S.; Tritsaris, G. A.; Kaxiras, E., Ab Initio Tight-Binding Hamiltonian for Transition Metal Dichalcogenides. *Physical Review B* **2015**, *92*, 205108.

38. Johari, P.; Shenoy, V. B., Tuning the Electronic Properties of Semiconducting Transition Metal Dichalcogenides by Applying Mechanical Strains. *ACS Nano* **2012**, *6*, 5449-5456.

39. Bertolazzi, S.; Brivio, J.; Kis, A., Stretching and Breaking of Ultrathin MoS_2. *ACS Nano* **2011**, *5*, 9703-9709.

40. Henkelman, G.; Uberuaga, B. P.; Jonsson, H., A Climbing Image Nudged Elastic Band Method for Finding Saddle Points and Minimum Energy Paths. *Journal of Chemical Physics* **2000**, *113*, 9901-9904.

41. Komsa, H. P.; Kurasch, S.; Lehtinen, O.; Kaiser, U.; Krasheninnikov, A. V., From Point to Extended Defects in Two-Dimensional MoS_2: Evolution of Atomic Structure under Electron Irradiation. *Physical Review B* **2013**, *88*, 035301.

42. Wang, S. S.; Lee, G. D.; Lee, S.; Yoon, E.; Warner, J. H., Detailed Atomic Reconstruction of Extended Line Defects in Monolayer MoS_2. *ACS Nano* **2016**, *10*, 5419-5430.

CHAPTER 7. MODELING THE EFFICIENCY OF PEROVSKITE-BASED SOLAR CELLS

In this chapter we establish a framework for bridging the properties of materials established from first-principles simulations to the performance of devices based on them, using a novel class of solar cells with hybrid organic-inorganic perovskite absorber as an example. This type of solar cell has seen a seven-fold increase of efficiency, from ~3% in early devices to over 20% in research prototypes.[1] Oft-cited reasons for this increase are: (i) a carrier diffusion length reaching hundreds of microns;[2] (ii) a low exciton binding energy;[3-4] and (iii) a high optical absorption coefficient. These hybrid organic-inorganic materials span a large chemical space with the perovskite structure. Here, using first-principles calculations and thermodynamic modelling, we establish that, given the range of band-gaps of the metal-halide-perovskites, the theoretical maximum efficiency limit is in the range of ~25-27%. Our conclusions are based on the effect of level alignment between the perovskite absorber layer and carrier-transporting materials on the performance of the solar cell as a whole. Our results provide a useful framework for experimental searches toward more efficient devices. This chapter is based on the work that is submitted as a publication: O. Granas, D.Vinichenko, E. Kaxiras. Establishing the limits of efficiency of perovskite solar cells from first-principles modeling.

7.1. Introduction

Photovoltaic applications of metal halide perovskite absorbers face a number of outstanding challenges, including materials stability, hysteresis effects of the current-voltage characteristics, and fine-tuning of the absorption properties.[5-6] The perovskite ABX_3 structure, where A and B are typically organic or inorganic cations, and X is an oxygen or halogen anion,

offers a large chemical phase space, allowing many properties to be tailored, albeit not independently. We argue that reaching the ideal efficiency, apart from optimizing the quality of materials and interfaces, is more challenging than optimizing the photo absorber; when absorption properties are tailored by adjusting the composition, band alignment and effective masses are also affected.

To establish the theoretical limits of efficiency, we assume ideal interfaces and defect-free crystals. We investigate how the presence of different ions on the A, B and X site of the perovskite structure impacts the photovoltaic performance, using electronic structure calculations and available experimental information, together with thermodynamic considerations. We show that the level alignment to the electron- and hole-transporting media (ETM and HTM) is central to reaching maximum efficiency for heterogeneous positive-intrinsic-negative (PIN) junctions. In fact, the efficiency limit for many of the perovskites is similar, in the ~25-27% range, given perfect band-alignment to ETM and HTM. Current state-of-the-art cells reach >80% of the theoretical maximum efficiency, indicating that higher performance is mostly a matter of interface engineering and the construction of multi-junction cells. Our results shed light on the performance differences arising from different perovskite compositions, choices of electron and hole transporting media or modification of the heterojunctions.[7-8] Thus, they serve as a guide to further work on what HTM and ETM are suitable for optimal device performance.

7.2. Model of the solar cell

We aim to establish theoretical limits for the power conversion efficiency (PCE) of the perovskite-based solar cells as a function of chemical composition of the perovskite layer and the electronic properties of the electron- and hole-transporting media, by using the thermodynamic approach pioneered by Shockley and Queisser.[9] Recent experiments indicate that exotic effects

such as ferroelectricity are not responsible for the reported high efficiencies.[10-11] The most important properties of the perovskites as photo-absorbers are the resilience to form recombination centers and the reasonable interface quality to many ETMs and HTMs. Thus, it is reasonable to determine efficiency limits of perovskite solar cells from arguments based on detailed balance. Experimental characterization of the heterojunctions suggests that perovskite solar cells are of PIN character,[12-13] where the perovskite itself comprises the central intrinsic semiconductor. We model the cell as a PIN heterojunction, where the carriers are generated by photon absorption in the intrinsic perovskite layer, separated and injected across a corresponding interface into the electron- or hole-transporting medium. We model interfaces between the perovskite and carrier transporting media as diodes $D_{ETM/HTM}$ with the constant voltage drop (CVD) approximation, that is, we ignore the band matching character. We also ignore back transfer of carriers, which has been experimentally proven a slow process.[14] We present in Figure 7-1(b) a schematic of the circuit. We assume a constant-entropy mode of operation; with that, the electrochemical potential of the cell can be expressed as:

$$\Delta\mu = eV_{ext} = E_G - k_B T \ln\left[\frac{2\pi k_B T}{h^2}\frac{4(m_e^* m_h^*)^{3/2}}{n_e n_h}\right] - \Delta E_{HTM} - \Delta E_{ETM} \qquad (7.2.1)$$

where E_G is the perovskite absorber band gap, $m_e^*,\ m_h^*$ are carrier effective masses, $n_e,\ n_h$ are carrier densities, and $\Delta E_{HTM},\ \Delta E_{ETM}$ are the potential drops at the boundaries between the perovskite layer and carrier-transporting media. The energy levels in our model are shown in Figure 7-1(c). We ignore shunt and series resistances, and consider only radiative recombination to provide the upper limit for experimental efforts. With these assumptions, the current is given by $J = e\left(N_{ph} - RR(V_{ext})\right)$, where N_{ph} is the density of incident photons, and $RR(V_{ext})$ is the rate of radiative recombination:

$$RR(V_{ext}) = e\frac{2\pi}{c^2h^3}\exp\left({eV_{ext}}/{k_BT}\right)\int_{E_G}^{\infty}\frac{E^2dE}{\exp\left({E}/{k_BT}\right)-1} \tag{7.2.2}$$

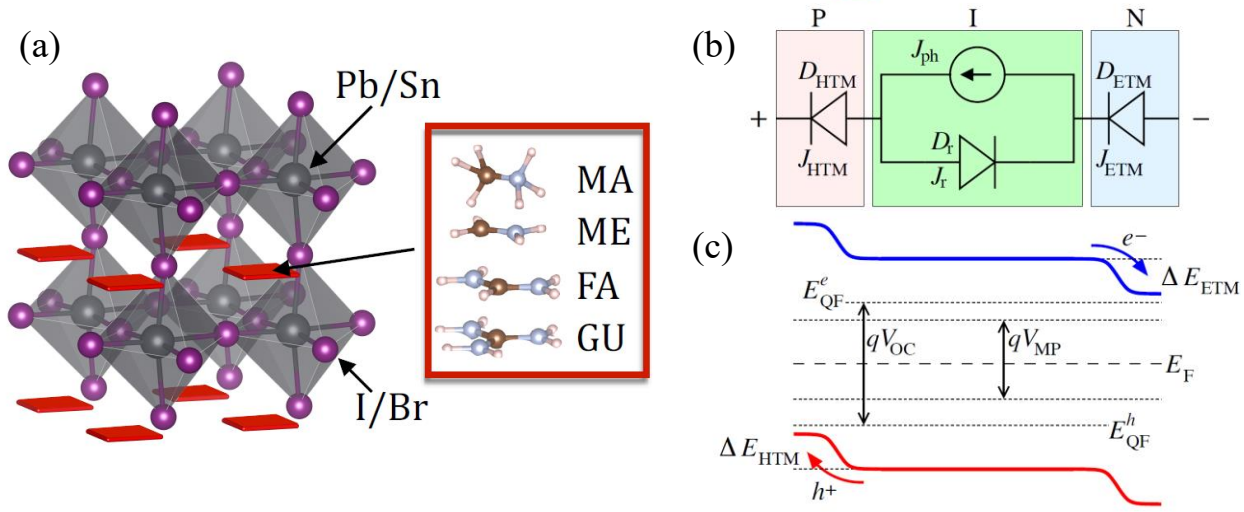

Figure 7-1. Perovskite-based solar cell.

(a) The crystal structure and composition of perovskites under consideration (in the organic cation, blue – N, brown – C, pink – H).

(b) Equivalent circuit of the PIN photo absorbing heterostructure. J_{ph} denotes the photon induced current, D_r the intrinsic diode (additional diode currents from Shockley-Reed-Hall recombination can be added), $D_{ETM/HTM}$represent rectifying diodes, the diode voltage drop, as well as forward direction is determined by the energy level alignment.

(c) A schematic of the energy level diagram of the intrinsic part, with band offsets ΔE_{ETM}, ΔE_{HTM} to ETM/HTM. E_F is the position of the Fermi level, E^h_{QF}, E^e_{QF} the quasi Fermi-levels under illumination with open circuit, V_{MP} the voltage for the maximum power point and V_{OC} the open-circuit voltage.

In order to compute the PCE, we use the NREL reference AM1.5 spectrum normalized to 1 kW/m^2. The maximal extractable power density is determined by maximizing $P = V_{ext}J(V_{ext})$. From these considerations, we can identify three main components that determine the efficiency of the perovskite-based solar cell: the band gap of the absorber, the carrier characteristics

relevant for entropic contribution, and the alignment between absorber and carrier transporting media. In the following we discuss our approach to determining these quantities using a combination of first-principles modelling and available experimental data.

The band gap of the perovskites is generally direct, or close to direct.[15] Together with band edge characters that allow for dipolar transitions this leads to an extraordinarily high optical absorption coefficient, which allows for a thin-film cell architecture. The onset of absorption is determined by the band gap, which is crucial for estimating the fraction of absorbed photon flux. In the context of first-principles electronic structure calculations with density functional theory (DFT), due to incorrect Coulomb interaction asymptotes in semi-local functionals,[16] the band gap is underestimated; for the hybrid organic-inorganic perovskites the error between different DFT calculations is fairly constant (for instance, the difference between the gaps of $MASnI_3$ and $MAPbI_3$ is only 0.07 eV, as inferred by comparison with GW results).[17-19] In order to obtain realistic band gaps for the compounds studied here, we use $MAPbI_3$ (for which reliable experimental data is available) as the reference material and calculate the difference in the DFT computed gap with all other compounds. Experimentally, the photo-induced excitons dissociate rapidly into free carriers due to low exciton binding energy of 10-50 meV;[3-4] for this reason it is not necessary to model excitonic effects.

For the DFT simulations, we use the PBE[20] exchange-correlation functional with van der Waals corrections,[21-22] as implemented in VASP.[23-26] We also employ the HSE hybrid density functional[27] for an independent estimate of the shifts in band gaps. We use a Γ-centered k-point grid of 8×8×8 and a plane-wave energy cutoff of 550 eV. The band gap of the perovskites under consideration ranges between 1.2 eV and 2.1 eV (in PBE) or between 0.9 eV and 2.3 eV (in HSE).

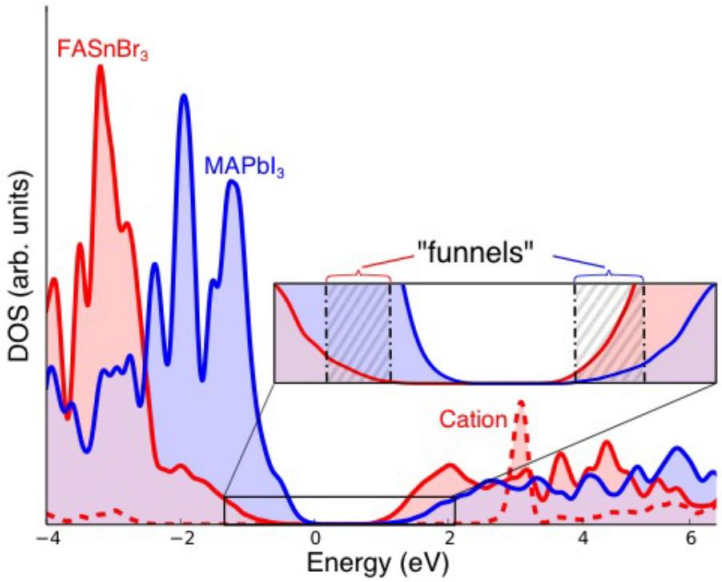

Figure 7-2. Perovskite electronic structure.

Densities of states for methylammonium lead iodide ($MAPbI_3$) and formamidinium tin bromide ($FASnBr_3$), showing contributions from the organic cation and the inorganic sublattice. The band edges are spanned by metal and halogen states, with the cation having only minor hybridization. The inset shows the shallow DOS ("funnel") near the VBM, characteristic for tin-based compounds, and shallow DOS near the CBM, typical of Pb-based compounds.

Charge mobility is a very important factor for efficient solar cells and is key to the high reported efficiencies. Long mean free paths have been observed in experiments: early work measured over 1 μm for polycrystalline samples[13] while more recent work reported values up to 175 μm under illumination and 3 mm under weak light[2]. Apart from the diffusion length, the low carrier masses decrease the entropic loss of the open-circuit voltage. Small masses are manifested as funnel-like structures on the electronic density of states, with a minimal number of states at the band edges. The perovskite compounds have a pronounced funnel effect (see Figure 7-2), which contributes to the high cell efficiency at operating temperatures. We compute the average of the effective mass tensor of the carriers from the band structure using a high density

(approximately $2 \cdot 10^7$ k-points/Å^{-3}) mesh. Finally, we determine the entropy contribution of the excited carriers to the quasi Fermi level splitting using an effective density of excited states of 10^{15} cm^{-3}.[28] The carrier masses of the compounds under consideration are smaller than, for example, Si which results in a smaller effect of the temperature arising from carrier entropy and a smaller deviation from the Shockley-Queisser limit.[29] The results show a clear trend: Pb compounds have lower conduction band mass and higher valence band mass in relation to the Sn counterparts. The effective masses range between 0.1 and 0.3 in units of m_e.

Properties like the natural level alignment and the absorption onset are difficult to measure from an experimental point of view. Interface and surface dipoles, defects and sample inhomogeneities result in substantial differences in available experimental data. For example, the depletion layer is reported to be somewhere between 45 to 300 nm,[30-31] indicating that level alignments play a crucial role for the carrier concentrations and for the potential gradient in the perovskite. The calculation of the natural band alignment from first principles using core levels or by inspecting how the average electrostatic potential changes on a site in different environments has been treated extensively in the literature.[32-33] We employ a method of alignment similar to what was first suggested by Massidda et al.,[34] and later refined by Wei and Zunger,[35] and by Lang et al.[36] We estimate the relative positions of valence and conduction bands of perovskites with different composition using the average electrostatic potential on the B site ion, or, in the case of B ion substitution, on the X site ion. We also use the equivalent formalism of core-levels as a test for consistency. The band alignments, from the relative position of band edges to those of $MAPbI_3$, are used to map the corresponding potential drops for the rectifying diodes appearing at the heterojunctions between the perovskite and the charge-transporting media. We do not explicitly model the P and N materials; instead, we map out the

efficiency as a function of the natural band alignment with respect to $MAPbI_3$. Recent computational work provided estimates for the potential alignment of a number of common semiconductors in relation to $MAPbI_3$; with our results, this can be used directly to estimate how the efficiency is altered.[37]

$MAPbI_3$ has a cubic structure at high temperature and undergoes a transition to a tetragonal structure close to room temperature. The two phases are known to co-exist to much lower temperatures. To reduce the complexity of calculating the level alignments, we concentrate on the cubic phase of the perovskite cell. This involves calculating: the core-level (CL) alignment in the super cell between the left (L) and right (R) perovskites; the CL alignment to valence band maxima (VBM) in the strained materials (with strain according to the relaxed heterostructure); the CL alignment to VBM in the relaxed structure. To account for possible steric effects from the orientation of the organic cations, we consider them oriented in the plane orthogonal to the super cell stacking, which resembles the tetragonal phase. This procedure allows us to determine the natural band alignment between $MAPbI_3$ and the other perovskites under consideration. Reviewing other theoretical approaches reveals a significant spread in reported data,[38] but the trends are consistent with our findings. Previous work on GaAs/AlAs heterojunctions estimated errors in absolute value on the order of 0.05 eV.[35] For the calculations reported here, we expect slightly larger errors due to the complex symmetry breaking by the organic cation. Previous studies using similar methods but without taking into account strain effects reported similar values of band alignment for the compounds under consideration.[39]

In cases where an increase (decrease) in the potential for electrons (holes) occurs, the voltage drop over the junction depends exponentially on the barrier height, with most of the external bias over the junction. In these cases, the quasi Fermi-level splitting goes to zero,

inducing strong recombination in the junction, and the value of J_{sc} effectively goes to zero as well. An accurate treatment of the junctions requires knowledge of the dielectric properties and density of states of specific HTMs and ETMs.

7.3. Results and discussion

We have investigated 16 hybrid organic-inorganic compounds with the perovskite structure of the general composition ABX_3. For A, we have considered 4 organic cations, the traditionally used methylammonium, and three other ions of similar size and conjugated π system of increasing size: methyleneiminium ($ME = CH_2NH_2^+$), formamidinium ($FA = CH(NH_2)_2^+$) and guanidinium ($GU = C(NH_2)_3^+$). We consider Pb^{2+} and Sn^{2+} ions for the B position, and Cl^- and Br^- ions for the X position. The most important feature of the electronic structure of these compounds is that, regardless of the size of the π system of the A cation, the band edges are spanned by *s*- and *p*-states of the B ion and by *p*-states of the X ion. The A ions act mostly as spacers and affect the electronic structure and properties of the perovskite through the changes they induce in lattice structure, as we discuss below.

The bottom of the conduction band is spanned mostly by the *p*-states of B ions with an admixture of *p*-states of X ions. There are two factors in action when the B site is changed: the relative electronegativity of Pb and Sn, and the strength of spin-orbit coupling, which leads to splitting of the bottom of the conduction band. Those factors lead to lowering of the absolute position of the CBM for Sn compared to Pb. Moreover, due to stronger spin-orbit coupling in Pb, the CBM splitting is larger, which leads to higher curvature of the CBM and, therefore, to lower effective masses for electrons in Pb-based perovskites (~0.1 m_e) compared to Sn-based perovskites (~0.2 – 0.3 m_e). With Br in the X position the CBM is higher due to the more covalent character of the bonds and larger splitting between bonding and antibonding orbitals.

The top of the valence band consists of an antibonding combination of X *p*-states and B *s*-states leading to a lower absolute position of the VBM with increasing electronegativity of the X ion. Substitution of Pb by Sn on the B site leads to an increase of the VBM energy due to larger overlap of the Sn *s*-states with the X *p*-orbitals. An increase in the size of the A cation leads to a decrease in the VBM energy due to the expansion of the lattice structure from the decrease in overlap between the *s*-orbitals of B and the *p*-orbitals of X, and the concomitant lowering of the antibonding level energy. Smaller overlap between the *s*-orbitals of B and the *p*-orbitals of X also leads to a less dispersive band and an increase in the effective carrier mass for the holes for Pb (~0.15–0.3 m_e) compared to Sn (~0.07–0.1 m_e). A similar effect is introduced by increasing the size of the A cation. Another factor is the Darwin term effect on the *s*-orbitals of B ion levels, which is larger for Pb, resulting in flatter bands.

From the absolute positions of the CBM and VBM we obtain the value of the band gap, which is in the 1.1 – 1.6 eV range for most compounds except for FA-, MA-, and GU-based perovskites with $PbBr_3$ lattice backbone. Based on our calculations, we conclude that all of the compounds studied (except for those mentioned above) demonstrate power conversion efficiency of 25-27% regardless of actual chemical composition, provided optimal band alignment is satisfied. This clearly demonstrates that the intrinsic properties of hybrid organic-inorganic compounds, that is, band gap and carrier effective masses, are not the limiting factor in determining the efficiency. Rather, it is the misalignment of the absolute positions of the band edges that can lead to substantial performance deterioration (about 5% PCE for every 0.2 eV of mismatch). We map out the band edge positions, which are equivalent to the optimal carrier-transporting material levels, and the PCE achieved with optimal level matching in Figure 7-3.

Our overall conclusion is that experimental efforts should be directed towards optimization of the device as a whole, using as a guide the results outlined above.

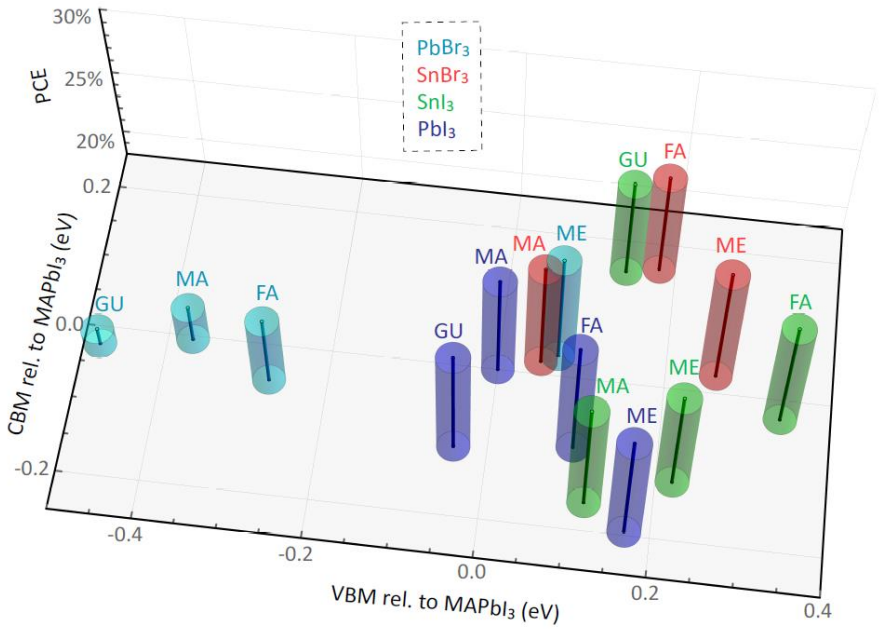

Figure 7-3. Cell efficiency as a function of level alignment in relation to $MAPbI_3$

Classes of compounds are colored according to their inorganic backbone: PbI_3 – purple, $PbBr_3$ – blue, SnI_3 – green, and $SnBr_3$ – red. The PCE determine the height of the cylinders. Note that the spread in VBM is much larger than the spread in CBM.

In particular, the chemical composition of the perovskite has a profound effect on the position of VBM and CBM levels: an increase in the size of the A ion leads to lowering of the VBM; compounds with Pb on the B site have lower VBM and higher CBM compared to Sn-based perovskites; Br-based compounds have lower VBM and higher CBM compared to I-based ones. These principles can be used to tailor the composition of the perovskite absorber in order to optimize the efficiency of the cell overall by ensuring optimal level matching with the carrier-transporting materials.

7.4. References

1. Chen, W., et al., Efficient and Stable Large-Area Perovskite Solar Cells with Inorganic Charge Extraction Layers. *Science* **2015**, *350*, 944-948.

2. Dong, Q. F.; Fang, Y. J.; Shao, Y. C.; Mulligan, P.; Qiu, J.; Cao, L.; Huang, J. S., Electron-Hole Diffusion Lengths > 175 μm in Solution-Grown $CH_3NH_3PbI_3$ Single Crystals. *Science* **2015**, *347*, 967-970.

3. Galkowski, K., et al., Determination of the Exciton Binding Energy and Effective Masses for Methylammonium and Formamidinium Lead Tri-Halide Perovskite Semiconductors. *Energy & Environmental Science* **2016**, *9*, 962-970.

4. Miyata, A.; Mitioglu, A.; Plochocka, P.; Portugall, O.; Wang, J. T. W.; Stranks, S. D.; Snaith, H. J.; Nicholas, R. J., Direct Measurement of the Exciton Binding Energy and Effective Masses for Charge Carriers in Organic-Inorganic Tri-Halide Perovskites. *Nat Phys* **2015**, *11*, 582-587.

5. Berry, J., et al., Hybrid Organic-Inorganic Perovskites (HOIPs): Opportunities and Challenges. *Advanced Materials* **2015**, *27*, 5102-5112.

6. Unger, E. L.; Hoke, E. T.; Bailie, C. D.; Nguyen, W. H.; Bowring, A. R.; Heumuller, T.; Christoforo, M. G.; McGehee, M. D., Hysteresis and Transient Behavior in Current-Voltage Measurements of Hybrid-Perovskite Absorber Solar Cells. *Energy & Environmental Science* **2014**, *7*, 3690-3698.

7. Wojciechowski, K., et al., Heterojunction Modification for Highly Efficient Organic - Inorganic Perovskite Solar Cells. *ACS Nano* **2014**, *8*, 12701-12709.

8. Stranks, S. D.; Snaith, H. J., Metal-Halide Perovskites for Photovoltaic and Light-Emitting Devices. *Nat Nanotechnol* **2015**, *10*, 391-402.

9. Shockley, W.; Queisser, H. J., Detailed Balance Limit of Efficiency of p-n Junction Solar Cells. *Journal of Applied Physics* **1961**, *32*, 510-519.

10. Meloni, S., et al., Ionic Polarization-Induced Current-Voltage Hysteresis in $CH_3NH_3PbX_3$ Perovskite Solar Cells. *Nat. Commun.* **2016**, *7*, 10334.

11. Sherkar, T. S.; Koster, L. J., Can Ferroelectric Polarization Explain the High Performance of Hybrid Halide Perovskite Solar Cells? *Phys Chem Chem Phys* **2016**, *18*, 331-338.

12. Bergmann, V. W.; Weber, S. A. L.; Ramos, F. J.; Nazeeruddin, M. K.; Gratzel, M.; Li, D.; Domanski, A. L.; Lieberwirth, I.; Ahmad, S.; Berger, R., Real-Space Observation of Unbalanced Charge Distribution inside a Perovskite-Sensitized Solar Cell. *Nat. Commun.* **2014**, *5*, 5001.

13. Edri, E.; Kirmayer, S.; Mukhopadhyay, S.; Gartsman, K.; Hodes, G.; Cahen, D., Elucidating the Charge Carrier Separation and Working Mechanism of $CH_3NH_3PbI_{3-x}Cl_x$ Perovskite Solar Cells. *Nat. Commun.* **2014**, *5*, 3461.

14. Marchioro, A.; Teuscher, J.; Friedrich, D.; Kunst, M.; van de Krol, R.; Moehl, T.; Gratzel, M.; Moser, J. E., Unravelling the Mechanism of Photoinduced Charge Transfer Processes in Lead Iodide Perovskite Solar Cells. *Nat. Photonics* **2014**, *8*, 250-255.

15. Motta, C.; El-Mellouhi, F.; Kais, S.; Tabet, N.; Alharbi, F.; Sanvito, S., Revealing the Role of Organic Cations in Hybrid Halide Perovskite $CH_3NH_3PbI_3$. *Nat Commun* **2015**, *6*, 7026.

16. Kummel, S.; Kronik, L., Orbital-Dependent Density Functionals: Theory and Applications. *Reviews of Modern Physics* **2008**, *80*, 3-60.

17. Umari, P.; Mosconi, E.; De Angelis, F., Relativistic GW Calculations on $CH_3NH_3PbI_3$ and $CH_3NH_3SnI_3$ Perovskites for Solar Cell Applications. *Sci Rep* **2014**, *4*, 4467.

18. Filip, M. R.; Eperon, G. E.; Snaith, H. J.; Giustino, F., Steric Engineering of Metal-Halide Perovskites with Tunable Optical Band Gaps. *Nat Commun* **2014**, *5*, 5757.

19. Frost, J. M.; Butler, K. T.; Brivio, F.; Hendon, C. H.; van Schilfgaarde, M.; Walsh, A., Atomistic Origins of High-Performance in Hybrid Halide Perovskite Solar Cells. *Nano Letters* **2014**, *14*, 2584-2590.

20. Perdew, J. P.; Burke, K.; Ernzerhof, M., Generalized Gradient Approximation Made Simple. *Physical Review Letters* **1996**, *77*, 3865-3868.

21. Bucko, T.; Lebegue, S.; Hafner, J.; Angyan, J. G., Improved Density Dependent Correction for the Description of London Dispersion Forces. *J Chem Theory Comput* **2013**, *9*, 4293-4299.

22. Tkatchenko, A.; Scheffler, M., Accurate Molecular Van Der Waals Interactions from Ground-State Electron Density and Free-Atom Reference Data. *Phys Rev Lett* **2009**, *102*, 073005.

23. Kresse, G.; Furthmuller, J., Efficient Iterative Schemes for Ab Initio Total-Energy Calculations Using a Plane-Wave Basis Set. *Physical Review B* **1996**, *54*, 11169-11186.

24. Kresse, G.; Hafner, J., Abinitio Molecular-Dynamics for Liquid-Metals. *Physical Review B* **1993**, *47*, 558-561.

25. Kresse, G.; Hafner, J., Norm-Conserving and Ultrasoft Pseudopotentials for First-Row and Transition-Elements. *J. Phys.-Condes. Matter* **1994**, *6*, 8245-8257.

26. Kresse, G.; Joubert, D., From Ultrasoft Pseudopotentials to the Projector Augmented-Wave Method. *Physical Review B* **1999**, *59*, 1758-1775.

27. Heyd, J.; Scuseria, G. E.; Ernzerhof, M., Hybrid Functionals Based on a Screened Coulomb Potential. *Journal of Chemical Physics* **2006**, *124*, 219906.

28. Osterloh, F. E., Maximum Theoretical Efficiency Limit of Photovoltaic Devices: Effect of Band Structure on Excited State Entropy. *J Phys Chem Lett* **2014**, *5*, 3354-3359.

29. Sha, W. E. I.; Ren, X. G.; Chen, L. Z.; Choy, W. C. H., The Efficiency Limit of $CH_3NH_3PbI_3$ Perovskite Solar Cells. *Applied Physics Letters* **2015**, *106*, 221104.

30. Dymshits, A.; Henning, A.; Segev, G.; Rosenwaks, Y.; Etgar, L., The Electronic Structure of Metal Oxide/Organo Metal Halide Perovskite Junctions in Perovskite Based Solar Cells. *Sci Rep* **2015**, *5*, 8704.

31. Jiang, C. S., et al., Carrier Separation and Transport in Perovskite Solar Cells Studied by Nanometre-Scale Profiling of Electrical Potential. *Nat Commun* **2015**, *6*, 8397.

32. Van de Walle, C. G., Band Lineups and Deformation Potentials in the Model-Solid Theory. *Phys Rev B Condens Matter* **1989**, *39*, 1871-1883.

33. Franciosi, A.; Van de Walle, C. G., Heterojunction Band Offset Engineering. *Surface Science Reports* **1996**, *25*, 1-140.

34. Massidda, S.; Min, B. I.; Freeman, A. J., Interface Phenomena at Semiconductor Heterojunctions - Local-Density Valence-Band Offset in GaAs/AlAs. *Physical Review B* **1987**, *35*, 9871-9874.

35. Wei, S. H.; Zunger, A., Calculated Natural Band Offsets of All II-VI and III-V Semiconductors: Chemical Trends and the Role of Cation *d* Orbitals. *Applied Physics Letters* **1998**, *72*, 2011-2013.

36. Lang, L.; Zhang, Y. Y.; Xu, P.; Chen, S. Y.; Xiang, H. J.; Gong, X. G., Three-Step Approach for Computing Band Offsets and Its Application to Inorganic ABX_3 Halide Perovskites. *Physical Review B* **2015**, *92*, 075102.

37. Butler, K. T.; Kumagai, Y.; Oba, F.; Walsh, A., Screening Procedure for Structurally and Electronically Matched Contact Layers for High-Performance Solar Cells: Hybrid Perovskites. *J Mater Chem C* **2016**, *4*, 1149-1158.

38. Mosconi, E.; Umari, P.; De Angelis, F., Electronic and Optical Properties of Mixed Sn-Pb Organohalide Perovskites: A First Principles Investigation. *Journal of Materials Chemistry A* **2015**, *3*, 9208-9215.

39. Butler, K. T.; Frost, J. M.; Walsh, A., Band Alignment of the Hybrid Halide Perovskites $CH_3NH_3PbCl_3$, $CH_3NH_3PbBr_3$ and $CH_3NH_3PbI_3$. *Mater Horiz* **2015**, *2*, 228-231.