Faculteit Wetenschappen
Departement Fysica

Automated in silico design of materials for energy and plasma applications

Geautomatiseerde in silico ontwikkeling van materialen voor energie- en plasmatoepassingen

Proefschrift voorgelegd tot het behalen van de graad van
doctor in de wetenschappen: fysica
aan de Universiteit Antwerpen te verdedigen door

Marnik Bercx

Promotors
Prof. dr. Dirk Lamoen
Prof. dr. Bart Partoens

Antwerpen, maart 2020

Abstract

Materials and their properties play a vital role in most applications we use on a daily basis. Many of the revolutions in industry are instigated by the discovery, by accident or design, of a new material that makes an application commercially viable. Historically, the study of materials has largely relied on an intuition-driven trial and error approach. However, considering the enormous design space of possible materials, as well as the fact that many of the straightforward materials have already been discovered, this process has become too expensive and time-consuming.

Since the middle of the 20th century, a new paradigm in materials science has been developing, where computer simulations are used to calculate the properties of materials from first principles. This new approach has steadily become more and more successful, pushed forward by the ever-increasing performance of modern computers and rapid progress in theoretical methods. During the past few decades, computational materials science has started evolving more and more into a predictive tool instead of simply offering theoretical insight into the physical processes of materials of interest. In combination with increasingly available tools for automating the required calculations, this has led to the concept of *in silico* materials design, where large numbers of compounds are investigated using computer simulations in order to gauge their potential for a specific application.

Among the most successful theoretical frameworks for computational materials science is density functional theory, which can determine the electronic structure of many compounds with ever increasing accuracy using a reasonable amount of computational resources. However, the connection between the electronic structure of a material and the property of interest for a specific application is rarely trivial. The main goal of this thesis is to provide or improve this connection, by analyzing existing metrics for flaws or anomalies, and developing new descriptors of material properties as well as the tools for calculating them using automated workflows. These methods are then applied to a set of topics, including solar cells, Li-ion batteries and ion-induced secondary electron emission. The structure of the thesis is as follows:

Chapter 1 briefly introduces the concept of *in silico* materials design, and provides a guide to the reader of this thesis for navigating and consulting the available resources.

Chapter 2 explains the density functional theory framework, as well as some practical computational techniques for calculating the electronic structure using this framework.

Chapter 3 outlines the workflows used for the automation of the required density functional theory calculations of each descriptor or metric.

Chapter 4 discusses the Shockley-Queisser limit and spectroscopic limited maximum efficiency, two metrics used to determine the potential of a material as the absorber layer of a single-junction solar cell. Next, it makes a comparison of the CuAu-like and chalcopyrite phase in the context of thin-film photovoltaics.

Chapter 5 presents an investigation of the stability of the oxygen framework of Li-rich Li_2MnO_3 and Li_2IrO_3 battery cathodes, as well as a limited substitution of Mn as a potential recipe for improving the structural stability of these materials. Moreover, it discusses the energy landscapes of $LiCB_{11}H_{12}$ and $NaCB_{11}H_{12}$ polyborane salts in the context of solid electrolytes.

Chapter 6 discloses a new model for calculating the secondary electron emission yield from ions neutralized at a semiconductor and metal surface, and applies this descriptor to a set of elemental surfaces spanning the periodic table in a high-throughput approach.

Abstract (Nederlands)

Materialen en hun eigenschappen spelen een cruciale rol in vele toepassingen die we dagelijks gebruiken. Veel van de revoluties in de industrie worden veroorzaakt door de ontdekking, per ongeluk of ontwerp, van een nieuw materiaal. De conventionele studie van materialen is grotendeels gebaseerd op een intuïtiegestuurde *trial and error* aanpak, maar gezien de enorme ontwerpruimte van mogelijke materialen is dit proces te duur en tijdrovend.

Sinds het midden van de 20^{e} eeuw is een nieuw paradigma in de materiaalwetenschap aan het opkomen, waarbij computersimulaties worden gebruikt om de eigenschappen van materialen *ab initio* te berekenen. In combinatie met de steeds toenemende performantie van computers en de snelle vooruitgang in theoretische methoden, biedt computationeel materiaalonderzoek meer en meer een betrouwbare manier om de eigenschappen van materialen te voorspellen. Dit heeft geleid tot het concept van *in silico* materiaalontwerp, waarbij een groot aantal materialen wordt onderzocht met behulp van computersimulaties om hun potentieel voor een specifieke toepassing na te gaan.

Een van de meest succesvolle theoretische kaders voor computationeel materiaalonderzoek is dichtheidsfunctionaaltheorie, waarmee de elektronische structuur van veel materialen met steeds toenemende nauwkeurigheid kan bepaald worden. Het verband tussen de elektronische structuur van een materiaal en de eigenschap die ons interesseert voor een specifieke toepassing is echter zelden triviaal. Het hoofddoel van dit proefschrift is om deze verbinding te bieden of te verbeteren, door bestaande metrieken voor fouten of afwijkingen te analyseren en nieuwe descriptoren van materiaaleigenschappen te ontwikkelen. Daarnaast worden deze methoden toegepast op een reeks onderwerpen, waaronder zonnecellen, Li-ion batterijen en secundaire elektronenemissie door neutralisatie van een ion aan een oppervlak. De structuur van het proefschrift is als volgt:

Hoofdstuk 1 introduceert kort het concept van *in silico* materiaalontwerp.

Hoofdstuk 2 legt dichtheidsfunctionaaltheorie uit, evenals enkele praktische computationele technieken voor het berekenen van de elektronische structuur binnen dit kader.

Hoofdstuk 3 schetst de workflows die worden gebruikt voor de automatisering van de vereiste *ab initio* berekeningen van elke descriptor of metriek.

Hoofdstuk 4 bespreekt de Shockley-Queisser limiet en de *spectroscopic limited maximum efficiency*, twee metrieken die worden gebruikt om het potentieel van een materiaal als de absorberende laag van een zonnecel te bepalen. Vervolgens maakt het een vergelijking van de *CuAu-like* en chalcopyrietfase in de context van dunne-film fotovoltaïsche cellen.

Hoofdstuk 5 presenteert een onderzoek naar de stabiliteit van het zuurstofkader van Li-rijke Li_2MnO_3- en Li_2IrO_3-batterijkathoden, evenals een beperkte substitutie van Mn als een potentieel recept voor het verbeteren van de structurele stabiliteit van deze materialen. Bovendien bespreekt het de energielandschappen van $LiCB_{11}H_{12}$ en $NaCB_{11}H_{12}$ polyboraanzouten in de context van vastestofelektrolyten.

Hoofdstuk 6 beschrijft een nieuw model voor het berekenen van de secundaire elektronenemissie van ionen geneutraliseerd op een halfgeleider- en metaaloppervlak, en past deze toe op een reeks elementaire oppervlakken in een *high-throughput* benadering.

Dedicated to my brother.

Acknowledgments

First and foremost, I would like to express my sincere gratitude to my supervisors Prof. Dr. Dirk Lamoen and Prof. Dr. Bart Partoens for providing continuous support, guidance and insightful feedback and discussion during the past four years. Second, I want to thank the many colleagues and collaborators I have had the pleasure to work with during my research, their valuable perspectives have been of great benefit to my work. A special mention here should go to Dr. Brandon Wood, for giving me the opportunity to work at Lawrence Livermore National Laboratory during the computational chemistry and materials science summer institute. My appreciation also extends to the academic and technical staff of the University of Antwerp and the Flemish Supercomputer service. Without their infrastructure and hard work, my research would not have been possible.

On a more personal note, I would like to thank my friends and family for their encouragement and contributions. Specifically, my thanks goes to Nils Deputter, who designed the cover of my thesis as well as some of the images from the chapter on solar cells. I am also very grateful to Elise Stynen, for always supporting me during my PhD and putting up with the long hours at work and times when stress or fatigue made my demeanor less than pleasant. Thank you for your understanding and love. Last but not least, I would like to thank my mother Tilda Stockmans. Her incredible strength and perseverance in face of the many hardships she has had to endure has always inspired me. It would not be an exaggeration to say that a large part of my motivation to succeed comes from a desire to make her proud. To make sure she knows that, no matter what has happened in the past, she has always been a wonderful mother, and the person I respect most in this world.

Contents

List of Acronyms

AD Auger De-excitation
AE All-electron
AIMD *Ab initio* molecular dynamics
AM Air mass
AN Auger neutralization
BZ Brillouin zone
BZ Irreducible Brillouin zone
CA CuAu-like
CH Chalcopyrite
DFT Density functional theory
DOS Density of states
EDX Energy-dispersive X-ray
GGA Generalised gradient approximation
HAADF High-angle annular dark-field
HF Hartree-Fock
HSE Heyd-Scuseria-Ernzerhof
KS Kohn-Sham
LDA Local density approximation
NEB Nudged elastic band
NMC Nickel-Manganese-Cobalt
PAW Projector augmented wave
PBE Perdew-Burke-Ernzerhof
PDOS Projected density of states
PP Pseudopotential
PS Pseudo
PV Photovoltaic
PXRD Powder X-ray diffraction
QUOTAS quantification of target surfaces
RN Resonance Neutralization
SCAN Strongly constrained and appropriately normed
SCF Self-consistent field
SEE Secondary electron emission
SLME Spectroscopic limited maximum efficiency
SQ Shockley-Queisser
SRH Shockley-Read-Hall
TM transition metal
VASP Vienna ab initio simulation package

Chapter 1

In Silico Materials Design

> *"If you can look into the seeds of time,*
> *And say which grain will grow and which will not,*
> *Speak then to me, who neither beg nor fear,*
> *Your favors nor your hate."*
>
> Banquo - *Macbeth Act I, Scene iii*

The discovery of new materials has driven many of the greatest technological revolutions in human history. In ancient times, the advent of the use of bronze and iron was an advancement that now marks an age. The invention of modern concrete changed the rules of civil engineering. More recently, the discovery of bakelite, the first truly synthetic material, led to the onset of the modern plastics industry. Clearly, material innovation is an essential component of technological development.

Currently, we know the properties of less than 1% of all materials, and it can require decades to identify new compounds for a technological application [1]. This process needs to be more efficient, by reducing both the time and cost of material development. Over the past few decades, there has been a significant increase in the capability of high-performance computing, which shows no signs of abating. Computer simulations that use sophisticated quantum models have the power to analyze many properties of materials, offering a relatively cheap and effective method to discover new potential candidates for any application. Recently developed computational high-throughput methods provide an automated procedure to screen large amounts of compounds [2, 3]. The combination of these rapidly evolving techniques with a suitable descriptor or selection metric has the potential of identifying the most promising materials from the many different possible structures.

In this chapter, I aim to provide a brief introduction to the concept of *in silico* materials design, as well as frame my research within this broader context. Section 1.1 starts by explaining the term *in silico*. Next, I combine this with the second part of the title of this chapter and connect it to the chapters of my thesis in Section 1.2. Finally, Section 1.3 discusses some conventions used throughout the thesis to help the reader navigate the document.

1.1 In Silico

Although the theoretical modeling of materials has a long history, the practice of describing physical relations using equations largely started during the scientific revolution in the 16th and 17th century. With the advent of modern computers, theoretical modeling was taken to the next level by facilitating the solution of these equations for more complex systems. This allowed the introduction of "computer experiments", i.e. simulations, whose results can be analyzed and compared with those from experiment. *In silico* refers to the use of such computer simulations, usually on a large computational infrastructure called a supercomputer, to solve intricate questions in a range of scientific disciplines.

For many problems in solid state physics, the equation that we aim to solve *in silico* is the Schrödinger equation. More specifically, we want to determine the electronic structure of a material, as many important properties can be derived from an understanding of the electronic states and their response to e.g. electromagnetic fields. Our method of choice for calculating the electronic structure *ab initio*, i.e. from first principles, is density functional theory (DFT), which is treated in Chapter 2. Although there are other methods available, DFT strikes a good balance between accuracy and computational workload, allowing for the study of a large number of systems.

1.2 Materials Design

Traditionally, many materials have been discovered experimentally either by chance or trial and error, largely driven by intuition. Such a trial and error approach can either be performed one material at a time, or in a so-called high-throughput fashion. A popular experimental example of the latter is how Edison tested over 6000 materials to serve as a filament for the incandescent light bulb [4]. Once a material is found that has the right properties, it is investigated in order to understand its structure and its relation to the origin of these properties. This approach, often referred to as direct design, has been the predominant source of new materials over the past centuries.

However, a potentially more effective approach to discover new materials is inverse design [5], where instead of finding materials and studying their properties, we start with a certain application in mind and then look for materials that exhibit the right properties for said application. There are different strategies to execute this approach, such as: high-throughput computational efforts that generate large databases [2, 6], that rely on evolutionary techniques to generate new and improved structures [7] or to use machine learning and data mining to discover trends and patterns [8]. The combination of these ideas with the development of ever more accurate computational techniques has put us on the brink of a revolution in material science, where large scale automated calculations can guide us to the next generation of materials for commercial applications.

When using DFT as our method of choice for calculating the properties of the materials under investigation, the relation between the fundamental quantities we can obtain from DFT and the property of interest is often complex. As phrased by Gerbrand Ceder, "there is no quantum operator for a better car" [9]. In order to obtain insight from the DFT results, we need domain-specific expertise to develop descriptors or metrics for the application of interest. During my PhD I have worked on a diverse set of topics, but the main goal has always been to use results from *ab initio* calculations to study materials with a certain application in mind. Moreover, I have developed workflows – discussed in Chapter 3 – for automating the DFT calculations and open-source Python packages for performing the subsequent analysis, i.e. transferring the DFT results to the task of materials design.

Here is a brief overview of how the various topics I have worked on can be framed in the context of *in silico* materials design:

> **Chapter 4** concerns the application and analysis of the spectroscopic limited maximum efficiency (SLME), a metric developed to judge the potential of materials for solar cells. Here, the combination of the workflow described in Section 3.3.1 with the post-processing tools for calculating the SLME allows for an automated high-throughput screening of materials for photovoltaic applications.
>
> **Chapter 5** discusses an investigation of the Li-rich layered oxides, a class of cathode materials that have demonstrated large energy densities, as well as the energy landscapes of $LiCB_{11}H_{12}$ and $NaCB_{11}H_{12}$ polyborane salts in the context of solid electrolytes. The computation of the properties of interest here is more involved, and hence the workflows described in Sections 3.3.2, 3.3.3 and 3.3.5 are necessary to make this process feasible for even a limited number of materials[1].
>
> **Chapter 6** details the calculation of the secondary electron emission from slow ions neutralized at a surface. Here we have both developed a model for calculating this descriptor from the DFT results, as well as a workflow (Sec. 3.3.4) for automating its calculation.

1.3 Guide to this thesis

One of the main objectives of this thesis is to provide the reader with the necessary information and tools to be able to reproduce my work, as well as extend their application to other materials. To achieve this, the workflows I designed are documented in Chapter 3, along with a description of the parts that they consist of. All of the underlying code is also freely available on github, and the necessary steps to set up the calculations and process the data are detailed in Jupyter notebooks. I have also added a more conventional description of the computational settings to Appendix A.2, in order to gather all of these details in one place and not unnecessarily interrupt the discussion of the results on the various topics.

To make this document easier to navigate, I have added a large amount of links to the text. These are consistently colored as follows: dark red text links to other parts of the thesis, whereas blue text links to a web page, which will be automatically opened in a tab of your default browser. Besides these links, each section that contains results also has two symbols next to its header:

jupyter

Links to the corresponding Jupyter notebooks that have been used to set up the calculations and process the data for the results presented in the section. Several results can also be explored interactively, depending on the topic. For most sections, the figures also have been set up in a Jupyter notebook.

Links to a section in Appendix A.2 which contains a more conventional description of the computational parameters used for the calculation of the properties presented in the section.

I highly recommend the reader to look up the key of their PDF reader of choice for returning to the previous page, i.e. before clicking on one of the links that navigates to a different section of the text. For Apple's preview, this is `Cmd+[`. For Acrobat reader, this should be `Alt+Leftarrow`.

---

[1] One of the Master's students I helped to supervise calculated that it would have taken him 1200 hours to manually perform and process the calculations for his thesis, which largely focused on a single material.

Bibliography

[1] A. Jain et al. “Commentary: The Materials Project: A materials genome approach to accelerating materials innovation”. *APL Materials* 1 (2013), p. 011002. DOI: `10.1063/1.4812323`.

[2] S. Curtarolo et al. “The high-throughput highway to computational materials design”. *Nat. Mater.* 12 (2013), pp. 191–201. DOI: `10.1038/nmat3568`.

[3] A. Jain et al. “Fireworks: A dynamic workflow system designed for high-throughput applications”. *Concurr. Comput.* 27 (2015), pp. 5037–5059. DOI: `10.1002/cpe.3505`.

[4] E. C. “Thomas A Edison: The Life-story of a Great American”. *Nature* 100 (1918), pp. 504–504. DOI: `10.1038/100504a0`.

[5] A. Zunger. “Inverse design in search of materials with target functionalities”. *Nat. Rev. Chem.* 2 (2018). DOI: `10.1038/s41570-018-0121`.

[6] A. Jain et al. “A high-throughput infrastructure for density functional theory calculations”. *Comput. Mater. Sci.* 50 (2011), pp. 2295–2310. DOI: `10.1016/j.commatsci.2011.02.023`.

[7] A. R. Oganov et al. “Computational Materials Discovery Using Evolutionary Algorithms”. In: *Computational Materials Discovery.* Royal Society of Chemistry, 2018, pp. 15–65. DOI: `10.1039/9781788010122-00015`.

[8] C. Draxl and M. Scheffler. “Big Data-Driven Materials Science and Its FAIR Data Infrastructure”. In: *Handbook of Materials Modeling.* Springer International Publishing, 2019, pp. 1–25. DOI: `10.1007/978-3-319-42913-7_104-1`.

[9] G. Ceder. “Opportunities and challenges for first-principles materials design and applications to Li battery materials”. *MRS Bull.* 35 (2010), pp. 693–701. DOI: `10.1557/mrs2010.681`.

Chapter 2

Density Functional Theory

> *"The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble. It therefore becomes desirable that approximate practical methods of applying quantum mechanics should be developed, which can lead to an explanation of the main features of complex atomic systems without too much computation."*
>
> Paul Dirac

> *"When teaching chemistry students, I explain that DFT is some algorithm meaning unreliable, while ab initio is Latin for too expensive."*
>
> Kieron Burke

Density functional theory (DFT) is a popular quantum mechanical method for calculating the electronic structure of a system, based on the idea that all of the relevant information on the electrons is stored in the electron density function. Within this theory, the properties of the system are defined as functionals, i.e. "a function of a function", of the electron density. It is one of the most widely used *ab initio* methods available, mainly because of its wide range of possible applications as well as its relative computational simplicity.

In this chapter, I discuss the theoretical underpinnings of DFT. Section 2.1 describes the many-body problem and explains the theoretical framework that allows us to tackle electronic structure calculations using DFT. Section 2.2 continues by explaining the important concepts of practical computations, such as the basis set for the expansion of the one-electron orbitals and the pseudopotential method. Section 2.3 discusses linear response theory and the calculation of the dielectric tensor. Finally, the chapter finishes with a brief summary of transition state theory in Section 2.4, as well as the nudged elastic band method, which is a valuable technique for determining the kinetic barrier of transitions.

2.1 Theoretical Background

We start our discussion by considering the quantum many-body problem. A crystal can be described by the wave function of the interacting particles, consisting of M nuclei and N electrons [1]:

$$\Psi(\mathbf{r}_1, \sigma_1, \ldots, \mathbf{r}_N, \sigma_N, \mathbf{R}_1, \Sigma_1, \ldots, \mathbf{R}_M, \Sigma_M), \tag{2.1}$$

where the many-body wave function Ψ depends on the position and spin coordinates of the electrons $(\mathbf{r}_i, \sigma_i)$ with mass m_e and the nuclei $(\mathbf{R}_k, \Sigma_k)$ with mass M_k. In the interest of making the equations manageable, it is conventional to write the coordinates as:

$$\begin{aligned}(\mathbf{r}_1, \sigma_1, \ldots, \mathbf{r}_N, \sigma_N) &= (\mathbf{x}_1, \ldots, \mathbf{x}_N) = \mathbf{x}\\ (\mathbf{R}_1, \Sigma_1, \ldots, \mathbf{R}_M, \Sigma_M) &= (\mathbf{X}_1, \ldots, \mathbf{X}_M) = \mathbf{X}.\end{aligned}$$

Using this notation, the time-independent Schrödinger equation becomes

$$\hat{H}\Psi(\mathbf{X}, \mathbf{x}) = E\Psi(\mathbf{X}, \mathbf{x}). \tag{2.2}$$

In the non-relativistic case, the Hamiltonian of the many-body problem is given by

$$\begin{aligned}\hat{H} = &-\sum_{k=1}^{M} \frac{\hbar^2}{2M_k}\Delta_{\mathbf{R}_k} - \sum_{j=1}^{N} \frac{\hbar^2}{2m_e}\Delta_{\mathbf{r}_j} + \frac{1}{2}\sum_{k_1 \neq k_2 = 1}^{M} \frac{1}{4\pi\varepsilon_0}\frac{Z_{k_1}Z_{k_2}e^2}{|\mathbf{R}_{k_1} - \mathbf{R}_{k_2}|}\\ &+\frac{1}{2}\sum_{j_1 \neq j_2 = 1}^{N} \frac{1}{4\pi\varepsilon_0}\frac{e^2}{|\mathbf{r}_{j_1} - \mathbf{r}_{j_2}|} - \sum_{k=1}^{M}\sum_{j=1}^{N} \frac{1}{4\pi\varepsilon_0}\frac{Z_k e^2}{|\mathbf{R}_k - \mathbf{r}_j|},\end{aligned} \tag{2.3}$$

where Z_k is the number of protons in the nuclei, ε_0 is the vacuum permittivity and $\hbar$ is the reduced Planck constant. The first two terms correspond to the kinetic energy of the nuclei and the electrons. The other three terms describe the potential energy for the nucleus-nucleus, electron-electron and the nucleus-electron interaction. In order to further shorten future notation, we write the Hamiltonian terms as

$$\hat{H} = \hat{H}_{k,n} + \hat{H}_{k,e} + \hat{H}_{p,n-n} + \hat{H}_{p,e-e} + \hat{H}_{p,n-e}, \tag{2.4}$$

where $\hat{H}_{k,n}$, $\hat{H}_{k,e}$ are the kinetic energy terms and $\hat{H}_{p,n-n}$, $\hat{H}_{p,e-e}$, $\hat{H}_{p,n-e}$ are the potential energy terms for the corresponding interactions.

2.1.1 The Born-Oppenheimer Approximation

Our first step in tackling the many-body problem is the Born-Oppenheimer approximation [2]. Since the mass of a nucleus is much larger than that of an electron, it is reasonable to assume that the nuclei move much more slowly than the electrons. As such, we can consider their position to be fixed, with the electrons reacting instantly to any change in the nuclear positions. Applying these assumptions leads to the following approximate expression for the wave function:

$$\Psi(\mathbf{X}, \mathbf{x}) = \Psi_n(\mathbf{X})\Psi_e(\mathbf{X}, \mathbf{x}). \tag{2.5}$$

The nuclear wave function $\Psi_n(\mathbf{X})$ depends solely on the nuclear coordinates, whereas the electronic wave function $\Psi_e(\mathbf{X}, \mathbf{x})$ depends directly on the electronic coordinates and considers the nuclear coordinates as parameters. If we substitute this product of wave functions into the Schrödinger equation and divide by the total wave function, we find

$$\frac{(\hat{H}_{k,e} + \hat{H}_{p,e-e} + \hat{H}_{p,n-e})\Psi_e(\mathbf{X},\mathbf{x})}{\Psi_e(\mathbf{X},\mathbf{x})} = E - \frac{(\hat{H}_{k,n} + \hat{H}_{p,n-n})\Psi_n(\mathbf{X})}{\Psi_n(\mathbf{X})}. \tag{2.6}$$

The right hand side of this equation is a function that only depends on the nuclear coordinates $\mathbf{X}$. If we write this function as $E_e(\mathbf{X})$ and multiply both sides with the electronic wave function, we find the Schrödinger equation for the electrons:

$$(\hat{H}_{k,e} + \hat{H}_{p,e-e} + \hat{H}_{p,n-e})\Psi_e(\mathbf{X},\mathbf{x}) = E_e(\mathbf{X})\Psi_e(\mathbf{X},\mathbf{x}). \tag{2.7}$$

We can see that the electronic wave function only depends parametrically on the nuclei coordinates, through the electrostatic interaction in $\hat{H}_{p,n-e}$. Changing to atomic units

$$\hbar = 1 \qquad m_e = 1 \qquad e = 1 \qquad 4\pi\varepsilon_0 = 1, \tag{2.8}$$

we finally obtain the electronic Schrödinger equation:

$$\left[-\frac{1}{2}\sum_{i=1}^{N}\Delta_{\mathbf{r}_i} + \frac{1}{2}\sum_{i_1\neq i_2=1}^{N}\frac{1}{|\mathbf{r}_{i_1}-\mathbf{r}_{i_2}|} - \sum_{k=1}^{M}\sum_{i=1}^{N}\frac{Z_k}{|\mathbf{R}_k-\mathbf{r}_i|}\right]\Psi_e(\mathbf{X},\mathbf{x}) = E_e(\mathbf{X})\Psi_e(\mathbf{X},\mathbf{x}). \tag{2.9}$$

The Born-Oppenheimer approximation significantly reduces the complexity of the many-body problem. Every set of fixed nuclei coordinates defines an external potential, for which we can solve the Schrödinger equation to obtain the electronic wave function. However, since the nuclei coordinates are considered to be fixed parameters, solving Eq. (2.9) does not provide any information on the lattice constants or atomic positions. We return to this issue in Section. 2.1.5.

2.1.2 Hohenberg-Kohn Theorems

Although the idea of using the electron density as the main variable when solving the many-body problem dates from the 1920's, it was not properly formalized until 1964. This is when Hohenberg and Kohn introduced two theorems that formulate density functional theory as an exact theory of many-body systems [3]. We begin our discussion of these theorems by defining the electron density function as

$$\rho(\mathbf{r}) = \langle\Psi_e|\,\hat{\rho}(\mathbf{r})\,|\Psi_e\rangle\,, \tag{2.10}$$

where $\hat{\rho}(\mathbf{r})$ is the electron density operator:

$$\hat{\rho}(\mathbf{r}) = \sum_{i=1}^{N}\delta(\mathbf{r}_i - \mathbf{r}). \tag{2.11}$$

and the electron density is normalized to the number of electrons:

$$\int \rho(\mathbf{r})d\mathbf{r} = N \tag{2.12}$$

The first of the Hohenberg-Kohn theorems states that there is a one to one correspondence between the ground state density and the external potential of our system. When solving the electronic Schrödinger equation for a certain external potential, we find the unique ground state wave function, which can in turn be used to calculate the ground state density. Because of the first Hohenberg-Kohn theorem, we know that the reverse is possible as well. In other words, we cannot have two different external potentials for a given electron density $\rho(\mathbf{r})$. This means that the electron density contains all the relevant information of the system. As a result, we can write any observable as a unique functional of the electron density [4]:

$$O[\rho] = \langle\Psi|\, \hat{O}\, |\Psi\rangle\,. \tag{2.13}$$

The second Hohenberg-Kohn theorem concerns the functional of the electronic energy $E_e[\rho]$. It states that the form of this functional is given by

$$E_e[\rho] = F_{HK}[\rho] + \int \rho(\mathbf{r}) V_{ext}(\mathbf{r}) d\mathbf{r}, \tag{2.14}$$

where $F_{HK}[\rho]$ is called the Hohenberg-Kohn energy functional and $V_{ext}(\mathbf{r})$ is the external potential generated by the nuclei:

$$V_{ext}(\mathbf{r}) = -\sum_{k=1}^{M} \frac{Z_k}{|\mathbf{R}_k - \mathbf{r}|}. \tag{2.15}$$

Furthermore, once we know the functional of the electronic energy, we can approximate the ground state energy by using a variational principle:

$$E_0 = E_e[\rho_0] \leq E_e[\rho], \tag{2.16}$$

with ρ_0 the ground state density of the electrons. Hence, if we know the functional $F_{HK}[\rho]$, we can provide an upper bound[1] for the ground state energy by inserting different functions for $\rho(\mathbf{r})$ and looking for the minimal value of Eq. (2.14).

An important concept in this discussion is the *universality* of the functional $F_{HK}[\rho]$. The second Hohenberg-Kohn theorem does not simply propose the existence of an energy functional $F_{HK}[\rho]$ for each external potential $V_{ext}(\mathbf{r})$, it declares that there is one such functional which is the same for every system. Unfortunately, we do not know the exact form of this functional, which means that we have to look for approximations. We know that $F_{HK}[\rho]$ consists of the kinetic and the interaction energy of the electrons:

$$F_{HK}[\rho] = T[\rho] + \frac{1}{2}\iint \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r} d\mathbf{r}' + E'_{xc}[\rho], \tag{2.17}$$

where the first term denotes the kinetic energy, the second term is due to the Coulomb interaction, and all other interaction energy is combined in the *exchange-correlation* energy $E'_{xc}[\rho]$. Finding a good expression for $E'_{xc}[\rho]$ is of great importance for DFT calculations, and is discussed in more detail in Section 2.1.4.

---

[1] Since we use a variational principle, the electron energy found by minimization is in principle an upper bound for the ground state electron energy. However, when approximations are used for the functional $F_{HK}[\rho]$, the minimized electron energy can no longer be considered an upper bound. This means that it is possible that DFT calculations produce a ground state energy below E_0.

The Hohenberg-Kohn theorems, as presented in this section, apply to the ground state of a system of spin-unpolarised electrons subjected to an external potential. However, they can be extended to include the presence of several types of particles. In order to study a spin-polarised system, we separate the electron density in a component for the spin-up orbitals and the spin-down orbitals [5]:

$$\rho(\mathbf{r}) = \rho(\mathbf{r}, \sigma = \uparrow) + \rho(\mathbf{r}, \sigma = \downarrow). \tag{2.18}$$

This results in a new functional for the electron energy, which now also depends on the spin density $s(\mathbf{r}) = \rho(\mathbf{r}, \sigma = \uparrow) - \rho(\mathbf{r}, \sigma = \downarrow)$:

$$E_e = E_e[\rho(\mathbf{r}), s(\mathbf{r})] = E_e^s[\rho^s(\mathbf{x})], \tag{2.19}$$

where in the final expression $\rho^s(\mathbf{x})$ is a function that depends on both the electrons' position and spin coordinates.

2.1.3 Kohn-Sham Equations

The Hohenberg-Kohn theorems turn the electron density into the fundamental variable of the system. The first Hohenberg-Kohn theorem guarantees that no information is lost when we consider the electron density as the main variable instead of the wave function. The second Hohenberg-Kohn theorem allows us to use a variational principle to find the ground state electron density.

The next step in solving the many-body problem using DFT was introduced by Kohn and Sham in 1965 [6]. They considered a system of non-interacting electrons that has exactly the same energy and electron density as the original problem. This fictitious system is subject to a different external potential[2] $V_{eff}(\mathbf{r})$, but has a much simpler expression for the electron energy functional:

$$E_e'[\rho] = T_0[\rho] + \int V_{eff}(\mathbf{r})\rho(\mathbf{r})d\mathbf{r}, \tag{2.20}$$

where $T_0[\rho]$ is the kinetic energy of the new system, which differs from $T[\rho]$. Since the particles are now non-interacting, the Hamiltonian has the form

$$\hat{H}_{KS} = \sum_{i=1}^{N} \left[-\frac{1}{2}\Delta_{\mathbf{r_i}} + V_{eff}(\mathbf{r}) \right]. \tag{2.21}$$

For this Hamiltonian, the solution of the Schrödinger equation can be written as a Slater determinant: $\Psi_e'(\mathbf{r}) = |\phi_1(\mathbf{r}), \phi_2(\mathbf{r}), \ldots, \phi_N(\mathbf{r})|$, where the single-particle orbitals $\phi_i(\mathbf{r})$ are solutions of the single-particle Schrödinger equations:

$$\left[-\frac{1}{2}\Delta_{\mathbf{r}_i} + V_{eff}(\mathbf{r}) \right] \phi_i(\mathbf{r}) = \epsilon_i \phi_i(\mathbf{r}). \tag{2.22}$$

---

[2] Note that for spin-polarised systems, the effective potential can also depend on the spin coordinates.

The electron density is now given by [7]

$$\rho(\mathbf{r}) = \sum_{i=1}^{N} |\phi_i(\mathbf{r})|^2 . \tag{2.23}$$

This non-interacting system is easier to solve, but we do not know $V_{eff}(\mathbf{r})$. However, we do know that by definition both systems have the same energy and electron density, which can be used to find an expression for $V_{eff}(\mathbf{r})$. Writing down the original many-body electron energy functional using Eq. (2.17):

$$\begin{aligned} E_e[\rho] &= F_{HK}[\rho] + \int \rho(\mathbf{r}) V_{ext}(\mathbf{r}) d\mathbf{r} \\ &= T[\rho] + \frac{1}{2} \iint \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r} d\mathbf{r}' + \int \rho(\mathbf{r}) V_{ext}(\mathbf{r}) d\mathbf{r} + E'_{xc}[\rho], \end{aligned}$$

we can equate it to the non-interacting total energy from Eq. (2.20) and rewrite it to find

$$\int \rho(\mathbf{r}) V_{eff}(\mathbf{r}) d\mathbf{r} = \int \rho(\mathbf{r}) V_{ext}(\mathbf{r}) d\mathbf{r} + \frac{1}{2} \iint \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r} d\mathbf{r}' + E_{xc}[\rho], \tag{2.24}$$

where we have redefined the exchange correlation energy by including the difference between the kinetic energy of the original and non-interacting system:

$$E_{xc}[\rho] = (T[\rho] - T_0[\rho]) + E'_{xc}[\rho]. \tag{2.25}$$

Finally, if we take the functional derivative of Eq. (2.24), we get an expression for the effective potential:

$$\begin{aligned} V_{eff}(\mathbf{r}) &= V_{ext}(\mathbf{r}) + \int \frac{\rho(\mathbf{r})}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}' + \frac{\delta E_{xc}}{\delta \rho} && (2.26) \\ &= V_{ext}(\mathbf{r}) + V_H(\mathbf{r}) + V_{xc}(\mathbf{r}), && (2.27) \end{aligned}$$

where $V_H(\mathbf{r})$ is the Hartree or Coulomb potential and $V_{xc}(\mathbf{r})$ is the exchange-correlation potential.

We call Eqs. (2.22), in combination with the expression for $V_{eff}(\mathbf{r})$ from Eq. (2.27), the *Kohn-Sham* (KS) equations. At first glance, we seem to be faced with a problem: In order to calculate the ground state electron density $\rho(\mathbf{r})$, we need to solve the KS equations to find the one electron wave functions $\phi_i(\mathbf{r})$. However, these equations require the knowledge of the effective potential $V_{eff}(\mathbf{r})$, whose calculation in turn requires the ground state density $\rho(\mathbf{r})$. To solve this issue, we perform what is called a self-consistent field (SCF) procedure. Starting from an initial guess for the electron density, we calculate the effective potential. Once we have the effective potential, we insert it into the Kohn-Sham equations to calculate the one electron wave functions. We then use these wave functions to calculate the new ground state density, from which we can once again derive a new effective potential. This provides us with an algorithm which converges to the ground state density after a sufficient amount of iterations. This algorithm is shown schematically in Fig. 2.1.

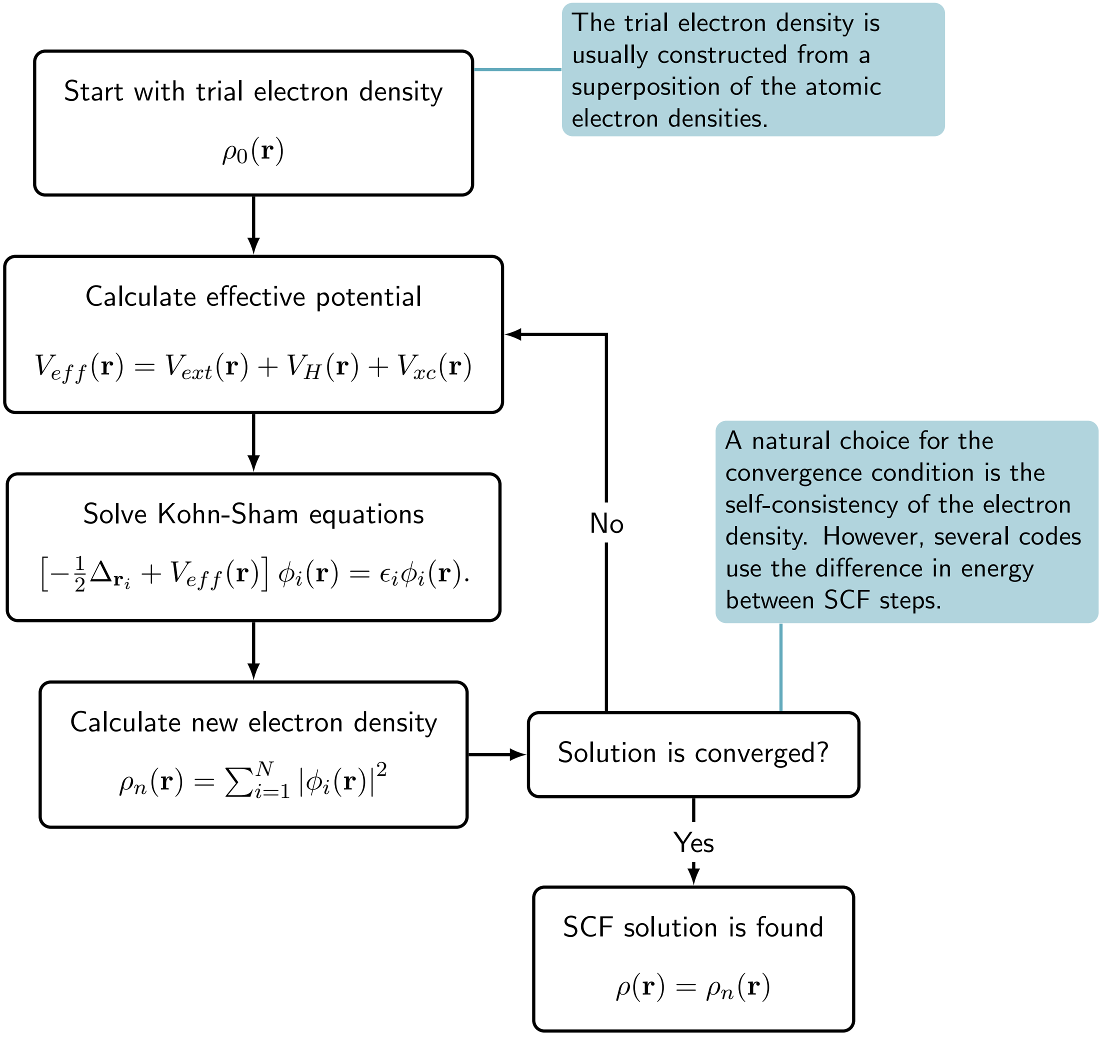


Figure 2.1: Flowchart representing the self consistent field calculation used to find the ground state density in DFT calculations.

The SCF procedure is converged when the electron density is *self-consistent*, i.e. solving the Kohn Sham equations using the effective potential for the converged electron density results in the same density. However, several implementations of the DFT formalism consider the difference between the total electronic energy of two SCF iteration steps for the convergence criterion.

2.1.4 Exchange Correlation Functional

The Kohn-Sham method is in principle exact, up to a certain precision determined by a parameter in the convergence procedure. However, we are still faced with one of the fundamental challenges of DFT: the expression of the exchange-correlation functional $E_{xc}[\rho]$ is unknown. Since the calculation of the effective potential for each iteration step requires this functional, we are forced to make use of approximations. One popular way of classifying the various approximations for the exchange-correlation energy is as rungs on Jacob's ladder for DFT [8], ultimately leading to chemical accuracy heaven (Fig. 2.2). These rungs are also often referred to as the *level* of the theory.

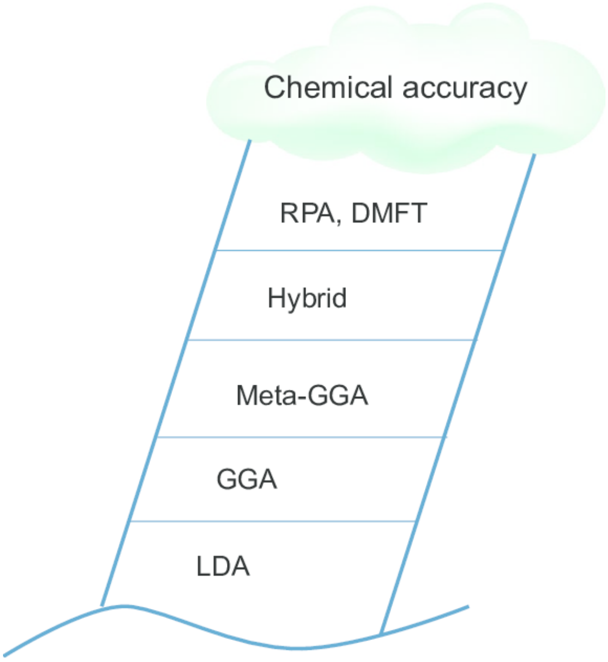


Figure 2.2: Jacob's ladder of density functional approximations for the exchange-correlation energy. Taken from [9].

The first rung on Jacob's ladder is the local density approximation (LDA), introduced in the original paper of Kohn and Sham [6]:

$$E_{xc}^{LDA}[\rho] = \int \rho(\mathbf{r})\epsilon_{xc}(\rho(\mathbf{r}))d\mathbf{r}, \tag{2.28}$$

where $\epsilon_{xc}(\rho(\mathbf{r}))$ is the exchange-correlation energy function, which in the LDA is assumed to depend only on the electron density $\rho(\mathbf{r})$. Most successful variants of the LDA calculate the exchange-correlation for the homogeneous electron gas, based on an exact expression for the exchange part and quantum Monte Carlo simulations for the correlation part.

In the next level of theory, some non-locality is introduced through the gradient of the electron density. In this way, we obtain the generalized gradient approximation (GGA) [10]:

$$E_{xc}^{GGA}[\rho] = \int \rho(\mathbf{r})f(\rho(\mathbf{r}), |\nabla\rho|)d\mathbf{r}, \tag{2.29}$$

where $f(\rho(\mathbf{r}), |\nabla\rho|)$ is a function of the electron density $\rho(\mathbf{r})$ and its gradient $|\nabla\rho|$. GGA's generally produce better results than the LDA, but there is some freedom in determining how exactly the gradient is included in the approximation through the definition of $f(\rho(\mathbf{r}), |\nabla\rho|)$. Some variants of the GGA use empirical data to fit the functional parameters to a wide range of experimental results [11, 12]. Alternatively, the functional form and parameters can be constructed based on theoretical constraints. An example of such a non-empirical GGA functional that is very popular in solid state physics is the *Perdew-Burke-Ernzerhof* (PBE) [13] functional:

$$E_{xc}^{PBE}[\rho] = \int \rho(\mathbf{r})\left[\epsilon_{xc}(\rho(\mathbf{r})) + H[\rho, t]\right]d\mathbf{r}, \tag{2.30}$$

with

$$H[\rho,t] = \left(\frac{e^2}{a_0}\right)\gamma\phi^3 \ln\left\{1+\frac{\beta}{\gamma}t^2\left[\frac{1+At^2}{1+At^2+A^2t^4}\right]\right\}, \tag{2.31}$$

where $t = |\nabla\rho|/(2k_s\rho)$ is a dimensionless density gradient, with k_s the Thomas-Fermi screening wave number, a_0 the Bohr radius and β, γ constants[3]. The function A has the following form:

$$A = \frac{\beta}{\gamma}\left[e^{-\epsilon_{xc}(\rho(\mathbf{r}))/(\gamma e^2/a_0)} - 1\right]^{-1}. \tag{2.32}$$

The next natural extension of the theory includes a dependency on $\nabla^2\rho(\mathbf{r})$, which leads to the third rung on Jacob's ladder: meta-GGA's. However, this term is also commonly used for functionals that include a dependence on the local kinetic energy density:

$$\tau(\mathbf{r}) = \sum_{i=1}^{N}\frac{1}{2}\left|\nabla\phi_i(\mathbf{r})\right|^2 \tag{2.33}$$

A recent meta-GGA functional based on the local kinetic energy density that has been shown to perform well with regard to the cohesive energies and lattice parameters of solids is the strongly constrained and appropriately normed (SCAN) [15] functional.

Although many LDA, GGA and meta-GGA functionals are quite effective at calculating the lattice parameters, atomic positions and binding energies, a common problem when using these functionals is that they tend to significantly underestimate the band gap of semiconductor materials [16]. Moreover, as they are often based on the homogeneous electron gas, they provide a poor description of the electronic structure of d and f valence orbitals, which are usually more localized [17]. In order to improve the value of the band gap and localization of electrons, it is possible to mix the DFT functional with a fraction of exact *Hartree-Fock* [18] (HF) exchange energy. We call the result of such a combination a *hybrid functional.* A very popular hybrid functional is the *Heyd-Scuseria-Ernzerhof* (HSE) functional [19], which makes a distinction between short and long range interaction, based on a range separation parameter ω:

$$E_{xc}^{HSE} = aE_x^{HF,SR}(\omega) + (1-a)E_x^{PBE,SR}(\omega) + E_x^{PBE,LR}(\omega) + E_c^{PBE}, \tag{2.34}$$

where a is the fraction of HF energy included in the short range interaction. Typical values for the parameters are $a = \frac{1}{4}$ and $\omega = 0.2$ Å, which is known as the HSE06 functional. It is important to note here that the mixing parameter a for hybrid functionals is often obtained by comparing the band gap result of the calculation that uses hybrid functionals to the band gap found from experiments. In this case, calculations using this hybrid functional can also no longer be considered truly *ab initio.*

Because they require the calculation of the exact exchange energy through the HF formalism, hybrid methods are significantly more computationally demanding than strictly DFT-based functionals. Depending on the computational resources available, this can severely restrict the systems of interest that can be feasibly investigated with hybrid functionals. For large systems with strongly localized electrons, another method that can improve the description of the electronic structure is the Hubbard U correction. The basic idea of this modification is to introduce an additional energy term that increases the repulsion between the electrons occupying d or f orbitals of the same atom. Several flavors of U corrections exist in the literature, but for

---

[3] $\beta = 0.066725$ and $\gamma = 0.031091$ [14].

the work in this thesis we have exclusively relied on the one proposed by Dudarev et al. [20]:

$$E_{DFT+U} = E_{DFT} + \frac{U_{eff}}{2}\sum_{\sigma}\left[\sum_{m}\rho^{\sigma}_{mm} - \sum_{m,m'}\rho^{\sigma}_{mm'}\rho^{\sigma}_{m'm}\right], \tag{2.35}$$

where E_{DFT} is the energy obtained from the uncorrected DFT result, and $\rho^{\sigma}_{mm'}$ is the spin orbital density matrix of the orbitals to which we want to apply the U_{eff} correction:

$$\rho^{\sigma}_{mm'} = \sum_{i,\sigma} f^{\sigma}_{i}\, \langle\phi^{\sigma}_{i}|\, \hat{P}^{I}_{mm'}\, |\phi^{\sigma}_{i}\rangle\,, \tag{2.36}$$

with f^{σ}_{i} the occupancy number of the ϕ^{σ}_{i} quasiparticle orbital and $\hat{P}^{I}_{mm'}$ the projector operator of the atom at site I and m, m' the magnetic quantum numbers of the orbitals to which we apply the U correction (e.g. $m = -2, -1, 0, 1, 2$ for d orbitals.). For more details on how to implement the U correction in the projector augmented wave [21, 22] (PAW) method (Sec. 2.2.4), we refer the reader to the paper of Bengone et al. [17]. An intuitive way to describe the effect of Eq. (2.35) is as applying a penalty to the energy for partially occupied states, pushing the system towards either fully occupied or unoccupied orbitals.

2.1.5 Hellman-Feynman theorem

So far we have only discussed the solution of the electronic Schrödinger equation (Eq. (2.9)), which considers the nuclei to be fixed. However, in order to find the lattice parameters or atomic positions, it is necessary to find the coordinates for which the atoms reach an equilibrium, i.e. for which the forces are zero [5]:

$$\mathbf{F}_{I} = -\frac{\partial E}{\partial \mathbf{R_I}} = 0 \qquad\qquad I = 1, \ldots, M, \tag{2.37}$$

where E is the energy of both the electrons and the nuclei:

$$E = E_{e} + \frac{1}{2}\sum_{k\neq l=1}^{M}\frac{Z_{k}Z_{l}}{|\mathbf{R}_{k} - \mathbf{R}_{l}|}. \tag{2.38}$$

Similarly, for the lattice parameters, we want to minimize the stress on the unit cell:

$$\sigma_{ij} = \frac{\partial E}{\partial \epsilon_{ij}}, \tag{2.39}$$

where σ_{ij} is the stress tensor and ϵ_{ij} the strain tensor.

We calculate the partial derivatives in Eqs. (2.37) and (2.39) by applying the *Hellmann-Feynman* theorem [23] [24]. It states that the derivative of the energy with respect to a parameter λ is equal to the expectation value of the derivative of the Hamiltonian operator to this parameter:

$$\frac{\partial E}{\partial \lambda} = \langle\Psi|\, \frac{\partial \hat{H}}{\partial \lambda}\, |\Psi\rangle\,. \tag{2.40}$$

In case of the electron energy E_e, the Hamiltonian is

$$\hat{H}_e = \hat{H}_{k,e} + \hat{H}_{p,e-e} + \hat{H}_{p,n-e}. \tag{2.41}$$

Evaluating the derivative in Eq.(2.37) by applying the Hellmann-Feynman theorem to $\hat{H} = \hat{H}_e + \hat{H}_n$, we find[4]

$$\mathbf{F}_I = -\underbrace{\langle\Psi|\frac{\partial \hat{H}_{k,e}}{\partial \mathbf{R_I}}|\Psi\rangle}_{=0} - \underbrace{\langle\Psi|\frac{\partial \hat{H}_{p,e-e}}{\partial \mathbf{R_I}}|\Psi\rangle}_{=0} - \langle\Psi|\frac{\partial \hat{H}_{p,n-e}}{\partial \mathbf{R_I}}|\Psi\rangle - \frac{\partial E_n}{\partial \mathbf{R_I}}, \tag{2.42}$$

where the first two terms are equal to zero because they do not depend on the position coordinates of the nuclei. Since $\hat{H}_{p,n-e}$ is the potential term describing the interaction between the nuclei and the electrons, the equation for the force becomes

$$\mathbf{F}_I = -\int \rho(\mathbf{r})\frac{\partial V_{ext}}{\partial \mathbf{R}_I} d\mathbf{r} - \frac{\partial E_n}{\partial \mathbf{R}_I}. \tag{2.43}$$

This formula allows us to calculate the forces on the atoms in the system, depending on the atom positions $\mathbf{R}_I$ and the electron density $\rho(\mathbf{r})$. We use it in combination with a suitable minimization algorithm to find the equilibrium positions of the atoms of the crystal.

2.2 Computational Techniques

Section 2.1 presents a short overview of the theoretical concepts that provide the foundation for density functional theory. However, when applying this framework to solving electronic structure problems on a computer, there are many practical considerations which have to be made in order to implement an effective software package. In this section I explain how to solve the electronic structure problems on a computational level.

The section begins by presenting the Bloch theorem for periodic systems and applying it to the Kohn-Sham orbitals. Next up is a discussion of the basis set, which is used for solving the Kohn-Sham equations. I continue by explaining the use of pseudopotentials to limit the size of the basis set, also presenting a more general method called projector augmented waves. Finally, the section concludes by presenting the use of a **k**-point mesh for sampling the first Brioullin zone.

2.2.1 Bloch Theorem

A first piece of information used in the solution of the Kohn-Sham equations is the periodicity of the crystal. A well known result of solid state physics is the Bloch theorem, which states that [25]: "*the eigenfunctions of the wave equation for a periodic potential are the product of a plane wave* $e^{i\mathbf{g}\cdot\mathbf{r}}$ *times a function* $u_{\mathbf{g}}(\mathbf{r})$ *with the periodicity of the lattice*". This means that the solutions of the Kohn-Sham equations can be written as

$$\phi_{\mathbf{g}}(\mathbf{r}) = u_{\mathbf{g}}(\mathbf{r})e^{i\mathbf{g}\cdot\mathbf{r}}, \tag{2.44}$$

---

[4] Because of the assumptions made in the Born-Oppenheimer approximation, the nuclear energy can be treated classically.

where $\mathbf{g}$ is a general reciprocal vector. Any reciprocal vector can be expressed as the sum of a reciprocal lattice vector $\mathbf{G}$ and a vector that lies in the first Brillouin zone $\mathbf{k}$:

$$\mathbf{g} = \mathbf{G} + \mathbf{k}. \tag{2.45}$$

This allows us to rewrite Bloch's theorem as

$$\phi_{\mathbf{g}}(\mathbf{r}) = \left(u_{\mathbf{g}}(\mathbf{r})e^{i\mathbf{G}\cdot\mathbf{r}}\right) e^{i\mathbf{k}\cdot\mathbf{r}}. \tag{2.46}$$

Because $e^{i\mathbf{G}\cdot\mathbf{r}}$ and $u_{\mathbf{g}}(\mathbf{r})$ are periodic over the lattice, the function between brackets in Eq. (2.46) is as well. We rename this function depending on the location of $\mathbf{g}$. If $\mathbf{g}$ is in the n^{th} Brillouin zone, we define it as $u_{n\mathbf{k}}(\mathbf{r})$:

$$\phi_{n\mathbf{k}}(\mathbf{r}) = u_{n\mathbf{k}}(\mathbf{r})e^{i\mathbf{k}\cdot\mathbf{r}}, \tag{2.47}$$

where n is called the *band index* [4]. This is in principle nothing more than an alternative way of labeling the wave vectors, since each reciprocal vector $\mathbf{g}$ corresponds unequivocally to one combination of n and $\mathbf{k}$. The result is that we have split up the Kohn-Sham orbitals into a plane wave $e^{i\mathbf{k}\cdot\mathbf{r}}$ and a periodic function $u_{n\mathbf{k}}(\mathbf{r})$, with $\mathbf{k}$ a wave vector that lies in the first Brillouin zone. This procedure is known as mapping the band in the reduced zone scheme, and is used for constructing the typical band structures used to study the properties of materials [25]. Figure 2.3 shows an example of the reduced zone scheme.

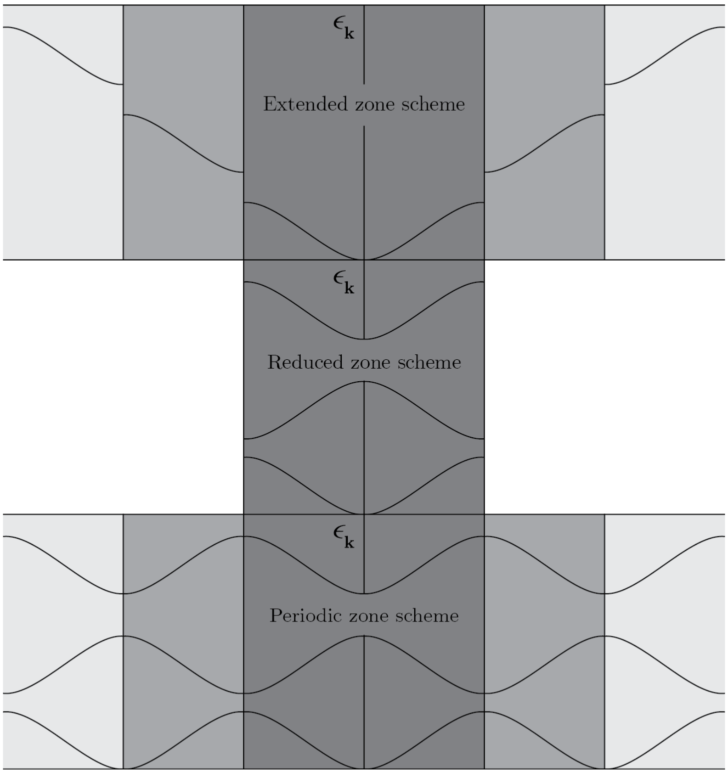


Figure 2.3: The reduced zone scheme, derived from the extended zone scheme by mapping all bands to the first Brillouin zone. If we repeat the reduced zone scheme for the whole reciprocal space, we obtain the so-called periodic zone scheme. Taken from [25].

2.2.2 Basis Set

In order to find the Kohn-Sham orbitals, the one electron wave functions are expanded in a certain basis set. Although in principle any basis set can be used, it is preferable to look for a set of basis functions that strikes a good balance between being *efficient* and *unbiased*. This means that the wave functions can be expanded in a relatively small number of basis functions and that the set is not tailored to suit a specific problem [4]. An example of a set of functions that are certainly unbiased are the plane waves $e^{i\mathbf{G}\cdot\mathbf{r}}$, where $\mathbf{G}$ is a reciprocal lattice vector. If we write $u_{n\mathbf{k}}(\mathbf{r})$ in Eq. (2.47) as a linear combination of these basis functions, we find

$$\begin{aligned} \phi_{n\mathbf{k}}(\mathbf{r}) &= \left(\sum_{\mathbf{G}} c^{\mathbf{G}}_{n\mathbf{k}} e^{i\mathbf{G}\cdot\mathbf{r}}\right) e^{i\mathbf{k}\cdot\mathbf{r}} && (2.48) \\ &= \sum_{\mathbf{G}} c^{\mathbf{G}}_{n\mathbf{k}} e^{i(\mathbf{G}+\mathbf{k})\cdot\mathbf{r}}. && (2.49) \end{aligned}$$

If we want to find the coefficients $c^{\mathbf{G}}_{n\mathbf{k}}$, we have to solve the eigenvalue problem derived from the Kohn-Sham equations. This requires diagonalisation of the Hamiltionian matrix in the basis set. In principle, the basis set can be infinitely large, but for practical calculations we have to put a limit on the number of basis functions. This limit is defined by only including terms with a reciprocal lattice vector $G \leq G_{max}$, which is usually determined by considering the maximum energy:

$$E_{cut} = \frac{\hbar^2 G^2_{max}}{2m_e}. \tag{2.50}$$

The *cutoff energy* E_{cut} is an input parameter for the calculations that determines how many plane waves are included in the basis set. In other words, increasing E_{cut} improves the precision of the calculation, but results in a more difficult diagonalization procedure for the Hamiltonian.

For non-periodic problems, such as the anion-cation system described in Section 5.4.2, it is often more efficient to rely on a localized basis set. A popular example of such a basis set is the Gaussian basis set, often used in molecular calculations. Here, the basis set consists of functions with the following form [26]:

$$\chi(\zeta, n, m, l; r, \theta, \phi) = N Y_{l,m}(\theta, \phi) r^{2n-2-l} e^{-\zeta r^2}, \tag{2.51}$$

where ζ is known as the exponent, n, m and l are the orbital quantum numbers, N is a normalization constant and $Y_{l,m}(\theta, \phi)$ are the spherical harmonic functions.

2.2.3 Pseudopotentials

The efficiency of the plane wave basis set depends on how large E_{cut} must be in order for Eq. (2.49) to correctly describe the wave function $\phi(\mathbf{r})$. This, in turn, is connected with how much $\phi(\mathbf{r})$ varies in the region around $\mathbf{r}$ [14]. When looking at Fig. 2.4, we can see that the radial part of the electron orbitals varies the most near the nucleus of an atom, where the interaction is stronger. However, the region which is important for the chemical bonds lies further from the nucleus, and here the wave function is relatively smooth. In the interest of keeping the calculations manageable, it is desirable to replace the quickly oscillating wave function with a smoother version, which still has the same value as the actual wave function in the region of the chemical bond.

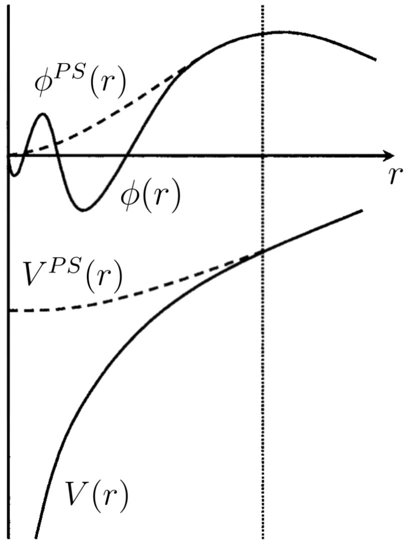


Figure 2.4: Schematic representation of the pseudopotential method. The true wave function (dashed curve) has several nodes, and is replaced by the smoothed version (full curve) after the introduction of the pseudopotential. Adapted from [27].

We can achieve this goal by replacing the potential of the nuclei by a *pseudopotential* (PP) [28]. Firstly, we make a distinction between two types of electrons in the atom: the *core* electrons are highly localized and are considered together with the nucleus, whereas the *valence electrons* are responsible for the bonding and are treated explicitly. In this formalism, we can construct the true valence wave functions $|\phi_v\rangle$ by combining a smooth wave function $|\tilde{\phi}_v\rangle$ with a linear combination of the core electron wave functions $|\phi_c\rangle$ [14]:

$$|\phi_v\rangle = |\tilde{\phi}_v\rangle + \sum_c \alpha_{vc} |\phi_c\rangle \,, \tag{2.52}$$

where $\alpha_{vc} = -\langle\phi_c|\tilde{\phi}_v\rangle$ in order to make the wave function $|\phi_v\rangle$ orthogonal to the core wave functions. In the case of DFT, we need to solve the Kohn-Sham equations (Eq. (2.22)):

$$\underbrace{\left[-\frac{1}{2}\Delta_{\mathbf{r}} + V_{eff}(\mathbf{r})\right]}_{\hat{H}_{KS}} |\phi_v\rangle = \epsilon_v |\phi_v\rangle \,. \tag{2.53}$$

Inserting $|\phi_v\rangle$ from Eq. (2.52) into Eq. (2.53), we find

$$\begin{aligned}
& \hat{H}_{KS} |\phi_v\rangle = \epsilon_v |\phi_v\rangle \\
\Leftrightarrow \quad & \hat{H}_{KS} |\tilde{\phi}_v\rangle + \sum_c \alpha_{vc}\epsilon_c |\phi_c\rangle = \epsilon_v |\tilde{\phi}_v\rangle + \sum_c \epsilon_v\alpha_{vc} |\phi_c\rangle \\
\Leftrightarrow \quad & \hat{H}_{KS} |\tilde{\phi}_v\rangle + \sum_c (\epsilon_c - \epsilon_v)\alpha_{vc} |\phi_c\rangle = \epsilon_v |\tilde{\phi}_v\rangle \\
\Leftrightarrow \quad & \left[\hat{H}_{KS} + \sum_c (\epsilon_v - \epsilon_c) |\phi_c\rangle \langle\phi_c|\right] |\tilde{\phi}_v\rangle = \epsilon_v |\tilde{\phi}_v\rangle \,.
\end{aligned}$$

We define the pseudopotential $\hat{V}_{PP}(\mathbf{r})$ by adding the operator

$$\sum_c (\epsilon_v - \epsilon_c) \left|\phi_c\right\rangle \left\langle\phi_c\right|$$

to the effective Kohn-Sham potential:

$$\hat{V}_{PP}(\mathbf{r}) = V_{eff}(\mathbf{r}) + \sum_c (\epsilon_v - \epsilon_c) \left|\phi_c\right\rangle \left\langle\phi_c\right| . \tag{2.54}$$

The Kohn-Sham equations are now transformed into:

$$\left[-\frac{1}{2}\Delta_{\mathbf{r}} + \hat{V}_{PP}(\mathbf{r})\right] |\tilde{\phi}_v\rangle = \epsilon_v |\tilde{\phi}_v\rangle , \tag{2.55}$$

which is the single particle equation for the smooth part of the wave function. Finding accurate approximations for the pseudopotential operator is a scientific discipline in its own right, and there is a large variety of possible methods of constructing pseudopotentials. For an overview, we refer the reader to [27].

2.2.4 Projector Augmented Waves

The main drawback of the pseudopotential method is that all the information on the full electron wave function near the the nuclei is lost. A more general approach is the projector augmented wave (PAW) method, which was first introduced by Blöchl in 1994 [21]. This method has similarities with the PP approach, in the sense that it tries to replace the true all-electron (AE) wave function with a smoother pseudo (PS) wave function, that agrees with the true wave function in the bonding region. Here I present a brief discussion of the concept, largely based on the excellent paper of Rostgaard [29].

We consider each nucleus to be enclosed by a spherical volume of radius R_a, called the *augmentation* region Ω_{R_a}, and define the remaining part of space as the *interstitial* region. We are looking for a linear operator that transforms the computationally convenient PS one electron wave functions $|\tilde{\phi}\rangle$ into the physically relevant AE one-electron wave functions $|\phi\rangle$:

$$|\phi\rangle = \hat{\mathcal{T}} |\tilde{\phi}\rangle . \tag{2.56}$$

Since we only want to modify the wave functions near the nuclei, we can write the transformation operator in the form

$$\hat{\mathcal{T}} = 1 + \sum_a \hat{\mathcal{T}}_{R_a}, \tag{2.57}$$

where each operator $\hat{\mathcal{T}}_{R_a}$ is a null operator[5] outside the augmentation region of atom a. Inside each augmentation region Ω_{R_a}, we expand the AE wave function in partial waves $|\varphi_i^a\rangle$ and the

---

[5] A null operator is an operator $\hat{A}$ which reduces any state to zero:

$$\hat{A}\,|\alpha\rangle = 0 \qquad \text{for any } |\alpha\rangle .$$

PS wave function in partial waves $|\tilde{\varphi}_i^a\rangle$:

$$|\phi^a\rangle = \sum_i c_i^a |\varphi_i^a\rangle \qquad |\tilde{\phi}^a\rangle = \sum_i \tilde{c}_i^a |\tilde{\varphi}_i^a\rangle , \tag{2.58}$$

and require that the partial waves can be transformed using the same operator as the AE and PS wave function:

$$\forall a, i: \qquad |\varphi_i^a\rangle = \left(1 + \hat{\mathcal{T}}_{R_a}\right) |\tilde{\varphi}_i^a\rangle \quad \Leftrightarrow \quad \hat{\mathcal{T}}_{R_a} |\tilde{\varphi}_i^a\rangle = |\varphi_i^a\rangle - |\tilde{\varphi}_i^a\rangle . \tag{2.59}$$

This completely defines $\hat{\mathcal{T}}$ for a given set of partial waves $\varphi_i^a, \tilde{\varphi}_i^a$. Since we have

$$\sum_i c_i^a |\varphi_i^a\rangle = |\phi^a\rangle = \hat{\mathcal{T}} |\tilde{\phi}^a\rangle = \sum_i \tilde{c}_i^a \hat{\mathcal{T}} |\tilde{\varphi}_i^a\rangle = \sum_i \tilde{c}_i^a |\varphi_i^a\rangle , \tag{2.60}$$

we can see that the coefficients for both expansions are the same ($c_i^a = \tilde{c}_i^a$). Also, because the operators $\hat{\mathcal{T}}_R$ are null operators outside their augmentation regions Ω_R, we can derive from Eq. (2.59) that the partial waves are the same in the interstitial region ($|\varphi_i^a\rangle = |\tilde{\varphi}_i^a\rangle$).

Since we are looking for a linear transformation $\hat{\mathcal{T}}$, the coefficients c_i^a have to be linear functionals of the PS wave function:

$$c_i^a = \langle \tilde{p}_i^a | \tilde{\phi}^a \rangle . \tag{2.61}$$

The $\langle \tilde{p}_i^a|$ are called the *projector functions*. Inside each augmentation region, the one-center expansion of the PS wave function is equal to itself:

$$\sum_i |\tilde{\varphi}_i^a\rangle \langle \tilde{p}_i^a | \tilde{\phi}^a\rangle = |\tilde{\phi}^a\rangle , \quad \text{from which} \quad \sum_i |\tilde{\varphi}_i^a\rangle \langle \tilde{p}_i^a| = 1. \tag{2.62}$$

Using this completeness relation in combination with Eq. (2.59), we find the following expression for the transformation operator in Ω_{R_a}:

$$\hat{\mathcal{T}}_{R_a} = \sum_i \hat{\mathcal{T}}_{R_a} |\tilde{\varphi}_i^a\rangle \langle \tilde{p}_i^a| = \sum_i (|\varphi_i^a\rangle - |\tilde{\varphi}_i^a\rangle) \langle \tilde{p}_i^a| , \tag{2.63}$$

from which we can conclude that the total transformation operator is equal to:

$$\hat{\mathcal{T}} = 1 + \sum_a \sum_i (|\varphi_i^a\rangle - |\tilde{\varphi}_i^a\rangle) \langle \tilde{p}_i^a| . \tag{2.64}$$

Using the PAW method changes the regular Kohn-Sham equations (Eq. (2.22)) for an AE wave function $|\phi\rangle$ with energy ϵ to an eigenvalue equation for the PS wave function:

$$\hat{\mathcal{T}}^\dagger \hat{H}_{KS} \hat{\mathcal{T}} |\tilde{\phi}\rangle = \epsilon \hat{\mathcal{T}}^\dagger \hat{\mathcal{T}} |\tilde{\phi}\rangle . \tag{2.65}$$

After solving Eq. (2.65), we can construct the AE wave function by using Eq. (2.56) and Eq. (2.64):

$$|\phi\rangle = \hat{\mathcal{T}}\,|\tilde{\phi}\rangle = \left(1 + \sum_a \hat{\mathcal{T}}_{R_a}\right)|\tilde{\phi}\rangle = |\tilde{\phi}\rangle + \sum_a \sum_i (|\varphi_i^a\rangle - |\tilde{\varphi}_i^a\rangle)\,\langle \tilde{p}_i^a|\tilde{\phi}\rangle\,, \tag{2.66}$$

which in position space becomes

$$\begin{aligned} \phi(\mathbf{r}) = \langle \mathbf{r}|\phi\rangle &= \langle \mathbf{r}|\tilde{\phi}\rangle + \sum_a \sum_i (\langle \mathbf{r}|\varphi_i^a\rangle - \langle \mathbf{r}|\tilde{\varphi}_i^a\rangle)\,\langle \tilde{p}_i^a|\tilde{\phi}\rangle & (2.67) \\ &= \tilde{\phi}(\mathbf{r}) + \sum_a \sum_i (\varphi_i^a(\mathbf{r}) - \tilde{\varphi}_i^a(\mathbf{r}))\,\langle \tilde{p}_i^a|\tilde{\phi}\rangle\,. & (2.68) \end{aligned}$$

The PAW method is implemented using large data sets for each element. Such a data set governs exactly how the PAW transformation works for each atomic site. One has a large degree of freedom when generating data sets for PAW transformations, which are determined by the partial waves $\varphi_i^a, \tilde{\varphi}_i^a$ and the projector augmented functions $\tilde{p}_i^a(\mathbf{r}) = \langle \mathbf{r}|\tilde{p}_i^a\rangle$. For more information, we refer the reader to [29].

2.2.5 K-point Mesh

For an infinite crystal lattice, **k** is a continuous variable, which means that the calculation of many properties of the system requires an integration over the first Brillouin zone (BZ). When doing numerical calculations, however, this is not practically possible. Instead, we make a selection of **k** points over which we sum the function in order to approximate the integration:

$$\frac{1}{\Omega_{BZ}} \int_{BZ} f(\mathbf{k}) d\mathbf{k} \simeq \frac{1}{N_{\mathbf{k}}} \sum_{\mathbf{k}} f(\mathbf{k}), \tag{2.69}$$

where $N_{\mathbf{k}}$ is the number of chosen **k**-points. This set of reciprocal points is commonly referred to as the *k-point mesh*. Figure 2.5 presents an example of a **k**-mesh for the BZ of the two dimensional square lattice. It also shows the irreducible Brillouin zone BZ, which is the smallest possible subspace of the Brillouin zone that can be used to construct the whole BZ by the symmetry operations of the crystal's space group.

As a result of the crystal's symmetry, which is determined by the lattice and atom positions, some of the **k**-points are equivalent. This means that they can be mapped on each other using the symmetry operations of the space group. For equivalent **k**-points, the value of the function $f(\mathbf{k})$ provides the same contribution to the sum in Eq. (2.69). It is therefore convenient to only calculate the value of $f(\mathbf{k})$ for the *irreducible* **k**-points of the **k**-mesh. This is the smallest set of inequivalent **k**-points of the **k**-mesh from which the whole **k**-mesh can be retrieved through a combination of symmetry operations. In case the **k**-mesh has the same symmetry as the space group, the irreducible **k**-points are found in the BZ.

If $\{\mathbf{k}_i\}$ is the set of irreducible **k**-points, the sum in Eq. (2.69) can be expressed as:

$$\frac{1}{N_{\mathbf{k}}} \sum_{\mathbf{k}} f(\mathbf{k}) = \sum_{\mathbf{k}_i} w_{\mathbf{k}_i} f(\mathbf{k}_i), \tag{2.70}$$

where $w_{\mathbf{k}_i}$ is the weight of $\mathbf{k}_i$. The weight of an irreducible **k**-point is defined as the number of equivalent **k**-points of the total mesh that correspond to this **k**-point, divided by the total number of **k**-points in the mesh (Fig. 2.5).

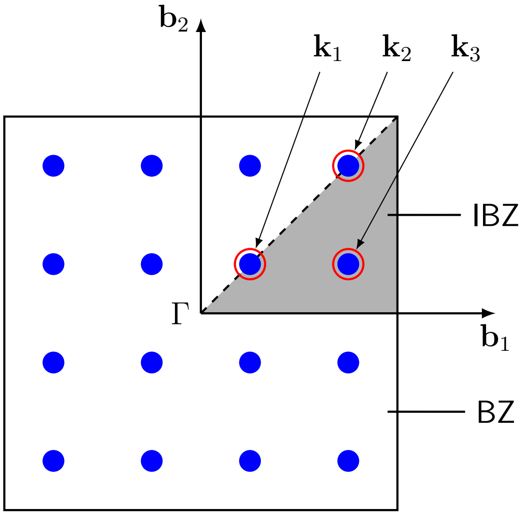


Figure 2.5: Illustration of the **k**-point mesh for a two dimensional square lattice. The irreducible **k**-points are highlighted by a red circle. The weights for the three irreducible **k**-points are: $w_{\mathbf{k}_1} = \frac{1}{4}$, $w_{\mathbf{k}_2} = \frac{1}{4}$, $w_{\mathbf{k}_3} = \frac{1}{2}$.

A common collection of **k**-points used in DFT calculations is the Monkhorst-Pack mesh [30], which is an equidistant grid defined as:

$$\mathbf{k}(n_1, n_2, n_3) = \sum_{i=1}^{3} \frac{2n_i - N_i - 1}{2N_i}\mathbf{b}_i, \tag{2.71}$$

where the $\mathbf{b}_i$ are the basis vectors of the reciprocal lattice and the n_i run from 1 to N_i. After the Monkhorst-Pack mesh is created, it is usually shifted in order to make it centered around the Γ-point. This way we can increase the symmetry of the **k**-mesh, which lowers the amount of irreducible **k**-points, making our calculations less cumbersome. This is illustrated for the two dimensional square lattice in Fig. 2.5.

2.3 Linear Response Theory

Many important properties of materials are based on its response to external electric fields. Optical properties, such as the absorption coefficient used for the calculation of the selection metric described in Section 4.3.3, depend on the dielectric response of the material in the long-wavelength limit. So does the electron energy loss function, used in Section 6.3.1 to introduce plasmonic excitations to our model of secondary electron emission. In this section I present a shortened derivation of the expression used to calculate the dielectric tensor, largely based on the excellent thesis of Judith Harl [31].

2.3.1 Response Function

The dielectric tensor expresses the optical properties of the material, which depend on the response of the material to external electric fields. In linear response theory, the response of the material to an external electric field is described by the response function $\chi(\mathbf{r},\mathbf{r}',t-t')$, which determines the change in density $\delta\rho$ at $(\mathbf{r},t)$ if the external potential undergoes a small change δv_{exp} at $(\mathbf{r}',t')$ [31]:

$$\delta\rho(\mathbf{r},t)=\int dt'\int d\mathbf{r}'\chi(\mathbf{r},\mathbf{r}',t-t')\delta v_{exp}(\mathbf{r}',t'). \tag{2.72}$$

Within the time-dependent Kohn-Sham formalism, the response function $\chi^{KS}(\mathbf{r},\mathbf{r}',t-t')$ describes the change in density in response to a small change in the effective potential $\delta V_{eff}(\mathbf{r},t)$:

$$\delta\rho(\mathbf{r},t)=\int dt'\int d\mathbf{r}'\chi^{KS}(\mathbf{r},\mathbf{r}',t-t')\delta V_{eff}(\mathbf{r}',t'). \tag{2.73}$$

Following the derivation of Pines and Nozières [32], it is possible to use perturbation theory to write the Kohn-Sham response function in frequency space as

$$\chi^{KS}(\mathbf{r},\mathbf{r}',\omega)=-\sum_{i,j}f_i\left(1-f_j\right)\left(\frac{\phi_j^*(\mathbf{r}')\phi_i(\mathbf{r}')\phi_i^*(\mathbf{r})\phi_j(\mathbf{r})}{\epsilon_j-\epsilon_i-\breve{\omega}}+\frac{\phi_i^*(\mathbf{r}')\phi_j(\mathbf{r}')\phi_j^*(\mathbf{r})\phi_i(\mathbf{r})}{\epsilon_j-\epsilon_i+\breve{\omega}}\right), \tag{2.74}$$

where $\phi_i(\mathbf{r})$ are the KS orbitals of the time-independent problem (Eqs. (2.22)), and $\breve{\omega}=\omega+i\eta$ is a complex deviation from the angular frequency ω, introduced in the perturbation calculation. The occupancies f_i are included in Eq (2.74) in such a way that the coefficient in the sum is equal to 1 only when the ith orbital is occupied and the jth orbital is unoccupied. Next, we use the periodicity of the crystal and take the Fourier transform of Eq. (2.74). Because the real space response function is invariant with respect to a shift by a lattice vector $\mathbf{R}$:

$$\chi^{KS}(\mathbf{r}+\mathbf{R},\mathbf{r}'+\mathbf{R},\omega)=\chi^{KS}(\mathbf{r},\mathbf{r}',\omega), \tag{2.75}$$

it is possible to demonstrate that the Fourier transform $\chi^{KS}(\mathbf{g},\mathbf{g}',\omega)$ is only nonzero when the two wave vectors $\mathbf{g}$ and $\mathbf{g}'$ differ by a reciprocal lattice vector: $\mathbf{g}=\mathbf{g}'+\mathbf{G}''$. Since we can write $\mathbf{g}'$ as the sum of a vector in the first Brillouin zone $\mathbf{q}$ and a reciprocal lattice vector $\mathbf{G}'$, we can do the same for $\mathbf{g}=\mathbf{g}+\mathbf{G}''=\mathbf{q}+\mathbf{G}+\mathbf{G}''=\mathbf{q}+\mathbf{G}'$. This can be used to express the momentum space response function as $\chi^{KS}_{\mathbf{G},\mathbf{G}'}(\mathbf{q},\omega)$, where $\mathbf{G}$ and $\mathbf{G}'$ are two reciprocal lattice vectors and $\mathbf{q}$ is a reciprocal vector lying in the first Brillouin zone. Next, we use the Bloch theorem to write the states as we did in Section 2.2.1. The combination of all these considerations allows us to formulate the response function in momentum space:

$$\begin{aligned}\chi^{KS}_{\mathbf{G},\mathbf{G}'}(\mathbf{q},\omega)=-\frac{1}{V}\sum_{n\mathbf{k},m\mathbf{k}'}f_{n\mathbf{k}}(1-f_{m\mathbf{k}'})&\left(\frac{\langle\phi_{m\mathbf{k}'}|e^{i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}}|\phi_{n\mathbf{k}}\rangle\langle\phi_{n\mathbf{k}}|e^{-i(\mathbf{q}+\mathbf{G}')\cdot\mathbf{r}'}|\phi_{m\mathbf{k}'}\rangle}{\epsilon_{m\mathbf{k}'}-\epsilon_{n\mathbf{k}}-\breve{\omega}}\right.\\&\left.+\frac{\langle\phi_{n\mathbf{k}}|e^{i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}}|\phi_{m\mathbf{k}'}\rangle\langle\phi_{m\mathbf{k}'}|e^{-i(\mathbf{q}+\mathbf{G}')\cdot\mathbf{r}'}|\phi_{n\mathbf{k}}\rangle}{\epsilon_{m\mathbf{k}'}-\epsilon_{n\mathbf{k}}+\breve{\omega}}\right).\end{aligned} \tag{2.76}$$

This can alternatively be written as [31]

$$\begin{aligned}\chi^{KS}_{\mathbf{G},\mathbf{G}'}(\mathbf{q},\omega)=-\frac{1}{V}\sum_{n,m;\mathbf{k}}&f_{n\mathbf{k}}\langle u_{m\mathbf{k}+\mathbf{q}}|e^{i\mathbf{G}\cdot\mathbf{r}}|u_{n\mathbf{k}}\rangle\langle u_{n\mathbf{k}}|e^{-i\mathbf{G}'\cdot\mathbf{r}'}|u_{m\mathbf{k}+\mathbf{q}}\rangle\\&\cdot\left(\frac{1}{\epsilon_{m\mathbf{k}+\mathbf{q}}-\epsilon_{n\mathbf{k}}-\breve{\omega}}+\frac{1}{\epsilon_{m\mathbf{k}+\mathbf{q}}-\epsilon_{n\mathbf{k}}+\breve{\omega}}\right),\end{aligned} \tag{2.77}$$

where

$$\langle u_{m\mathbf{k}+\mathbf{q}}|e^{i\mathbf{G}\cdot\mathbf{r}}|u_{n\mathbf{k}}\rangle \equiv \int_V d\mathbf{r}\; u^*_{m\mathbf{k}+\mathbf{q}}(\mathbf{r})e^{i\mathbf{G}\cdot\mathbf{r}}u_{n\mathbf{k}}(\mathbf{r}) \tag{2.78}$$

are integrals over the volume V of the unit cell.

2.3.2 Dielectric Tensor

In order to derive the microscopic dielectric function, we use the *random phase approximation* (RPA), which connects the dielectric function to the Kohn-Sham response function as follows [31]:

$$\varepsilon_{\mathbf{G},\mathbf{G}'}(\mathbf{q},\omega) = \delta_{\mathbf{G},\mathbf{G}'} - \frac{4\pi}{|\mathbf{G}+\mathbf{q}||\mathbf{G}'+\mathbf{q}|}\chi^{KS}_{\mathbf{G},\mathbf{G}'}(\mathbf{q},\omega), \tag{2.79}$$

where $\delta_{\mathbf{G},\mathbf{G}'}$ is the Kronecker delta. However, for wavelengths that are much larger than the periodicity of the system, we are interested in macroscopic averages of the dielectric response, not the rapid fluctuations at the microscopic level. Following the derivation of Harl [31], the relation between the microscopic and macroscopic dielectric function is

$$\varepsilon_{\mathrm{mac}}(\mathbf{q},\omega) = (\varepsilon^{-1}_{0,0}(\mathbf{q},\omega))^{-1}. \tag{2.80}$$

When we ignore local field effects, the off-diagonal elements of the microscopic dielectric tensor are neglected, and the macroscopic dielectric function can be approximated as:

$$\varepsilon_{\mathrm{mac}}(\mathbf{q},\omega) = \varepsilon_{0,0}(\mathbf{q},\omega) \underset{(2.79)}{=} 1 - \frac{4\pi}{q^2}\chi^{KS}_{0,0}(\mathbf{q},\omega). \tag{2.81}$$

For optical properties, we consider the long-wavelength limit ($\mathbf{q} \to 0$) of the macroscopic dielectric function, also written as $\varepsilon_\infty(\hat{\mathbf{q}},\omega)$:

$$\varepsilon_\infty(\hat{\mathbf{q}},\omega) = \lim_{q\to 0}\varepsilon_{\mathrm{mac}}(\mathbf{q},\omega) = 1 - \lim_{q\to 0}\frac{4\pi}{q^2}\chi^{KS}_{0,0}(\mathbf{q},\omega). \tag{2.82}$$

Using these approximations in combination with Eq. (2.77) allows us to evaluate the imaginary part of the macroscopic dielectric function $\varepsilon_\infty(\hat{\mathbf{q}},\omega) = \varepsilon^{(1)}_\infty(\hat{\mathbf{q}},\omega) + i\varepsilon^{(2)}_\infty(\hat{\mathbf{q}},\omega)$ [31]:

$$\varepsilon^{(2)}_\infty(\hat{\mathbf{q}},\omega) = \frac{4\pi}{V}\lim_{q\to 0}\frac{1}{q^2}\sum_{n,m,\mathbf{k}} f_{n\mathbf{k}}|\langle u_{m\mathbf{k}+\mathbf{q}}|u_{n\mathbf{k}}\rangle|^2 \cdot [\delta(\epsilon_{m\mathbf{k}} - \epsilon_{n\mathbf{k}} - \omega) - \delta(\epsilon_{m\mathbf{k}} - \epsilon_{n\mathbf{k}} + \omega)]. \tag{2.83}$$

We introduce the macroscopic dielectric tensor by setting:

$$\varepsilon^{(2)}_\infty(\hat{\mathbf{q}},\omega) = \sum_{\alpha,\beta}\hat{\mathbf{q}}_\alpha\varepsilon^{(2)}_{\alpha\beta}(\omega)\hat{\mathbf{q}}_\beta. \tag{2.84}$$

The imaginary part of the dielectric tensor is then calculated as:

$$\varepsilon^{(2)}_{\alpha\beta}(\omega) = \frac{4\pi}{V}\lim_{q\to 0}\frac{1}{q^2}\sum_{n,m;\mathbf{k}} f_{n\mathbf{k}}\langle u_{m\mathbf{k}+q\mathbf{e}_\alpha}|u_{n\mathbf{k}}\rangle\langle u_{n\mathbf{k}}|u_{m\mathbf{k}+q\mathbf{e}_\beta}\rangle \cdot [\delta(\epsilon_{m\mathbf{k}}-\epsilon_{n\mathbf{k}}-\omega) - \delta(\epsilon_{m\mathbf{k}}-\epsilon_{n\mathbf{k}}+\omega)], \tag{2.85}$$

whereas the real part is found using the Kramers-Kronig transformation:

$$\varepsilon^{(1)}_{\alpha\beta}(\omega) = 1 + \frac{2}{\pi}\mathcal{P}\int_0^\infty \frac{\varepsilon^{(2)}_{\alpha\beta}(\omega')\omega'}{(\omega')^2-\omega^2}d\omega', \tag{2.86}$$

where $\mathcal{P}$ is the Cauchy Principal value [33].

I finish this section with a couple of notes. First, by introducing the **k**-mesh in Section 2.2.5, we can restrict the summation in Eq. (2.85) to the irreducible **k**-points:

$$\begin{aligned}\varepsilon^{(2)}_{\alpha\beta}(\omega) = \frac{4\pi}{V}\lim_{q\to 0}\frac{1}{q^2}\sum_{n,m;\mathbf{k}_i} f_{n\mathbf{k}_i} w_{\mathbf{k}_i} \langle u_{m\mathbf{k}_i+q\mathbf{e}_\alpha}|u_{n\mathbf{k}_i}\rangle \langle u_{n\mathbf{k}_i}|u_{m\mathbf{k}_i+q\mathbf{e}_\beta}\rangle \\ \cdot[\delta(\epsilon_{m\mathbf{k}_i}-\epsilon_{n\mathbf{k}_i}-\omega)-\delta(\epsilon_{m\mathbf{k}_i}-\epsilon_{n\mathbf{k}_i}+\omega)],\end{aligned}$$

where $\mathbf{k}_i$ are the irreducible **k**-points with corresponding weights $w_{\mathbf{k}_i}$. Second, for semiconductors we can only consider interband transitions ($n \neq m$) between the valence and conduction band in the calculation of the dielectric tensor. Moreover, in this case the valence bands are fully occupied ($f_{n\mathbf{k}} = 2$), simplifying the expression further:

$$\varepsilon^{(2)}_{\alpha\beta}(\omega) = \frac{4\pi}{V}\lim_{q\to 0}\frac{1}{q^2}\sum_{c,v,\mathbf{k}_i} 2w_{\mathbf{k}_i}\ \langle u_{c\mathbf{k}_i+q\mathbf{e}_\alpha}|u_{v\mathbf{k}_i}\rangle \langle u_{v\mathbf{k}_i}|u_{c\mathbf{k}_i+q\mathbf{e}_\beta}\rangle \cdot \delta(\epsilon_{c\mathbf{k}_i}-\epsilon_{v\mathbf{k}_i}-\omega)), \tag{2.87}$$

where c and v denote the conduction and valence orbitals.

Finally, we note that for each frequency ω it is possible to diagonalize the dielectric tensor. If we write the diagonal components as $\varepsilon_{11}(\omega) = \varepsilon_x(\omega)$, $\varepsilon_{22}(\omega) = \varepsilon_y(\omega)$ and $\varepsilon_{33}(\omega) = \varepsilon_z(\omega)$, we can calculate the complex index of refraction from the diagonal components of the diagonalized dielectric tensor as:

$$(n_c)_\alpha(\omega) = \sqrt{\varepsilon_\alpha(\omega)} \qquad \alpha = x, y, z. \tag{2.88}$$

Using this relation, we can find the real index of refraction and extinction coefficient with the rules for taking the square root of a complex number:

$$\begin{aligned} n_\alpha(\omega) &= \sqrt{\frac{|\varepsilon_\alpha(\omega)| + \varepsilon^{(1)}_\alpha(\omega)}{2}} \\ k_\alpha(\omega) &= \sqrt{\frac{|\varepsilon_\alpha(\omega)| - \varepsilon^{(1)}_\alpha(\omega)}{2}} \end{aligned} \qquad \alpha = x, y, z. \tag{2.89}$$

2.3.3 Drude model

For metals, bands can be partially occupied, which means that we have to consider intraband transitions, i.e. where $n = m$ in the expressions of the previous section. In the long-wavelength limit, the intraband part of the dielectric tensor is included using the Drude model [31]:

$$\varepsilon^{(2),intra}_{\alpha\beta}(\omega) = \frac{\gamma\omega^2_{\alpha\beta}}{\omega(\omega^2+\gamma^2)}, \tag{2.90}$$

where γ is the damping parameter. The so-called intraband plasma frequency (squared) $\omega^2_{\alpha\beta}$ is calculated from first-principles using the expression [31]:

$$\omega^2_{\alpha\beta} = \frac{4\pi}{V} \sum_{n,\mathbf{k}} \frac{\partial f_{n\mathbf{k}}}{\partial \epsilon_{n\mathbf{k}}} \left(\mathbf{e}_\alpha \frac{\partial \epsilon_{n\mathbf{k}}}{\partial \mathbf{k}} \right) \left(\mathbf{e}_\beta \frac{\partial \epsilon_{n\mathbf{k}}}{\partial \mathbf{k}} \right). \tag{2.91}$$

The imaginary part of the total dielectric tensor is the sum of the inter- and intraband contributions:

$$\varepsilon^{(2)}_{\alpha\beta}(\omega) = \varepsilon^{(2),inter}_{\alpha\beta}(\omega) + \varepsilon^{(2),intra}_{\alpha\beta}(\omega), \tag{2.92}$$

where the interband part is calculated via Eq. (2.85). The real part is once again obtained via the Kramers-Kronig relationship (Eq. (2.86)).

2.4 Transition state theory

When considering a transition between two different states of a system, a first property of interest is the difference in energy ΔE between the final and initial state of the system, also referred to as the *reaction energy*. For the comparison of the chalcopyrite and CuAu phases in Section 4.4, for example, comparing the formation energy of both structures can help us understand the likelihood of each phase being present after synthesis. In the case of the O-O dimerization in Section 5.2.3, the difference in energy between the initial and final state provides information on whether the formation of a specific dimer stabilizes the structure.

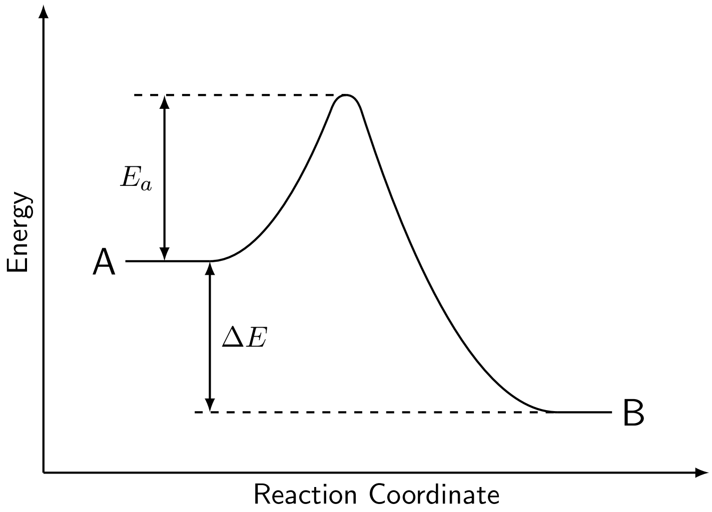


Figure 2.6: Schematic of a transition from state A to state B.

However, often we are also interested in the kinetics of the transition, i.e. the rate at which a reaction will occur. One way of investigating this is by performing *ab initio* molecular dynamics (AIMD) calculations. Unfortunately, for many reactions performing AIMD simulations over a sufficiently long time scale to obtain relevant statistics is far too computationally demanding. A popular alternative for studying the kinetics is by obtaining the kinetic barrier, also known as the activation energy E_a, directly by determining the minimum energy path for the transition. Figure 2.6 shows a schematic of the energy barrier of a transition.

2.4.1 Nudged elastic band method

A well established method for finding the minimum energy path between the initial and final state of the system is the nudged elastic band (NEB) method [34, 35]. In this technique, a path is constructed by considering a number of images in between the initial and final state, usually obtained by linearly interpolating the atomic positions of the initial and final structures. These images are then connected via fictitious springs in order to keep them spread somewhat evenly along the path, as simply optimizing the geometry of the images would revert them to either the initial or final state. This connected system is then optimized until the total force, i.e. a suitably chosen combination of the spring force and the *true* force, is zero for each atom of each image.

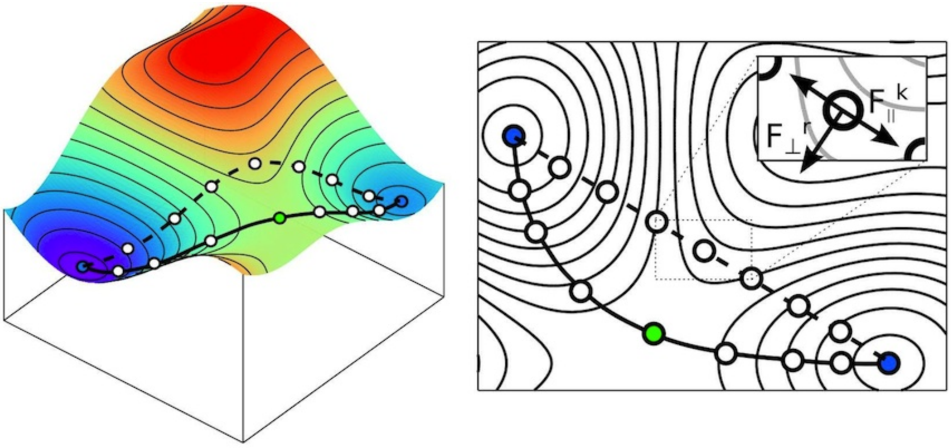


Figure 2.7: Nudged elastic band method. Taken from [36].

Write the $N+1$ images as $[\mathbf{R}_0, \mathbf{R}_1, \mathbf{R}_2, ..., \mathbf{R}_N]$, where $\mathbf{R}_0$ and $\mathbf{R}_N$ are the coordinates of the initial and final state, respectively. The basic idea of the NEB method is that the total force acting on an image is the sum of the spring force along the local tangent and the perpendicular part of the true force, i.e. the force on the image produced by the energy landscape:

$$\mathbf{F}_i = \mathbf{F}_i^s|_{\parallel} - \nabla E(\mathbf{R}_i)|_{\perp}. \tag{2.93}$$

The perpendicular part of the true force is calculated from the total true force and the projection on the local tangent $\hat{\boldsymbol{\tau}}_i$:

$$\nabla E(\mathbf{R}_i)|_{\perp} = \nabla E(\mathbf{R}_i) - \nabla E(\mathbf{R}_i)|_{\parallel} = \nabla E(\mathbf{R}_i) - [\nabla E(\mathbf{R}_i) \cdot \hat{\boldsymbol{\tau}}_i] \cdot \hat{\boldsymbol{\tau}}_i, \tag{2.94}$$

whereas the spring force depends on the spring constant k:

$$\mathbf{F}_i^s|_{\parallel} = k\left(|\mathbf{R}_{i+1} - \mathbf{R}_i| - |\mathbf{R}_i - \mathbf{R}_{i-1}|\right)\hat{\boldsymbol{\tau}}_i. \tag{2.95}$$

The path is then optimized to the saddle point by a suitable minimization algorithm using the force from Eq. (2.93). Figure 2.7 shows an example of the initial and final path of the NEB method for a simple two dimensional energy landscape. Note that the simplest definition of the tangent

$$\hat{\boldsymbol{\tau}}_i = \frac{\mathbf{R}_{i+1} - \mathbf{R}_i}{|\mathbf{R}_{i+1} - \mathbf{R}_i|}, \tag{2.96}$$

has been shown to lead to kinks in the minimum energy path, which interferes with the convergence of the NEB. An improved tangent was developed by Henkelman and Jonsson [35] by bisecting the two unit vectors

$$\boldsymbol{\tau}_i = \frac{\mathbf{R}_i - \mathbf{R}_{i-1}}{|\mathbf{R}_i - \mathbf{R}_{i-1}|} + \frac{\mathbf{R}_{i+1} - \mathbf{R}_i}{|\mathbf{R}_{i+1} - \mathbf{R}_i|}, \tag{2.97}$$

and subsequently normalizing the tangent vector $\hat{\boldsymbol{\tau}}_i = \boldsymbol{\tau}_i / |\boldsymbol{\tau}_i|$.

2.4.2 Climbing Image Modification

A common issue with the NEB method presented above is that the saddle point is in between two images along the path, which can lead to an underestimation of the activation energy. Although a spline interpolation of the energy barrier can improve the result, a better solution to this problem is the so-called climbing image method [37], which introduces a small modification to the NEB algorithm. After several iterations of a regular NEB calculation, the algorithm searches the image with the highest energy along the path. For this image, the force defined in Eq. 2.93 is substituted by:

$$\mathbf{F}_i^{\max} = -\nabla E(\mathbf{R}_i^{\max}) + 2\nabla E(\mathbf{R}_i^{\max})|_{\parallel} \tag{2.98}$$

Simply put, the direction of the component of the true force along the path is inverted, pushing the image towards the saddle point. Note that the image with the maximum energy is no longer affected by the spring force defined in Eq. 2.95. Figure 2.8 compares the difference in minimum energy path between both methods for CH_4 dissociative adsorption on a Ir(111) surface. We can see that the activation energy is significantly higher when the climbing image modification is applied.

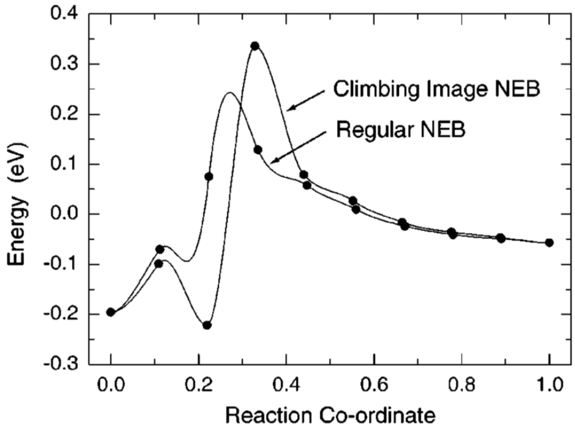


Figure 2.8: NEB calculation of the minimum energy path for CH_4 dissociative adsorption on a Ir(111) surface, using both the regular NEB algorithm as well as the climbing image modification. [37]

Bibliography

[1] M. Springborg and Y. Dong. “Density functional theory”. In: *Handbook of Solid State Chemistry*. Weinheim, Germany: Wiley-VCH Verlag GmbH & Co. KGaA, 2017, pp. 117–138. ISBN: 9781420045451. DOI: `10.1201/9781420045451`.

[2] M. Born and R. Oppenheimer. “Zur Quantentheorie der Molekeln”. *Ann. Phys.* 389 (1927), pp. 457–484. DOI: `10.1002/andp.19273892002`.

[3] P. Hohenberg and W. Kohn. “Inhomogeneous electron gas”. *Phys. Rev.* 136 (1964), B864–B871. DOI: `10.1103/PhysRev.136.B864`.

[4] S. Cottenier. “Density Functional Theory and the Family of (L)APW-methods: a step-by-step introduction”. ISBN 978-90-807215-1-7 (freely available at `http://susi.theochem.tuwien.ac.at/reg_user/textbooks/DFT_and_LAPW_2nd.pdf`). 2002-2013 (2$^{\text{nd}}$ edition).

[5] R. M. Martin. *Electronic Structure: Basic Theory and Practical Methods*. Cambridge University Press, 2012. ISBN: 9780511805769. DOI: `10.1017/CBO9780511805769`.

[6] W. Kohn and L. J. Sham. “Self-consistent equations including exchange and correlation effects”. *Phys. Rev.* 140 (1965), A1133–A1138. DOI: `10.1103/PhysRev.140.A1133`.

[7] N. Argaman and G. Makov. “Density functional theory: An introduction”. *Am. J. Phys.* 68 (2002), pp. 69–79. DOI: `10.1119/1.19375`.

[8] J. P. Perdew. “Jacob’s ladder of density functional approximations for the exchange-correlation energy”. In: *AIP Conf. Proc.* Vol. 577. 1. AIP, 2001, pp. 1–20. DOI: `10.1063/1.1390175`.

[9] L. W. Wang. “Divide-and-conquer quantum mechanical material simulations with exascale supercomputers”. *Natl. Sci. Rev.* 1 (2014), pp. 604–617. DOI: `10.1093/nsr/nwu060`.

[10] J. P. Perdew et al. “Atoms, molecules, solids, and surfaces: Applications of the generalized gradient approximation for exchange and correlation”. *Phys. Rev. B* 46 (1992), pp. 6671–6687. DOI: `10.1103/PhysRevB.46.6671`.

[11] A. D. Becke. “Density-functional exchange-energy approximation with correct asymptotic behavior”. *Phys. Rev. A* 38 (1988), pp. 3098–3100. DOI: `10.1103/PhysRevA.38.3098`.

[12] T. W. Keal and D. J. Tozer. “A semiempirical generalized gradient approximation exchange-correlation functional”. *J. Chem. Phys.* 121 (2004), pp. 5654–5660. DOI: `10.1063/1.1784777`.

[13] J. P. Perdew, K. Burke, and M. Ernzerhof. “Generalized gradient approximation made simple”. *Phys. Rev. Lett.* 77 (1996), pp. 3865–3868. DOI: `10.1103/PhysRevLett.77.3865`.

[14] J. Kohanoff. *Electronic structure calculations for solids and molecules: Theory and computational methods*. Cambridge University Press, 2006. ISBN: 9780511755613. DOI: `10.1017/CBO9780511755613`.

[15] J. Sun, A. Ruzsinszky, and J. Perdew. “Strongly Constrained and Appropriately Normed Semilocal Density Functional”. *Phys. Rev. Lett.* 115 (2015), p. 036402. DOI: `10.1103/PhysRevLett.115.036402`.

[16] F. Tran, P. Blaha, and K. Schwarz. “Band gap calculations with Becke-Johnson exchange potential”. *J. Phys. Condens. Matter* 19 (2007), p. 196208. DOI: `10.1088/0953-8984/19/19/196208`.

[17] O. Bengone et al. “Implementation of the projector augmented-wave LDA+U method: Application to the electronic structure of NiO”. *Phys. Rev. B* 62 (2000), pp. 16392–16401. DOI: `10.1103/PhysRevB.62.16392`.

[18] J. C. Slater. “A simplification of the Hartree-Fock method”. *Phys. Rev.* 81 (1951), pp. 385–390. DOI: `10.1103/PhysRev.81.385`.

[19] J. Heyd, G. E. Scuseria, and M. Ernzerhof. “Hybrid functionals based on a screened Coulomb potential”. *J. Chem. Phys.* 118 (2003), pp. 8207–8215. DOI: `10.1063/1.1564060`.

[20] S. L. Dudarev et al. “Electron-energy-loss spectra and the structural stability of nickel oxide: An LSDA+U study”. *Phys. Rev. B* 57 (1998), pp. 1505–1509. DOI: `10.1103/PhysRevB.57.1505`.

[21] P. E. Blöchl. “Projector augmented-wave method”. *Phys. Rev. B* 50 (1994), pp. 17953–17979. DOI: `10.1103/PhysRevB.50.17953`.

[22] D. Joubert. “From ultrasoft pseudopotentials to the projector augmented-wave method”. *Phys. Rev. B - Condens. Matter Mater. Phys.* 59 (1999), pp. 1758–1775. DOI: `10.1103/PhysRevB.59.1758`.

[23] R. P. Feynman. “Forces in molecules”. *Phys. Rev.* 56 (1939), pp. 340–343. DOI: `10.1103/PhysRev.56.340`.

[24] H. Hellmann and F. Deuticke. “Einführung in die Quantenchemie”. *Angew. Chemie* 54 (1937), pp. 156–156. DOI: `10.1002/ange.19410541109`.

[25] C. Kittel. *Introduction to Solid State Physics.* John Wiley & Sons Inc, Oct. 28, 2004. 704 pp. ISBN: 978-0-471-41526-8.

[26] J. G. Hill. “Gaussian basis sets for molecular applications”. *Int. J. Quantum Chem.* 113 (2013), pp. 21–34. DOI: `10.1002/qua.24355`.

[27] D. J. Singh and L. Nordstrom. *Planewaves, pseudopotentials and the LAPW method: Second edition.* Springer US, 2006. ISBN: 0387287809. DOI: `10.1007/978-0-387-29684-5`.

[28] J. C. Phillips and L. Kleinman. “New method for calculating wave functions in crystals and molecules”. *Phys. Rev.* 116 (1959), pp. 287–294. DOI: `10.1103/PhysRev.116.287`.

[29] C. Rostgaard. “The Projector Augmented-wave Method”. Available as e-print: arXiv:0910.1921v2. 2010.

[30] H. J. Monkhorst and J. D. Pack. “Special points for Brillouin-zone integrations”. *Phys. Rev. B* 13 (1976), pp. 5188–5192. DOI: `10.1103/PhysRevB.13.5188`.

[31] J. Harl. “The linear response function in density functional theory: Optical spectra and improved description of the electron correlation”. PhD thesis. University of Vienna, 2008.

[32] P. Nozieres and D. Pines. *Theory of Quantum Liquids: Superfluid Bose Liquids.* Westview Press, 1994. ISBN: 0201408414.

[33] A. O. Gogolin. *Lectures on Complex Integration.* Ed. by E. G. Tsitsishvili and A. Komnik. Undergraduate Lecture Notes in Physics. Cham: Springer International Publishing, 2014. ISBN: 978-3-319-00211-8. DOI: `10.1007/978-3-319-00212-5`.

[34] H. Jónsson, G. Mills, and K. W. Jacobsen. “Nudged elastic band method for finding minimum energy paths of transitions”. In: *Classical and Quantum Dynamics in Condensed Phase Simulations.* World Scientific, 1998, pp. 385–404. ISBN: 978-981-02-3498-0. DOI: `10.1142/9789812839664_0016`.

[35] G. Henkelman and H. Jónsson. “Improved tangent estimate in the nudged elastic band method for finding minimum energy paths and saddle points”. *J. Chem. Phys.* 113 (2000), pp. 9978–9985. DOI: `10.1063/1.1323224`.

[36] *Nudged Elastic Band method.* URL: `https://tinyurl.com/u6eycye` (visited on 01/02/2020).

[37] G. Henkelman, B. P. Uberuaga, and H. Jónsson. “A climbing image nudged elastic band method for finding saddle points and minimum energy paths”. *J. Chem. Phys.* 113 (2000), p. 9901. DOI: `10.1063/1.1329672`.

Chapter 3

Automation

> *"Give me six hours to chop down a tree and I will spend the first four sharpening the axe."*
>
> Abraham Lincoln

Once the budding computational scientist has learned the basics on how to use their *ab initio* code of choice, he or she will begin their career by manually setting up the required input files, usually from a set copied from a previous calculation. This approach, however, has several downsides. Manually adjusting the input tags for each calculation is very time consuming, especially when a lot of calculations have to be performed. Certainly, any supervisor can agree this time would be better spent analysing results and writing papers. Moreover, some properties cannot be feasibly calculated in such a manner, as the time required to set up all of the calculations would be prohibitively large. Finally, adjusting input settings manually often leads to user-related errors, especially when calculations are set up late at night.

When a large amount of similar calculations have to be performed, it is much more sensible to design an automated workflow that performs the required steps for whatever series of calculations you have in mind. Although it is certainly possible to write your own scripts, there are already several software packages available that provide a useful framework for this purpose, based on for example pymatgen [1] or the Atomic Simulation Environment (ASE) [2]. For my work, I have exclusively relied on Fireworks [3], a powerful Python package for designing, executing and managing computational workflows.

In this chapter I give an overview of all of the workflows that were used to calculate the various properties required for my research. I also explain some more technical details of performing certain calculations using our chosen implementation of the DFT formalism: the Vienna ab initio simulation package (VASP). After the short introduction in Section 3.1, the chapter continues with detailing the building blocks of a workflow in Section 3.2. Next, Section 3.3 describes the workflows used to obtain the results presented in future chapters. For each workflow, there is a corresponding example Jupyter notebook that details and executes the various steps performed to set up the calculation and run the workflow. Finally, Section 3.4 discusses some details on how to adjust the computational settings used for the workflows.

3.1 Introduction

Many *ab initio* methods for calculating properties rely on a fairly fixed algorithm of reoccurring steps. Even when the exact steps depend on the result of a previous calculation, it is usually feasible to come up with some sort of well-defined schematic, or workflow, on how to best calculate the desired property for any structure. Fireworks is a well-established Python package for designing such a workflow.

As an example, say you want to calculate the density of states of a certain structure, e.g. the chalcopyrite phase of $\mathrm{CuInSe_2}$. After obtaining an initial geometry, either by constructing it manually, or - more conveniently - by extracting it from a trustworthy database [4], the first step is to optimize the structure computationally. Once you have obtained the optimized geometry, the next step is calculate the density of states using two consecutive VASP calculations: first one to calculate the charge density of the electrons using a relatively sparse k-point sampling of the Brillouin zone, then a calculation that keeps the charge density fixed and determines the energies of the electronic states with a much denser k-point mesh. Such a fixed process for calculating a specific property is an ideal application of a workflow, especially if you want to obtain this property for a large number of structures.

Workflows are assembled from connected steps, called `Firework`s, which in turn consist of a set of consecutive `Firetask`s. How the various tasks that need to be performed are organized in `Firework`s and `Firetask`s is up to the user, but the developers do provide some general guidelines on how to design your workflows. All of my workflows were constructed by bundling all of the tasks required for a single VASP calculation into one `Firework`, and then connecting these `Firework`s to create a workflow for each of the properties required for my research.

3.2 Firetasks

Let's begin with an overview of the smallest building blocks of the workflows: the `Firetask`. Each `Firework` of the workflow consists of a list of `Firetask`s, which are executed sequentially in the order specified by the user when initializing the `Firework`. While the workflow is running, all `Firetask`s have access to the *spec* of their parent `Firework` (`_fw_spec`). This allows the user to easily pass information between tasks. The spec also contains the launch directory of the `Firework`, so the initial `Firetask` is executed in this directory.

3.2.1 `WriteVaspFromIOSet`

As the name implies, the `WriteVaspFromIOSet` task writes the VASP input files into the current directory. Each calculation is defined by an input set, i.e. a class derived from `pymatgen`'s `DictSet` class. This class does a few things:

- Load the basic configuration from a YAML[1] file. This file contains the settings for each of the VASP input files (See Appendix A.1). Here's an (truncated) example:

```
INCAR:
  ALGO: Fast
  EDIFF: 1.0e-06
  ENCUT: 500
  ISMEAR: -5
  ...
KPOINTS:
  grid_density: 3000
```

---

[1] YAML Ain't Markup Language; IT people love recursion, also in their acronyms.

```
POTCAR:
  Ac: Ac
  Ag: Ag
  Al: Al
  Am: Am
  Ar: Ar
  As: As
  At: At_d
  ...
```

Besides the `INCAR` tags, the YAML file also specifies the sampling of the Brillouin zone. This is done based on a specified density, as the number of k-points should of course depend on the size of the reciprocal unit cell. Finally, the `POTCAR` file is specified for each element, which mainly determines the number of valence (and core) electrons to consider in the PAW formalism. Each `POTCAR` file has a specific name, e.g. default file for At only considers the $6p$ electrons as valence electrons, but the `At_d` file also includes the $5d$ electrons.

- Make adjustments to the configuration based on the structure and what calculation has to be performed. For example, in case the provided structure has magnetic moments specified, the `ISPIN` tag is set to 2 in order to perform a spin polarized calculation. Or, for a geometry optimization of a slab, selective dynamics is applied to fix a certain selection of the atoms based on user-provided settings. Based on the input set, the configuration of the YAML file is also adjusted, e.g. `SlabOptimizeSet` is based on the `optimizeSet.yaml` file, but automatically sets `ISIF` to 2, fixing the unit cell of the calculation.
- Write the output files using the `write_input()` method. Every `DictSet` subclass requires the structure as an input argument, so based on the structure and configuration, all of the VASP input files can be written. Note that the `MAGMOM` tag, which specifies the magnetic moments in the structure, is also automatically added to the `INCAR` file, which can be a real headache to do manually for large complicated structures.

The default method for initializing the `WriteVaspFromIOSet` class is by providing either an initialized input set or the structure and the name of the input set of the calculation. However, in some cases the structure cannot be provided when the workflow is set up, as it relies on a previous geometry optimization. In this case the structure will be extracted from the directory of the parent `Firework`. Note that if the input set is defined by the name, either a structure or parent `Firework` must be provided, else the `Firetask` has no recourse for obtaining the structure.

3.2.2 VaspParallelizationTask

As most DFT calculations require a substantial amount of computational resources, it is important to properly parallelize the workload over the various nodes/cores of a cluster, as failing to do so can severely affect the performance. For VASP, the parallelization can be controlled by several parameters:

- `KPAR`: Determines the parallelization over the k-points. VASP treats `KPAR` k-points in parallel, i.e. it divides the number of available cores (`#CORES`) and k-points in `KPAR` groups and then assigns each group of cores to work on a group of k-points, one k-point at a time. As `KPAR` divides the k-points of the calculation in groups, `KPAR` is optimally chosen to be a divisor of the number of k-points. However, this rule should **not** be set in stone, as often choosing a higher `KPAR` that is not a divisor of the number of nodes can result in an improved optimization.

- `NPAR`: A second way in which VASP allows parallelization is over the electronic bands. For each k-point, there are `#CORES/KPAR` cores working on it at a single time. Using `NPAR`, we can further divide the cores per k-point into `NPAR` groups, i.e. treat `NPAR` bands in parallel using `#CORES/KPAR/NPAR` cores. Note that similar to the `KPAR` and the number of k-points, `NPAR` should optimally be a divisor of the number of bands to avoid inactive cores. VASP actually enforces this by setting the number of bands (`NBANDS`) to the smallest multiple of `NPAR` higher than the requested (or default) bands.
- `NCORE`: The number of cores that are working together on an individual band. For a specified `NPAR`, `NCORE` is automatically set using the following relation:

  $$\texttt{NCORE} = \frac{\texttt{\#CORES / KPAR}}{\texttt{NPAR}}$$

  In other words, the user should specify either `NPAR` or `NCORE`, as they are directly connected. As is said in the VASP manual, `NPAR` is preferred, i.e. if you do set both `NPAR` and `NCORE`, the `NCORE` value is ignored.
- `NSIM`: This input tag defines the number of bands that are treated simultaneously in the blocked mode of the `RMM-DIIS`[2] algorithm. This allows VASP to exploit matrix BLAS[3] operations instead of matrix-vector operations, which could lead to a speed up on some machines. Peter Larsson wrote an interesting blog post on the topic. Overall, the gains are smaller compared to successfully utilizing `NPAR` and `KPAR`, but it might be worth experimenting with this value in case you rely on the `RMM-DIIS` algorithm.

As an example, consider the following parallelization settings for a calculation that uses 32 cores: `KPAR = 4`; `NPAR = 2`, shown schematically in Fig. 3.1. First, the total number of cores is divided among `KPAR = 4` groups, resulting in `#CORES/KPAR = 8` cores working on each group of kpoints. Next, the cores are split into `NPAR = 2` groups, which means that `NCORE = #CORES/KPAR/NPAR` = 4 cores will be working simultaneously on each band.

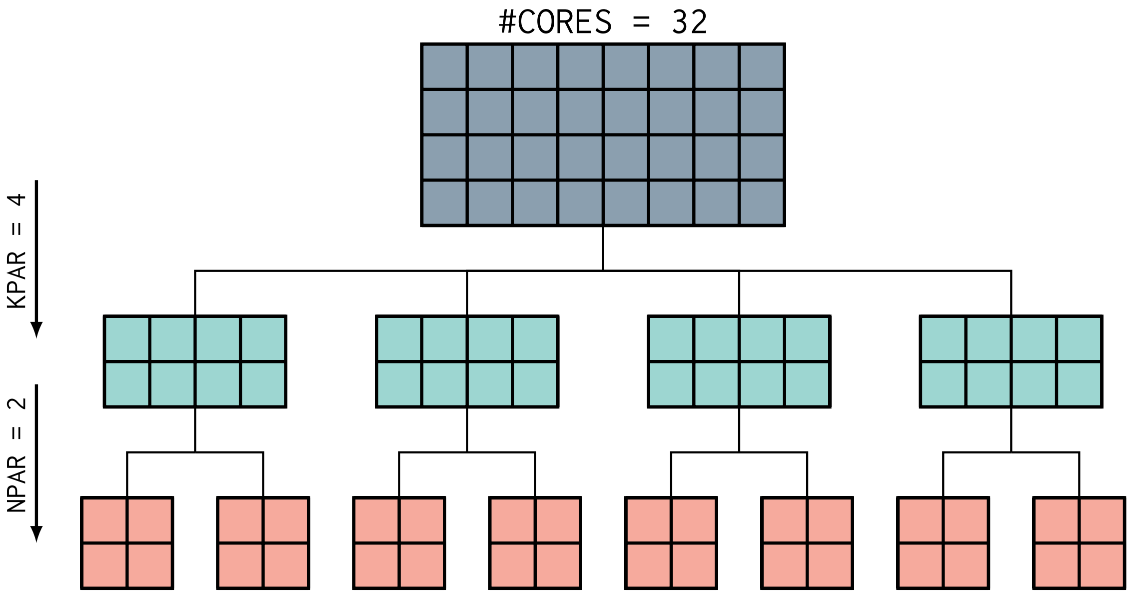


Figure 3.1: Schematic of a parallelization for 32 cores, KPAR = 4 and NPAR = 2.

---

[2] Residual minimisation/direct inversion in the iterative subspace [5].

[3] Basic linear algebra subprograms.

Most applications for calculation time on a higher Tier system, such as the BrEniac cluster of the Flemish supercomputer center (VSC), or the PRACE infrastructure of the European Union, require the user to demonstrate the scaling of their requested calculations is optimal. This involves testing a range of parallelization settings for an increasing number of nodes, and comparing the speed of the calculation at higher node usage in order to analyse the efficiency. An example of such an analysis can be found in Appendix A.3. However, performing these tests for every structure/calculation in a workflow would defeat the purpose, as you would spend more time on running tests than actually producing results.

`VaspParallelizationTask` is my humble approach to automating this process, based on a series of tests which I have performed on the machines I use most frequently. This is still very much a work in progress, but the algorithm in its current form works as follows:

- Make a list of reasonable `KPAR` values. The algorithm enforces that this is a divisor of the number of cores, and then calculates the amount of resources that would be wasted in each run over the k-points. To illustrate this, consider the following example: say you want to do a geometry optimization with a k-point sampling that has 7 irreducible points in the IBZ. You're running your calculation on a cluster that has 20 cores per node, and so there is no shared divisor for these two numbers larger than one. That said, choosing `KPAR = 1` would be a mistake, since a well configured parallelization can more than make up for the time lost due to inactive cores. So, say you choose `KPAR = 2`. In this case, VASP divides the cores over two groups, and hence solves the Kohn-Sham equations of two k-points simultaneously.

  After finishing 6 of the k-points, there is only one k-point left, so while the final k-point is being calculated there will be 10 inactive cores. `VaspParallelizationTask` defines the core waste as the number of inactive cores per run divided by the number of k-point groups, i.e. 4 in the example here, so the core waste is 2.5. As long as the core waste is smaller than both the number of cores per node[4] and 2/3rd of the total cores, the value of `KPAR` is accepted.

- Attempt to get `NPAR` or `NCORE` as close as possible to a specified value, depending on whether the calculation is a hybrid calculation or not, respectively. Based on the tests I have run, this is either `NPAR = 8` or `NCORE = 7`. Then look for the largest accepted `KPAR` that allows for this `NPAR/NCORE` value.

For the Davidson block iteration scheme and conjugate gradient algorithm, this scheme usually results in a fairly good guess for the optimal parallelization settings. However, for the `RMM-DIIS` algorithm, it seems that the optimal value of `NPAR/NCORE` is different. More testing is required to optimize this problem further. Moreover, it is probably better to use something a little more rigorous for tuning the automation of these settings, e.g. a machine learning model based on a large set of data. In the chessboard plot of the parallelization analysis I have performed, the setting chosen by the `VaspParallelizationTask` is indicated by marking the edge of the corresponding square in red.

3.2.3 `CustodianTask`

When running large numbers of calculations using workflows, you are bound to run into issues. Calculations might fail to converge or raise errors, and nodes might crash because they run out of memory or *stercus accidit*. In some cases the Fireworks flask-based graphics user interface will show these calculations as *fizzled*, so the user can find these issues easily. However, sometimes

---

[4] In this case, you might as well run the calculation on fewer nodes. This is also true in a sense for the cores, but using partial nodes is often impractical.

the issue might pass silently and only be discovered during the analysis. Moreover, dealing with common errors which have a solution that can be programmed should be part of a robust workflow.

Custodian is a just-in-time (JIT) Python package for automated error recovery and allows the user to make much more robust workflows. In short, the `CustodianTask` runs the calculation inside a `Custodian` instance, which has `ErrorHandler`s specified during its initialization which check for errors and apply pre-configured corrections (using the aptly named methods `check()` and `correct()`). For the workflows I have used during my research, I have designed the following `ErrorHandler`s:

- `ElectronicConvergenceMonitor`: Monitors the calculation and applies a linear fit to the total residual charge (i.e. the integration of the charge density difference for an electronic step) over a range of electronic steps (`max_fit_range`). Besides the maximum tolerated incline of the fit (`max_allowed_incline`), the user can also specify the minimum number of electronic steps that a calculation must have run before a correction is triggered (`min_electronic_steps`). If the `ErrorHandler` thinks the calculation is not converging, it will terminate the run and restart the calculation with a more stable algorithm (`ALGO=VeryFast→Fast→Normal→All`). If the algorithm is already set to `All`, the `ElectronicConvergenceMonitor` will attempt to change the charge mixing settings.

- `QuotasErrorHandler`: This `ErrorHandler` was copied and stripped from Custodian's own VaspErrorHandler in order to design an `ErrorHandler` that deals specifically with the issues we encounter for the quotas workflow.

- `ParallelizationTestMonitor`: When running parallelization tests, we usually don't need the calculation to finish, since we are only interested in the average time per electronic step, which is fairly consistent for most electronic optimizations. Using this `ErrorHandler` allows the user to specify how many electronic steps are desired before the calculation should be aborted using VASP's `STOPCAR` file (`max_elec_steps`). Moreover, some parallelization settings can lead to very slow calculations, which could result in such a high electronic time step that the whole workflow is slowed down needlessly (as obviously these settings are not optimal). Hence, the user is able to specify a maximum allowed step time (`max_elec_step_time`). If `ParallelizationTestMonitor` finds that one of the previous steps is above this time, the calculation is also aborted.

Note that if any of these `ErrorHandler`s are unable to correct the calculation, e.g. because they have exhausted all options, a `NonRecoverableError` is thrown, so any failed calculations are easily found via the Fireworks management system. `CustodianTask` is the `Firetask` that actually performs the VASP calculation, within a `Custodian` with specified `ErrorHandler`s. These `ErrorHandler`s are initialized by the user when designing the workflow, and passed as a `list` to the `Custodian` input argument of any `Firework`.

3.2.4 PulayTask

A known issue with DFT calculations is the so-called Pulay stress, an almost isotropic error on the diagonal components of the stress tensor introduced by the incompleteness of the basis set [6]. Although Pulay stresses are less of an issue when a plane wave basis set is used compared to basis sets that rely on the ionic positions, one should still take care when performing geometry optimization that allow the lattice vectors to change. Figure 3.2 demonstrates the influence of a large change in lattice vectors on the effective cutoff of the basis set.

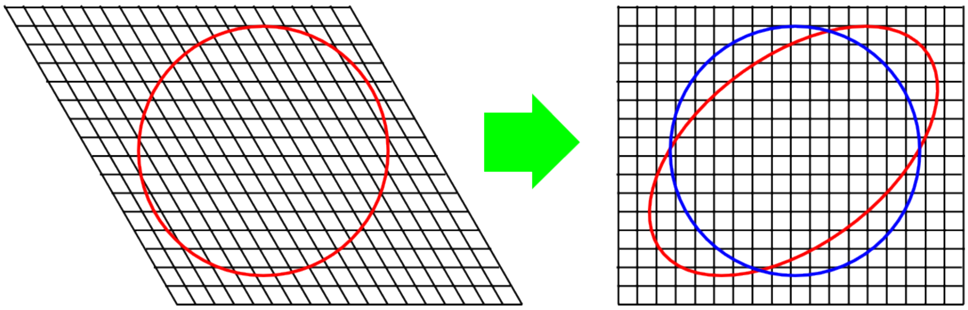


Figure 3.2: Demonstration of the influence of a change in lattice vectors on the effective energy cutoff. When the lattice changes from hexagonal (left) to a cubic (right), the corresponding energy cutoff surface changes from a sphere to an ellipsoid, effectively reducing the cutoff energy of the corresponding basis set. Taken from [7].

A common way of dealing with this issue is to perform an additional geometry optimization, as the basis set is initialized again at the start of this secondary calculation. However, if the lattice vectors change significantly during this second geometry optimization, the same problem might repeat itself. `PulayTask` tries to deal with this problem automatically by checking a user-specified condition versus a certain tolerance, and running the geometry optimization again in case this condition is not met. If the condition is met, the workflow continues as it was initially designed[5]. If not, another geometry optimization is performed based on the final structure.

3.2.5 `ConfigurationTask`/`EnergyConfTask`

These two tasks make up the complete configuration workflow 3.3.2, which can be used to calculate the energy of a large set of configurations for any system with degrees of configurational freedom in the placement of atoms on specified sites (e.g. lithium configurations of a charged cathode). `ConfigurationTask` generates the requested configurations based on the algorithm of Hart et al. [8] for the provided structure and a number of user-specified settings:

- `substitution_sites (list)`: Sites of the structure which should be considered for substitution.
- `element_list (list)`: Which elements are substituted into the structure to generate the configurations.
- `size (list)` : A list of allowed unit cell sizes for generating the configurations. Note that the algorithm is able to generate more than just supercells. Any unit cell that can reconstruct the structure with a number of sites equal to a size in the list times the number of sites in the original structure is considered.
- `concentration_restrictions (dict)`: Allows the user to specify restrictions on the *fractional* concentrations of the elements.

Once the configurations have been generated, `EnergyConfTask` sets up a two-step workflow for each configuration, consisting of a geometry optimization (`OptimizeFW`) and a static calculation to obtain a more precise value for the energy using the tetrahedron method (`StaticFW`).

3.2.6 Other

Here I give a brief overview of some smaller `Firetask`s which are also part of the workflows described in Section 3.3.

---

[5] In the context of the Fireworks lingo, `PulayTask` returns an `FWAction` that adds the next geometry optimization as an `addition`.

- `IncreaseNumberOfBands`: For some calculations, such as a calculation of the dielectric tensor, it is important to have a sufficient amount of empty bands. Often, this is done by setting the total number of bands to a multiple of the number occupied bands from a previous calculation. This `Firetask` automates this step, based on a user-specified multiplier.
- `VaspTask`: Similar to `CustodianTask`, this `Firetask` runs VASP in the directory of the `Firework`. However, this is just a bare VASP run without any error recovery based on a `Custodian`. In case the user does not specify any `ErrorHandlers` when initializing a `Firework` that runs VASP, this class is used instead of `CustodianTask`.
- `NWChemTask`: `Firetask` that runs an NWChem calculation in the specified directory. This is exclusively used for the calculation of the cation energy landscapes around polyborate anions, see Sections 3.3.5 and 5.4.2.
- `AddFinalGeometryToSpec`: This small but handy `Firework` extracts the final geometry from the current directory and adds it to the `_fw_spec`, so it can be used by both other `Firetask`s in the `Firework`, as well as future `Firework`s.
- `ScriptTask`: This is a standard task included in the Fireworks package which runs a user-defined command line script.
- `PyTask`: Versatile task included in Fireworks package that can run any Python method. Note that the method is defined as the string which would be used to import that method in a Python module, e.g. `pybat.workflow.workflows.get_wfs_noneq_dimers`. In case a `PyTask` is used in my workflows, the flowchart will detail the Python method used as well as a short explanation.

3.3 Workflows

In this section I present an overview of the workflows used to calculate the results presented in the rest of my thesis. As mentioned previously, all of these workflows consist of several `Firework`s, each of which represent one VASP calculation. As these `Firework`s are largely just a collection of `Firetask`s, I won't describe them separately. It should be clear from the description of the workflow what each `Firework` does.

3.3.1 Optical properties

Section 2.3 describes the theoretical framework for calculating the electromagnetic response of a material. Here I explain the practical steps you have to take to calculate the frequency dependent dielectric tensor for any given material, which have been employed to calculate the optical properties used in Chapter 4. As the number of steps required here is rather limited, it is an ideal problem to demonstrate a simple workflow.

Figure 3.3 shows the workflow applied to each structure. As the unit is often constructed by replacing elements in a template of the unit cell with the required spacegroup, the first step is to optimize the geometry of the structure[6] in the `OptimizeFW`. Here we allow for a full optimization of the unit cell (`ISIF` $= 3$), and as such the lattice vectors can change. In case the energy of the

---

[6] Note that in case a hybrid functional is used, it is efficient to first optimize the geometry using a less computationally demanding functional, especially when we generate structures that might have very different lattice parameters than the final structure.

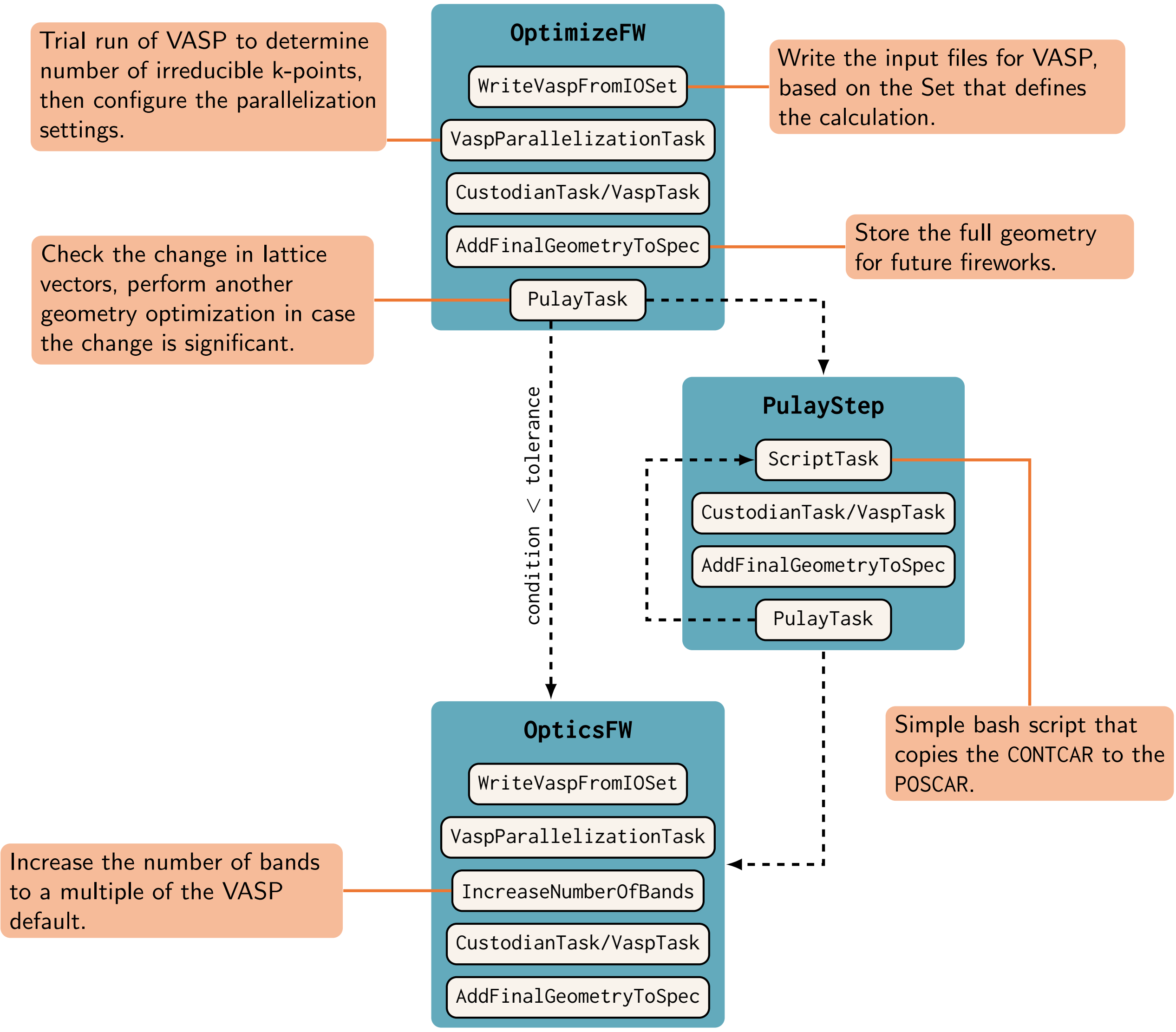


Figure 3.3: Flowchart for the optics workflow.

initial structure and final structure differs more than 1 meV per atom, the `PulayTask` copies the final structure `CONTCAR` to the `POSCAR` and performs another geometry optimization in a `PulayStep`.

Once the geometry is obtained, the workflow continues by calculating the dielectric tensor in the `OpticsFW`. This requires an increase in the number of bands, as to have enough unoccupied bands to calculate the dielectric tensor for sufficiently high energy transitions. Besides this, the `OpticsFW` also configures some default settings for dielectric function calculations, e.g. `LOPTICS`=True, a more dense energy grid (`NEDOS`=2000) and a more strict electronic convergence criterion (`EDIFF`=10^{-6}). Finally, the `Firework` also sets the `CSHIFT` parameter to 0.01, which should be sufficiently low to make sure that VASP does not overwrite the original imaginary part of the dielectric tensor with the one obtained from the Kramers-Kronig relation (See Appendix A.4). This introduces a broadening to the imaginary part of the dielectric tensor, which is passed to the absorption coefficient through Eq. (4.3). In effect, this reduces the band gap of the material, which has a large influence on the calculated efficiency described in Section 4.3.1.

3.3.2 Configurations

For various research questions, it is important to be able to investigate a whole range of configurations for a specific property. For example, when investigating the structural stability of a charged battery cathode, it is important to consider the most stable lithium configuration (Sec. 5.2.1). Figure 3.4 shows the configuration workflow, which at first only consists of a single `Firework` that generates all configurations for a specified set of restrictions using the algorithm of Hart et al. [8], as implemented in `ConfigurationTask`. The resulting list of symmetrically non-equivalent configurations is passed to the `EnergyConfTask`, which sets up a `Workflow` that optimizes the geometry of each configuration and subsequently determines the total energy based on a static calculation.

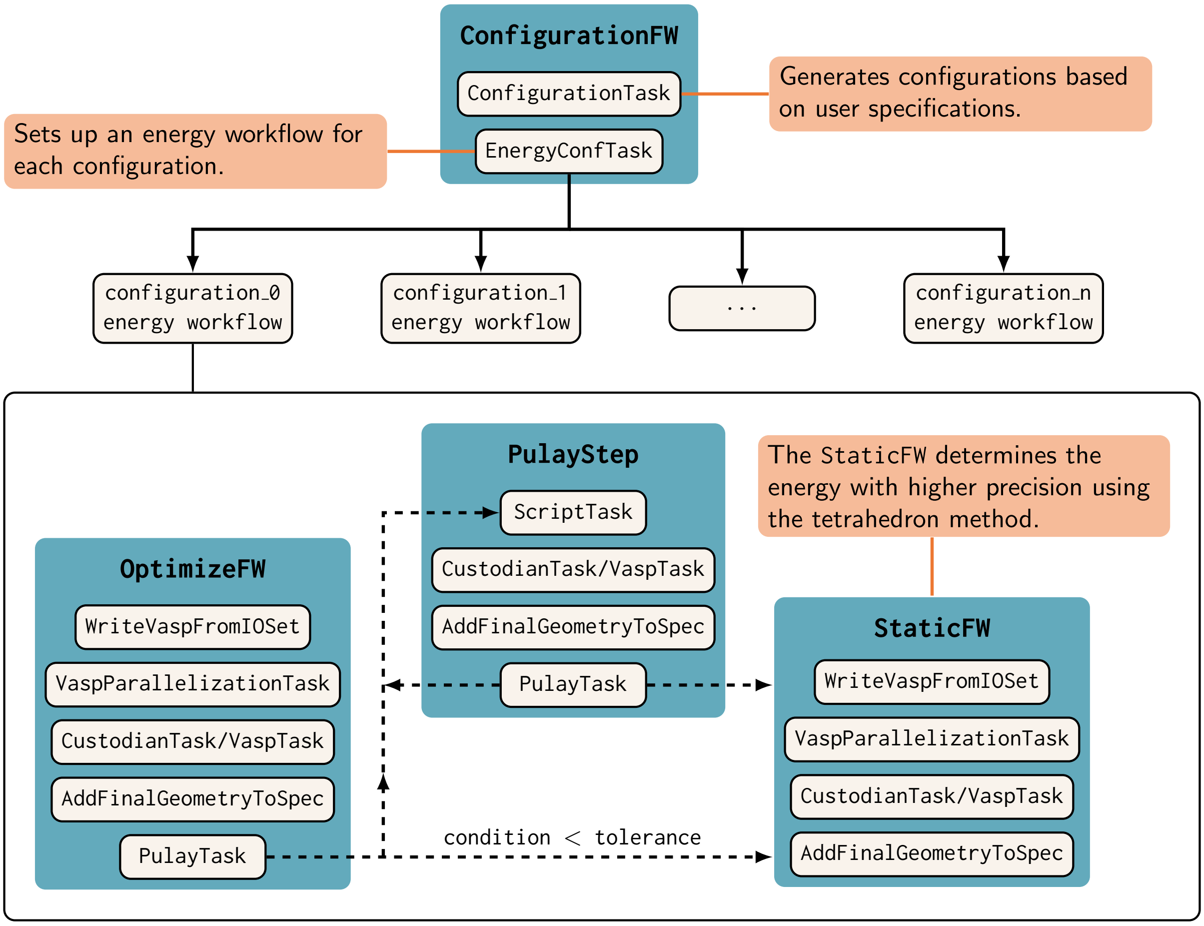


Figure 3.4: Flowchart for the configurations workflow.

3.3.3 Kinetic barriers - dimer workflow

As explained in Section 2.4, investigating a transition from one state of a system to another requires the calculation of both reaction energy and the kinetic barrier. After optimizing the geometry of two different states of a material, we can compare their energies in order to get an idea of the thermodynamic driving force for the transition. If such a reaction is favorable,

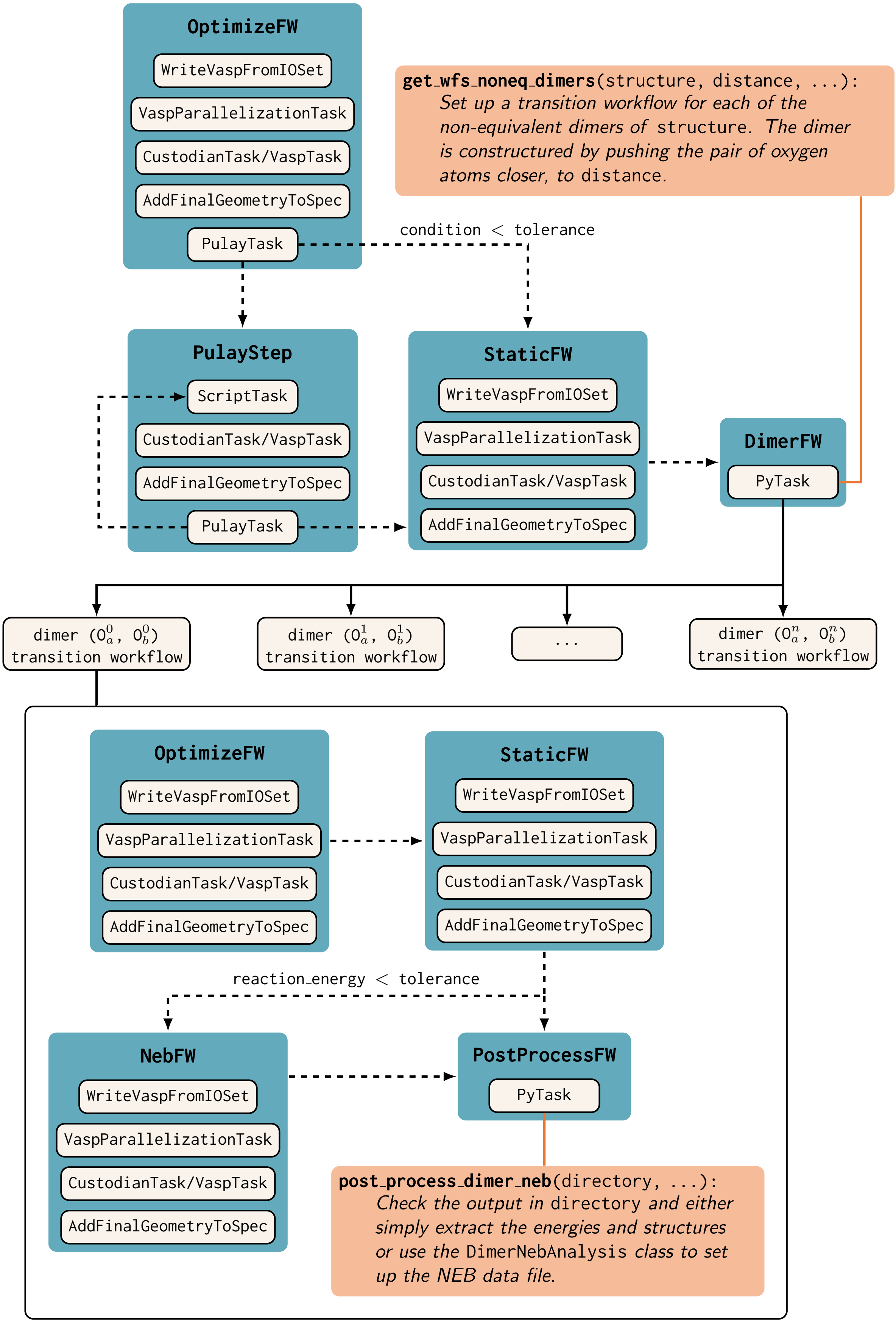


Figure 3.5: Flowchart for the dimer workflow.

i.e. its reaction energy is found to be negative, we still have to make sure that the transition is kinetically feasible. One method for doing this is calculating the kinetic barrier directly using the nudged elastic band method.

The investigation of the stability of the oxygen framework presented in Section 5.2.3 requires the thermodynamics and kinetics of oxygen dimer formation. Figure 3.5 shows the workflow that executes the necessary steps. First, the geometry of the provided structure is fully optimized (`ISIF`=3), and the energy is calculated more precisely using a static calculation. The final structure and energy is passed on to a small `Firework` that sets up a transition workflow for each non-equivalent dimer. This transition work flow consists of a geometry optimization of the dimer structure, where only the atomic positions are allowed to optimize (`ISIF`=2. Once again a static calculation is performed to determine the energy, which is then compared to the energy of the initial, unperturbed structure in order to obtain the reaction energy. If this reaction energy is smaller than a certain tolerance[7], the kinetic barrier is calculated using a nudged elastic band calculation.

3.3.4 Surface properties

The calculation of the secondary electron emission in the quantification of target surfaces (QUOTAS) project (Chapter 6) requires the DOS of the surface states, as well as the surface work function. However, because of the three dimensional periodic boundary conditions in VASP, simulating an isolated surface is not possible. The usual approach for dealing with this issue is by setting up a so-called slab geometry, where we introduce a layer of vacuum to the unit cell. In effect, this means that we have *two* opposite surfaces in our unit cell, hence the term slab is used instead of surface. Moreover, the use of periodic boundary conditions means that we have an infinite amount of slabs, but in case the vacuum is chosen to be sufficiently thick, the interaction between the surfaces of the slabs is negligible, and we can reasonably simulate a surface for the calculation or surface energies, work functions and other properties.

Let's describe the process of constructing such a slab unit cell in more detail. Usually, the surface is defined based on its Miller indices in the conventional unit cell of the bulk structure. However, for many structures, there are several possible surfaces corresponding to a set of Miller indices, depending on where we apply the cleavage plane in the structure. This leads to several possible terminations for each set of Miller indices. It is good practise to consider all terminations for each surface and calculate the surface energy to determine the most stable surface, and continue further calculations with this termination. Note that this approach does not consider the possibility of surface reconstruction. This is a difficult problem to solve, as probing all possible surface reconstructions is an insurmountable task. Surface reconstruction is especially common for polar surfaces, where the ionic nature of the bonds in the structure results in a dipole moment across the surface. Tasker [9] distinguished three types of ionic surfaces:

- Type I: Individual layers are charge neutral, and hence there is no surface dipole.
- Type II: Individual layers are charged, but groups of layers can be chosen that are charge neutral and non-polar.
- Type III: Groups of layers have a net dipole.

For the work presented in this thesis, I have only performed calculations on elemental surfaces, which are always of Tasker Type I and hence non-polar.

---

[7] Note that the reaction energy is defined as the final energy minus the initial energy. Hence, a transition with a negative reaction energy is thermodynamically favorable, i.e. setting the `tolerance` to zero only calculates the kinetic barrier for thermodynamically favorable dimers.

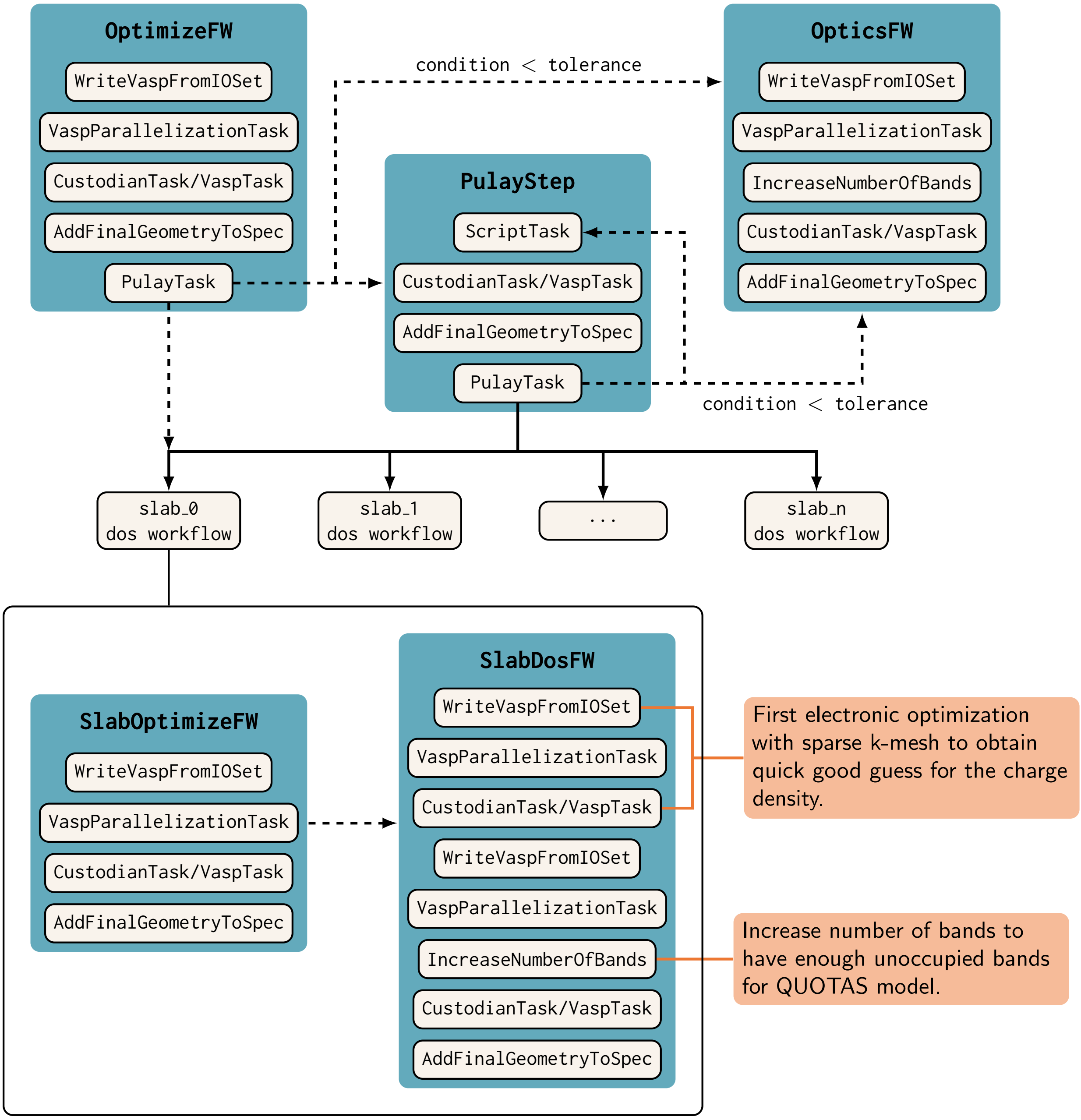


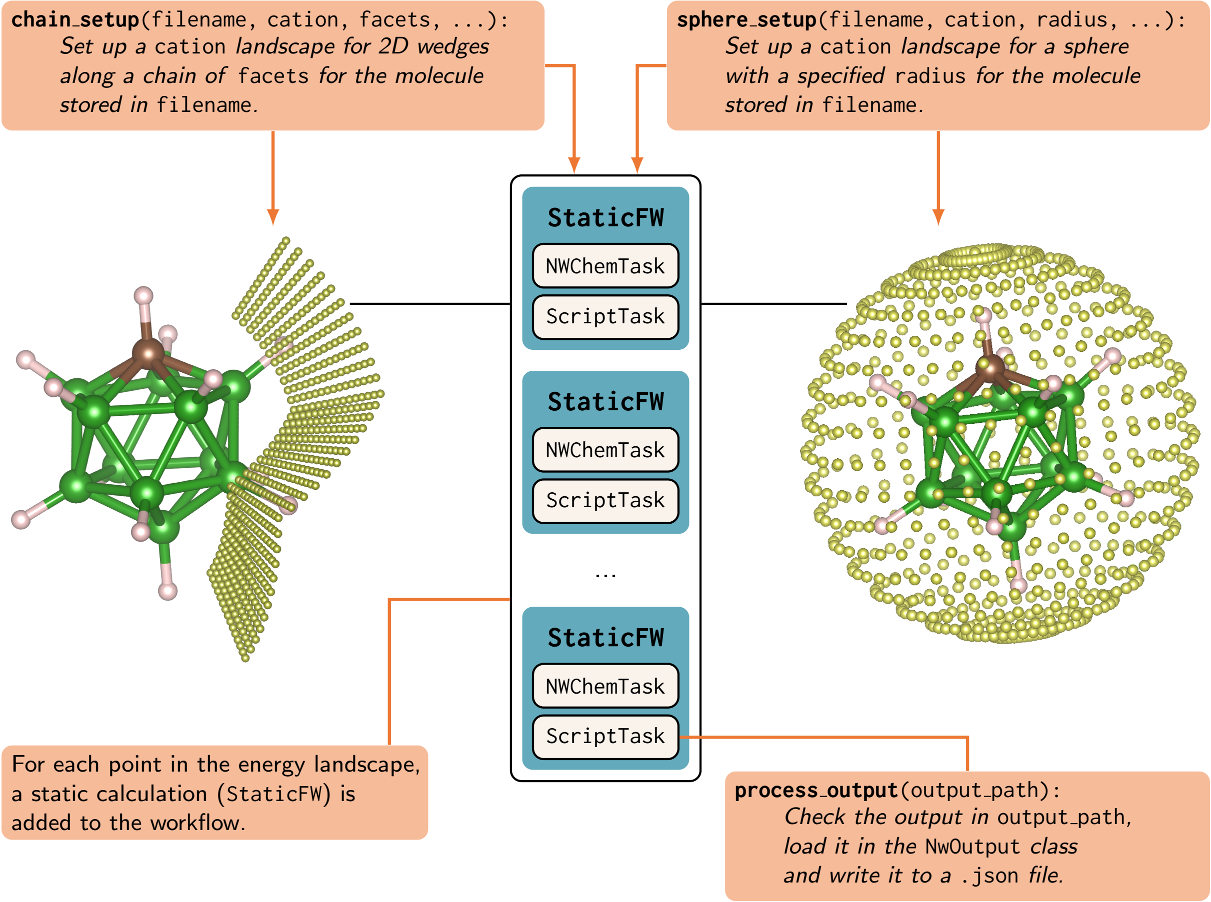
Figure 3.6: Flowchart for the QUOTAS workflow.

Once the surface and its termination have been decided, there are still several elements to consider before finalizing the unit cell for the calculations. The first thing to specify is the thickness of the slab, i.e. the amount of atomic layers, which should be large enough that the center of the slab behaves like the bulk of the material under investigation. Second, the thickness of the vacuum layer, i.e. the distance between two slabs in the periodic boundary conditions, must be set by suitably adjusting the unit cell and atom coordinates. In order to make sure both are sufficiently thick, convergence of the property of interest versus both slab and vacuum thickness should be checked.

The model for secondary electron emission described in Section 6.2 requires the density of states and the vacuum level. Starting from the bulk structure, the desired surface Miller indices need to be provided, as well as the `slab_settings`, e.g. the number of `free_layers` on each side of the slab. Once the geometry of the bulk structure has been fully optimized, the algorithm continues with calculating the dielectric tensor using a `OpticsFW`. Simultaneously, all terminations of the requested surfaces are generated, and a DOS workflow is initialized for each resulting slab. Each of these workflows starts with an optimization that *relaxes* the surface atoms, depending on how many free layers are specified. Finally, the workflow calculates the slab DOS in two steps: one run with a sparse k-point mesh in order to obtain a reasonable charge density, and one with a dense k-point mesh in order to obtain a precise DOS and local potential.

3.3.5 Landscapes

The energy landscapes presented in Section 5.4.2 require knowledge about the energy of the anion-cation system for a range of cation positions on a two dimensional surface. We represent the landscape as a grid of points or coordinates, for which we determine the energy of the system with a static DFT calculation. Although it is possible to simulate molecules in VASP – despite its use of three dimensional periodic boundary conditions – by considering a large unit cell, we choose to use NWChem [10]. This software package is more suitable for the study of molecules as it does not enforce the use of periodic boundary conditions.



chain_setup(filename, cation, facets, ...):
Set up a cation *landscape for 2D wedges along a chain of* facets *for the molecule stored in* filename.

sphere_setup(filename, cation, radius, ...):
Set up a cation *landscape for a sphere with a specified* radius *for the molecule stored in* filename.

For each point in the energy landscape, a static calculation (StaticFW) is added to the workflow.

process_output(output_path):
Check the output in output_path*, load it in the* NwOutput *class and write it to a* .json *file.*

Figure 3.7: Flowchart for the landscape workflow.

The landscape workflows, shown schematically in Fig. 3.7, rely on two setup scripts, `setup_chain` and `setup_sphere`, which initialize the `Landscape` according to the parameters specified by the user. For each coordinate in the landscape grid, a static calculation is added to the workflow, followed by a small `ScriptTask` that processes the output from the NWchem output file into a more succinct `.json` file. The data for each segment of the landscape is then mapped back onto the landscape using the `LandscapeAnalyzer` class, and subsequently tied together. Note that a single landscape can consist of hundreds or even thousands of grid points, which makes its calculation impractical without the use of a workflow.

3.4 Computational Details

Although all the workflows have defaults for computational settings such as the density of k-points, energy cutoff for the plane waves etc., most calculations will require the user to adjust these settings based on a set of convergence tests. These can be passed to the `Firework`s of the workflow using the `vasp_input_params` input argument. This is a Python `dictionary` whose key value pairs are passed to the `WriteVaspFromIOSet`'s input set as keyword arguments. Any input argument of the parent `DictSet` class can be adjusted using `vasp_input_params`, but here is a list of some commonly used ones:

- `user_incar_settings`: The most important of the keys, this allows the user to specify any INCAR tag. Any tag specified will override whatever configuration is generated by the input set, even when they are derived from the structure (e.g. `MAGMOM`).
- `user_kpoints_settings`: Change the density of the k-point mesh, or specify the number of k-points in each direction.
- `force_gamma`: Force a gamma centered mesh for the k-points.

For each of the results sections in this thesis, the most important computational details are provided in Appendix A.2, organised by chapter and in order of the presentation of the results. Although you are welcome to read them now, there is also a little silicon chip symbol next to each section header that contains computational results, that links to the corresponding section in the appendix.

Bibliography

[1] S. P. Ong et al. "Python Materials Genomics (pymatgen): A robust, open-source python library for materials analysis". *Comput. Mater. Sci.* 68 (2013), pp. 314–319. DOI: `10.1016/j.commatsci.2012.10.028`.

[2] A. H. Larsen et al. "The atomic simulation environment - A Python library for working with atoms". *J. Phys.: Condens. Matter* 29 (2017). DOI: `10.1088/1361-648X/aa680e`.

[3] A. Jain et al. "Fireworks: A dynamic workflow system designed for high-throughput applications". *Concurr. Comput.* 27 (2015), pp. 5037–5059. DOI: `10.1002/cpe.3505`.

[4] A. Jain et al. "Commentary: The Materials Project: A materials genome approach to accelerating materials innovation". *APL Materials* 1 (2013), p. 011002. DOI: `10.1063/1.4812323`.

[5] D. M. Wood and A Zunger. "A new method for diagonalising large matrices". *J. Phys. A: Math. Gen.* 18 (1985), pp. 1343–1359. DOI: `10.1088/0305-4470/18/9/018`.

[6] G. P. Francis and M. C. Payne. "Finite basis set corrections to total energy pseudopotential calculations". *J. Phys. Condens. Matter* 2 (1990), pp. 4395–4404. DOI: `10.1088/0953-8984/2/19/007`.

[7] F. Hachenberg. *Pulay Stress Visualization.* URL: `https://en.wikipedia.org/wiki/File:PulayStressVisualization.png` (visited on 12/14/2019).

[8] G. L. Hart et al. "Direct enumeration of alloy configurations for electronic structural properties". *Appl. Phys. Lett.* 87 (2005), pp. 1–3. DOI: `10.1063/1.2142091`.

[9] P. W. Tasker. "The stability of ionic crystal surfaces". *J. Phys. C: Solid State Phys.* 12 (1979), pp. 4977–4984. DOI: `10.1088/0022-3719/12/22/036`.

[10] M. Valiev et al. "NWChem: A comprehensive and scalable open-source solution for large scale molecular simulations". *Comput. Phys. Commun.* 181 (2010), pp. 1477–1489. DOI: `10.1016/j.cpc.2010.04.018`.

Chapter 4

Solar Cells

> *"We are star stuff harvesting sunlight."*
>
> Carl Sagan

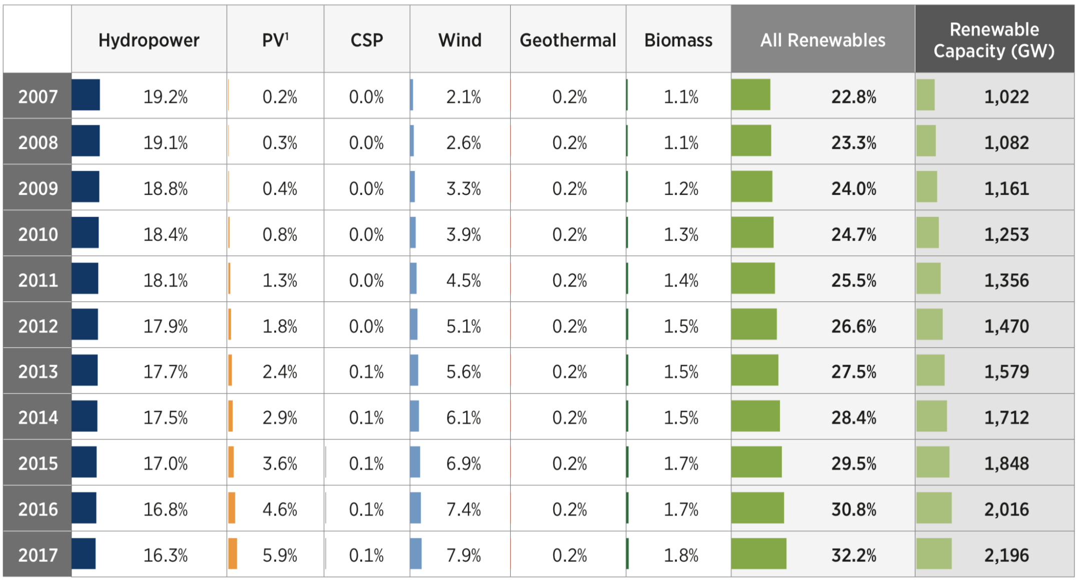
Solar cell technology is a rapidly developing field, largely driven by the discovery of cheaper and more efficient materials. However, the conventional search for materials that can serve as absorber layers in photovoltaic devices is both expensive and time consuming. Modern high-throughput computational methods have the power to screen large amounts of materials relatively quickly, providing valuable information that allows experimental work to focus on promising compounds. In order to accurately screen materials, a proper selection metric is required. The spectroscopic limited maximum efficiency (SLME) attempts to improve upon the traditional Shockley-Queisser (SQ) limit by including the first-principles calculated absorption spectrum of the material in the determination of the efficiency. It also allows researchers to investigate the thickness dependence of the efficiency, which is particularly interesting for thin-film solar cell research.

In this chapter I give an overview of my work on solar cell absorber materials, which was largely performed during the first year of my PhD. After a brief introduction, the chapter starts by explaining the basic working principles of a solar cell in Section 4.2. Next, Section 4.3 continues by discussing the SLME, as well as its predecessor, the Shockley Queisser limit. In Section 4.4, this selection metric is applied to compare the potential of a range of ternary I-III-VI$_2$ compounds in both the CuAu-like and the chalcopyrite phase. Finally, some crucial aspects of our chosen selection metric are investigated in Section 4.5.

4.1 Introduction

Nearly all sources of energy found on earth are in some way derived from the sun. Both animals and plants depend on solar energy to produce the heat and sustenance they need to survive. Fossil fuels are nothing more than long buried organisms, exposed to millions of years of heat and pressure in the earth's crust. Wind energy would not exist without the air currents that are a product of solar heated air and the rotation of the earth. Solar cells are photovoltaic (PV) devices that try to directly convert the sun's light into electricity, by absorbing the incoming photons and using the resulting energy to create a current of moving electrons. They have the advantage of being a renewable source of clean energy, whose application can be much more distributed than more conventional sources of electricity [1].

The first practical photovoltaic devices were constructed in the 1950s. Over the course of the next decades, solar cell technology was mainly developed by the space industry, which required a reliable source of energy for its satellite applications. In the 1980s, solar cells received an increased amount of attention due to the oil crisis and a growing demand for power supply in remote areas that are not connected to the electricity grid. More recently, the threat of global warming has expanded the interest in sustainable energy sources. Advances in technology have increased the efficiency and longevity of solar cells, while reducing the costs of production and maintenance. Moreover, governments around the world have started initiatives to increase the percentage of the world's renewable energy supply. Figure 4.1 shows the results of these efforts. We can see that the PV market has increased significantly over the past decade, now contributing 5.9% of the total global (grid-connected) electrical energy production. An important caveat here is that the production and recycling of solar modules also requires energy, and has an environmental impact. With regards to the production, Bhandari et al. [2] report an energy payback time (EPBT) of 1 to 4.1 years depending on the type of module. Considering a lifetime of 25 years for the average solar cell, efficient end-of-life (EOL) strategies will have to developed to ensure the sustainability of photovoltaic devices [3].



	Hydropower	PV[1]	CSP	Wind	Geothermal	Biomass	All Renewables	Renewable Capacity (GW)
2007	19.2%	0.2%	0.0%	2.1%	0.2%	1.1%	22.8%	1,022
2008	19.1%	0.3%	0.0%	2.6%	0.2%	1.1%	23.3%	1,082
2009	18.8%	0.4%	0.0%	3.3%	0.2%	1.2%	24.0%	1,161
2010	18.4%	0.8%	0.0%	3.9%	0.2%	1.3%	24.7%	1,253
2011	18.1%	1.3%	0.0%	4.5%	0.2%	1.4%	25.5%	1,356
2012	17.9%	1.8%	0.0%	5.1%	0.2%	1.5%	26.6%	1,470
2013	17.7%	2.4%	0.1%	5.6%	0.2%	1.5%	27.5%	1,579
2014	17.5%	2.9%	0.1%	6.1%	0.2%	1.5%	28.4%	1,712
2015	17.0%	3.6%	0.1%	6.9%	0.2%	1.7%	29.5%	1,848
2016	16.8%	4.6%	0.1%	7.4%	0.2%	1.7%	30.8%	2,016
2017	16.3%	5.9%	0.1%	7.9%	0.2%	1.8%	32.2%	2,196

Figure 4.1: Renewable electricity as a percentage of the total installed global electricity capacity. Note that the PV numbers here are for grid-connected production only, represented by the index 1. CSP stands for concentrated solar power. Taken from [4].

Although the growth of the PV market is promising, continued efforts must be made to increase the share of renewable energy in the global energy production. In order to make solar cells more economically competitive with conventional sources of electricity, new materials have to be found that either increase the efficiency or lower the cost of PV devices. In this chapter, we investigate a selection metric that determines the potential of a material for solar cell applications.

4.2 Solar Cell Basics

This section discusses the basics of solar cells, starting from a presentation of the solar spectrum and the absorption coefficient. It continues by explaining recombination effects, an important limiting factor for solar cell efficiency. The P-N junction, as well as the relevant equations, are the next topic of this section. Finally, we present the working principles of the solar cell, as well as a discussion of its I-V characteristic, which is essential for understanding the selection metrics introduced in the Section 4.3.

4.2.1 Solar Spectrum

Before discussing the operation of solar cells, we briefly take a look at the solar spectrum itself (Fig. 4.3). In good approximation, the sun emits electromagnetic radiation as a black body, a perfectly absorbing mass whose emission spectrum is determined by Planck's radiation law [5]. Due to the influence of the sun's atmosphere, the spectral distribution that reaches the outside of earth's atmosphere differs from the ideal black-body spectrum. Before reaching the planetary surface, the incoming radiation intensity is further attenuated by various scattering and absorption effects [6]. The degree of attenuation depends on the angle at which the light enters the atmosphere. The solar spectra are usually classified by their *Air Mass* (AM) [7]:

$$\text{Air Mass} = \frac{1}{\cos\theta_z}, \tag{4.1}$$

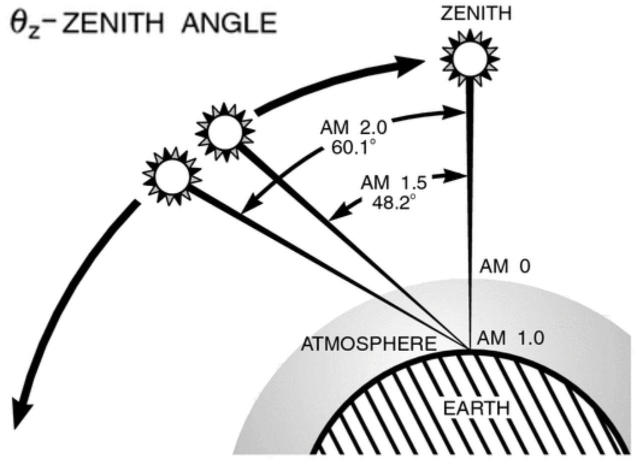


Figure 4.2: Relation between zenith and Air Mass [8].

where θ_z is the zenith angle of the incoming light (Fig. 4.2). The use of these standard spectra allows the performance of devices to be judged fairly, by exposing them to the same agreed-upon spectrum. In this work, we use the AM 1.5G irradiance spectrum for all of our calculations, which is a good representation of the illumination conditions on a tilted flat PV array. The 'G' stands for global tilt and refers to the angle between the normal of the surface and the direction of the incoming sunlight.

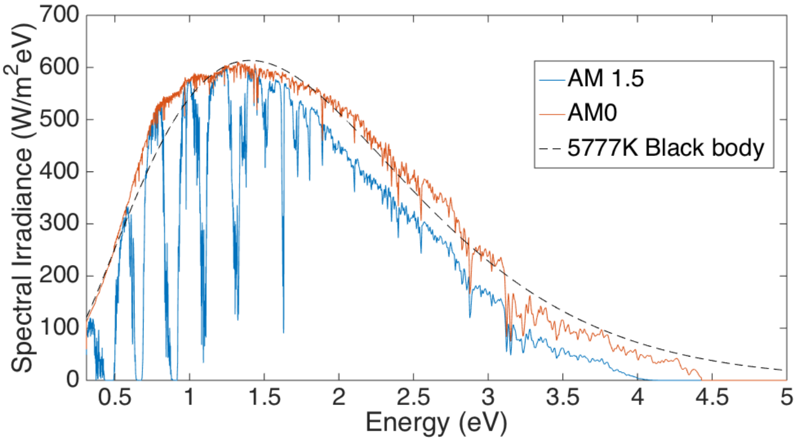


Figure 4.3: Solar Spectra for the extraterrestrial AM 0 (orange) and AM 1.5 (blue) measurements. A black-body spectrum is added as a reference point. Data taken from [9].

4.2.2 Absorption

When light passes through a semiconductor, a fraction of the photons is absorbed by the material, which causes the system to undergo a transition into an excited state. An example of such a higher energy state is an electron that is excited from the valence band (VB) to the conduction band (CB). When an electron undergoes this transition, it leaves behind a space in the valence band for other electrons to move into. Rather than keeping track of the valence band electrons, the movement of the empty space is usually represented by a positively charged particle, referred to as a *hole*. In this framework, the absorption of light is considered as the creation of an electron-hole pair, called an *exciton*, through the annihilation of an incoming photon. If the electron-hole pair can be separated, they become free charge carriers that can move through the semiconductor.

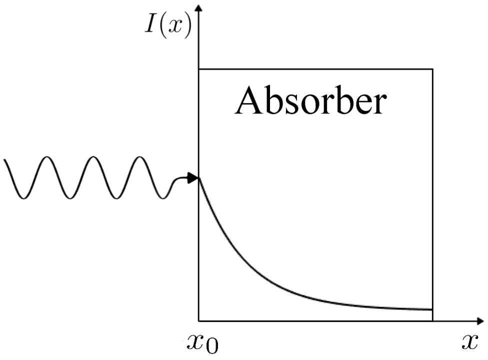


Figure 4.4: Attenuation of the incoming light.

The absorption process is essential for the generation of carriers that produce the current in photovoltaic devices. The strength of the absorption of a material for photons of energy E is described by the *absorption coefficient* $\alpha(E)$, which determines the attenuation of the incoming monochromatic light [7]:

$$I(x) = I(x_0)e^{-\alpha(E)(x-x_0)}, \tag{4.2}$$

where $I(x_0)$ is the intensity upon entering the semiconductor (Fig. 4.4). The absorption coefficient is related to the extinction coefficient $\hat{k}$, which is defined as the imaginary part of the complex index of refraction $\hat{n}_c(E) = n(E) + ik(E)$:

$$\alpha(E) = \frac{4\pi E \cdot k(E)}{hc}, \tag{4.3}$$

where h is Planck's constant and c is the speed of light. Both the real and imaginary parts of the index of refraction can be calculated from the dielectric tensor, see Section 2.3.

An important distinction to make here is the difference between *direct* and *indirect* transitions (Fig. 4.5). When a photon is absorbed to excite an electron to a higher energy level, there must

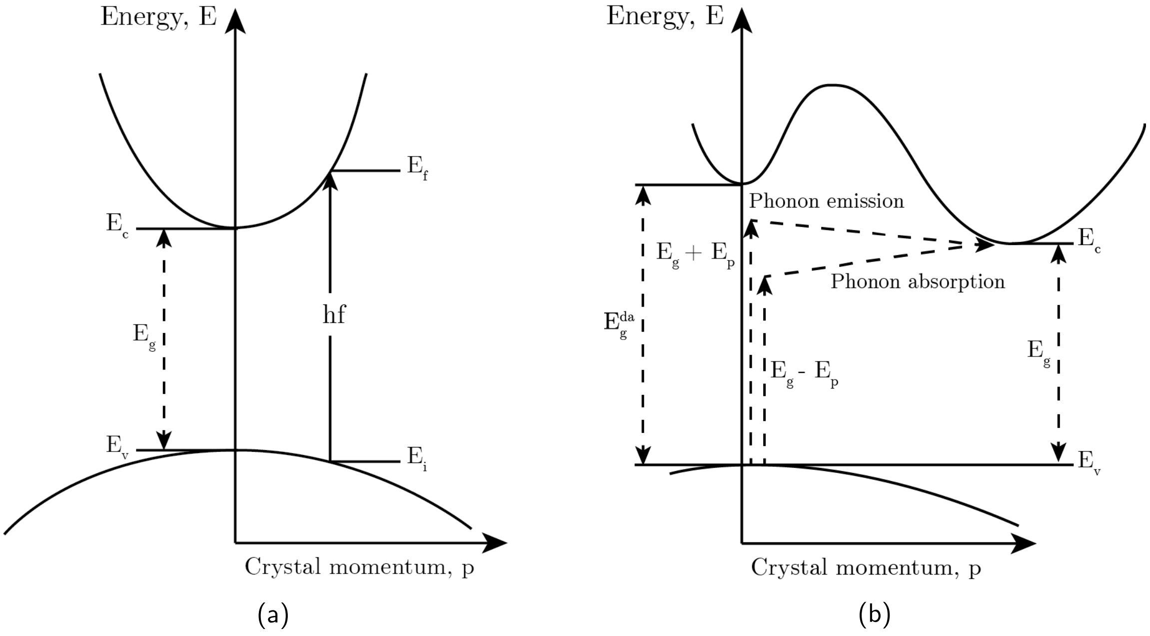


Figure 4.5: Direct (a) and indirect (b) absorption. Adapted from [7].

be a conservation of energy and momentum. For direct transitions, the electron absorbs both the photon energy and momentum, the latter of which is very small. However, for indirect band gap semiconductors, the minimum energy of the conduction band occurs at a very different value of the crystal momentum than the maximum energy of the valence band. In this case, absorption of photons with energies close to the band gap involves the participation of phonons – quasi particles describing the mechanical vibrations in the crystal – in order to provide the necessary change in momentum. For these materials, it is useful to make a distinction between the *fundamental* band gap E_g and the *direct allowed* band gap E_g^{da}, the latter of which is defined as the minimal difference in energy of the valence and conduction band at the same crystal momentum.

4.2.3 Recombination

The electrons that are excited to the conduction band in a semiconductor are in a meta stable state and eventually return back to the valence band, effectively also removing the holes they left after their transition to the conduction band. This process is called *recombination.* The lifetime of an electron-hole pair depends on the rate of the different recombination mechanisms present in the semiconductor material. In this work, we make a distinction between three types of recombination:

Radiative Recombination is an interband process that can be considered the reverse of the absorption process. An excited electron in the conduction band recombines with a hole in the valence band, producing a photon which is emitted by the diode (Fig. 4.6). The energy of the photon is given by $h\nu = E_c - E_v = E_g$. This recombination mechanism is prevalent in direct band gap absorbers such as GaAs, a material used for the design of light-emitting diodes [7].

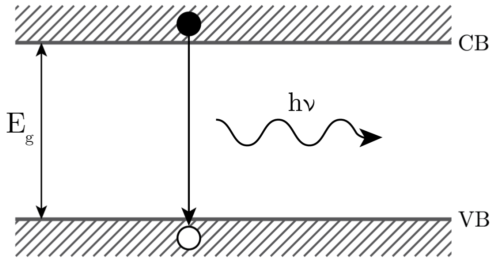


Figure 4.6: Radiative Recombination.

Auger Recombination [10] is a three particle mechanism, where the energy produced by the electron-hole recombination is transferred to another electron (either in the conduction or valence band), instead of emitting a photon. The second electron then returns back to its original energy via thermal relaxation (Fig. 4.7). Since no light is emitted in this process, it is often referred to as *non-radiative* recombination. This type of recombination is dominant in indirect band gap materials, of which the most important example is silicon [11].

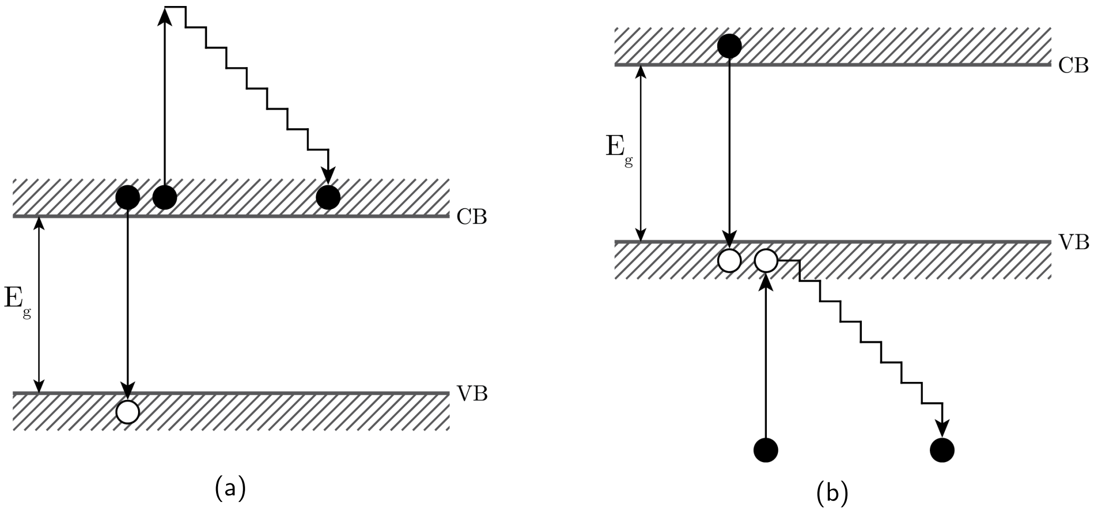


Figure 4.7: Auger Recombination. The recombination energy is either passed to an electron in the conduction band (a) or the valence band (b).

Shockley-Read-Hall Recombination (SRH) or trap-assisted recombination [12][13]. This mechanism uses energy levels in the band gap, usually produced by defects or impurities, in order to relax from the conduction band through a two step process. The electron first undergoes a transition to the energy level created by the defect, after which it recombines with a hole in the valence band (Fig. 4.8). Since SRH recombination uses energy levels in the band gap which are produced by defects, this mechanism is more important in materials which are heavily doped. It is also possible to demonstrate that defect levels situated near the middle of the band gap are the most effective for the SRH recombination, providing larger recombination rates [7].

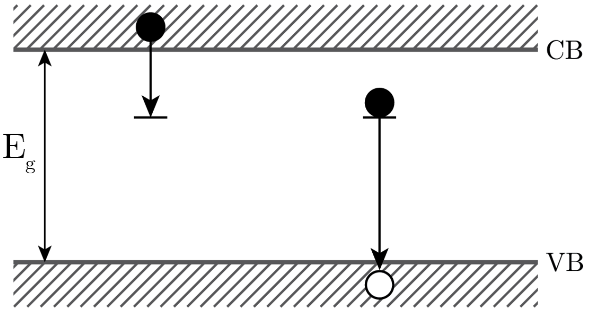


Figure 4.8: Shockley-Read-Hall Recombination.

4.2.4 PN - Junction

When a semiconductor is doped with impurity atoms that have more valence electrons, there are more electrons in the conduction band that can contribute to the conductivity of the material. This is called a *n-type* semiconductor. If atoms are introduced that have less valence electrons, there is an increased amount of holes in the valence band, also improving the conductivity in what is referred to as a *p-type* semiconductor. By joining a n-type and p-type semiconductor, we form what is known as a *P-N junction diode* [14] (Fig. 4.9).

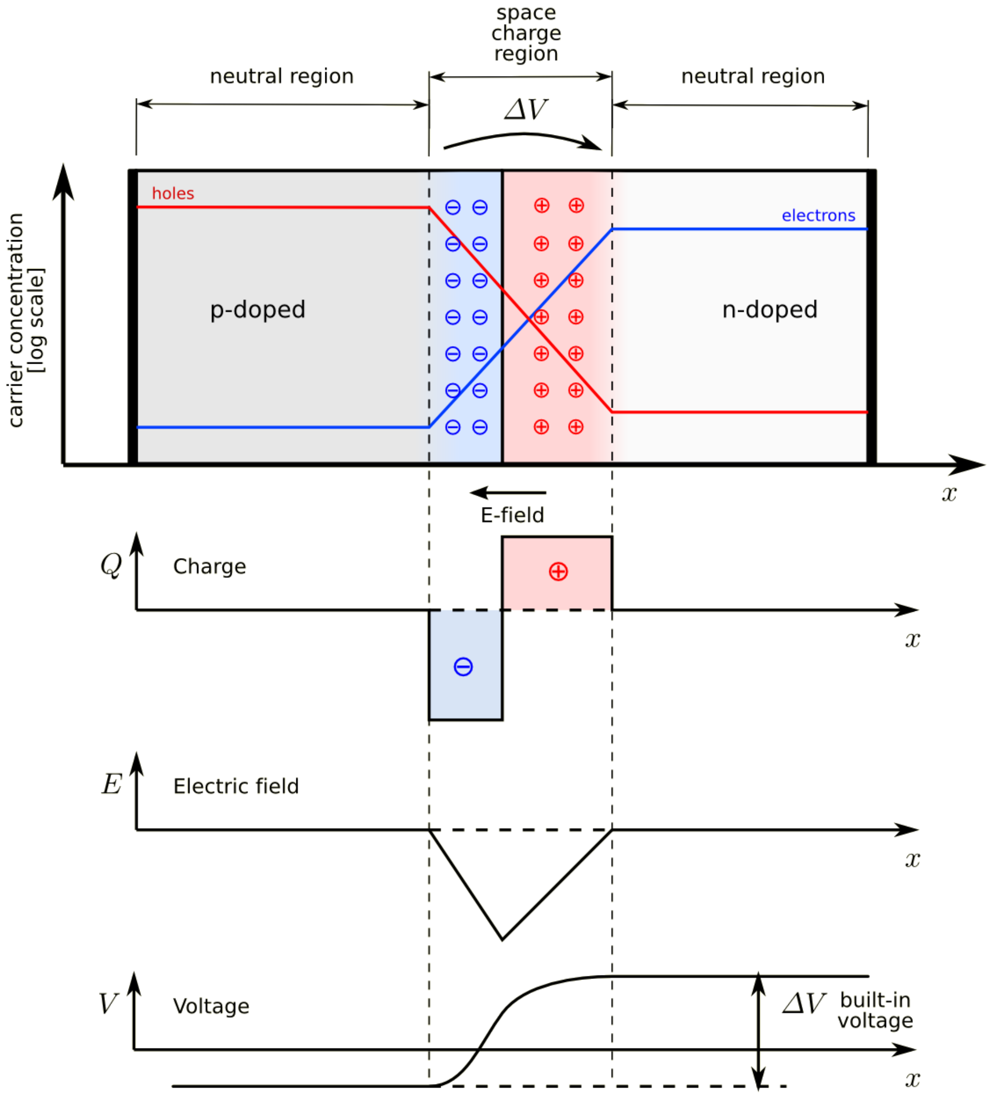


Figure 4.9: The P-N Junction. [15]

An important effect caused by putting these two types of materials together is the creation a *depletion region*, which is a result of the electrons and holes looking for an equilibrium by moving to the other side of the P-N junction. When the electrons move from the n-type to the p-type region, they leave behind positively charged ion cores. Similarly, holes moving across the junction create a negative surplus charge in the p-type semiconductor. This charge imbalance creates an electric field at the center region of the diode, which eventually puts a stop to the migration of electrons and holes. The result is a potential difference across the P-N junction, which is required to separate the charge carriers generated by the absorption of incoming light.

By applying a voltage across the P-N junction, we can influence the strength of the electric field in the depletion region. Under *forward bias*, we decrease the voltage difference over the diode, making it easier for the charge carriers to move across the junction. This has the potential of increasing the power output, but also raises the possibility that the electrons and holes can recombine. In the *reverse* bias case, the magnitude of the electric field across the junction is increased, and the charge carriers are more or less confined to their respective regions.

4.2.5 Ideal Diode Law

In order to properly model the functioning of a solar cell, we need to know the I-V characteristic of the P-N junction. If we write down the electron and hole densities as $n(x)$ and $p(x)$, along with their respective current densities J, mobilities μ and diffusivity constants D, the dynamics of the electrons and holes in the different regions of the P-N diode can be characterized by a set of basic equations [14]:

1. **Poisson's Equation:**

$$\frac{\partial E_x}{\partial x} = \frac{\rho}{\epsilon}, \tag{4.4}$$

2. **Transport Equations:**

$$J_n = e\mu_n n(x)E_x + eD_n\frac{\partial n}{\partial x}, \qquad J_p = e\mu_p p(x)E_x - eD_p\frac{\partial p}{\partial x}, \tag{4.5}$$

3. **Continuity Equations:**

$$\frac{\partial n}{\partial t} = \frac{1}{e}\frac{\partial J_n}{\partial x} - (U - G), \qquad \frac{\partial p}{\partial t} = -\frac{1}{e}\frac{\partial J_p}{\partial x} - (U - G), \tag{4.6}$$

where E_x is the electric field, ρ is the charge density, ϵ is the material permittivity and $e = 1.602 \cdot 10^{-19}$ C is the standard electronic charge. The generation rate G and recombination rate U describe the creation and annihilation of electron-hole pairs, and are determined by the absorption and recombination processes described in Sections 4.2.2 and 4.2.3.

These equations can be readily solved using numerical approaches. However, by applying a few approximations, it is possible to derive a general relation for the I-V characteristic of a P-N junction under dark conditions, known as the *Ideal Diode Law* [14]:

$$I = I_0(e^{\frac{eV}{k_B T}} - 1), \tag{4.7}$$

with k_B Boltzmann's constant and T the temperature of the diode. The current I_0 in Eq. 4.7 is called the *reverse saturation* current, which is a measure of the recombination in the device. In light of its connection with recombination effects, I_0 is also referred to as the recombination current [16].

4.2.6 Working Principles of a Solar Cell

We now turn to a basic discussion of the working principles of photovoltaic devices [17]. Figure 4.10 shows the design of a conventional solar cell. The top and base layer usually consist of a n-type and p-type doped material, not necessarily derived from the same compound. The layers form a P-N junction, which is connected on both sides to electrodes that are responsible for extracting the charge carriers from the photovoltaic device. In order to prevent losing too much of the incoming light to reflection, an *anti-reflective* (AR) coating [18] is applied to the top layer. Many solar cells also use a reflective back surface (not shown in figure) to increase the path length of the incoming light through the absorber layer.

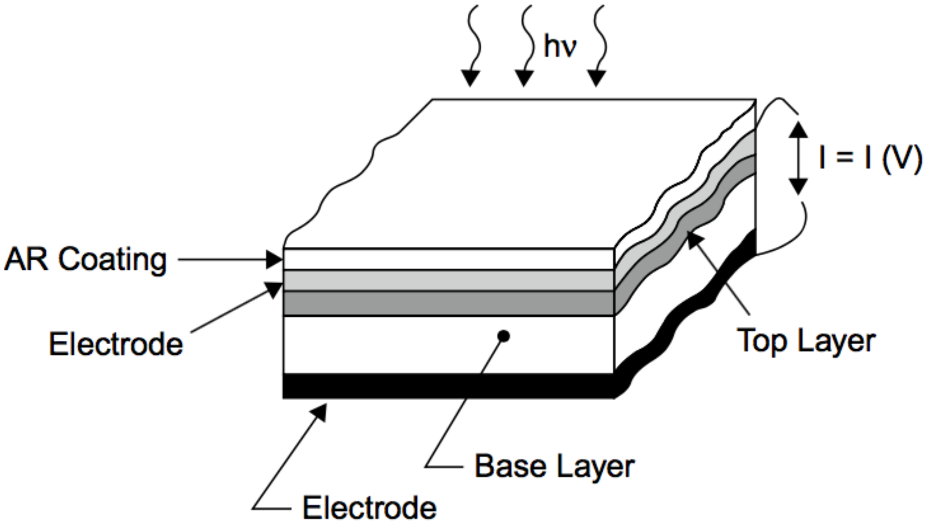


Figure 4.10: Cross section of a typical solar cell. Taken from [17].

When the absorber layer captures an incoming photon, the material is brought into a higher energy state, usually by the creation of an exciton or free electron-hole pair. In the case an exciton is produced, it must first be dissociated into an electron and a hole. Once the charge carriers are free to move, they go to their respective electrode interface. During this stage it is crucial that the electron does not recombine with another hole before it is extracted from the device. Once the electron reaches the cathode, it makes its way to an external load, where it transfers its energy before moving to the anode. Finally, the electron reaches the bottom layer and recombines with a hole.

The process of generating a current from a solar cell can be summarized in five steps, which are visualized by drawing an energy band diagram for the absorption process in Fig. 4.11:

I. Transition into an excited state by the absorption of an incoming photon.
II. Conversion of the excited state into at least one electron-hole pair.
III. Transport of the charge carriers to their respective electrodes.
IV. Transfer of the electron's energy at an exterior load.
V. Recombination of the electron with the hole.

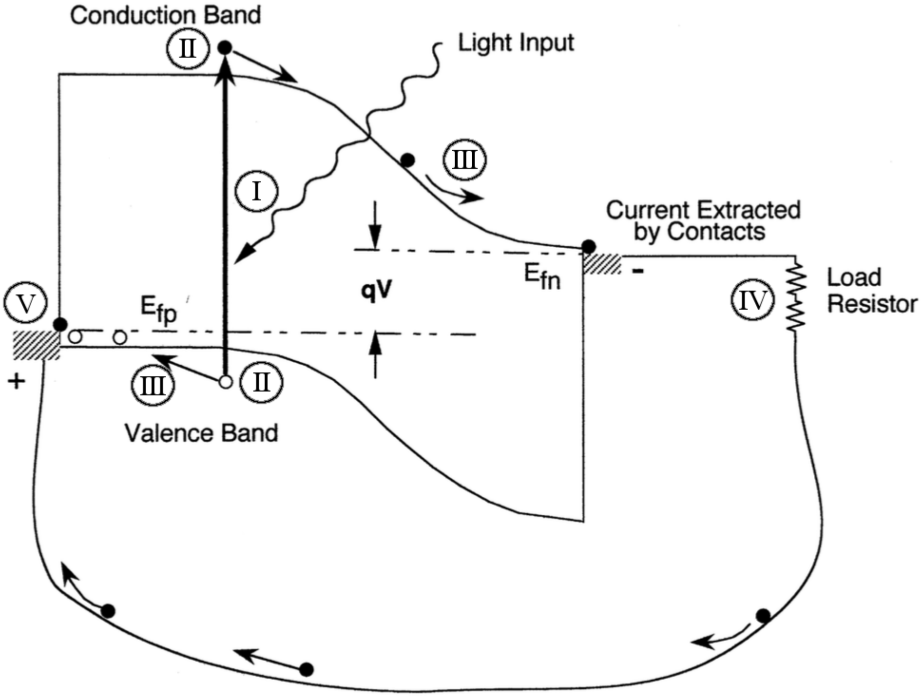


Figure 4.11: Step by step visualization of the different processes in a solar cell. Adapted from [19].

The I-V characteristic of the solar cell under dark conditions is given by the ideal diode law (Eq. (4.7)). When the solar cell is illuminated, the I-V equation becomes[1] [20]:

$$I = I_{sc} - I_0(e^{\frac{eV}{k_BT}} - 1), \tag{4.8}$$

where I_{sc} is the *short circuit* current. This current is a result of the generation of charge carriers due to absorption of incoming photons. In many cases, the currents in Eq. 4.8 are

---

[1] Note that in principle, the I-V curve is given by $I = I_0(e^{\frac{eV}{k_BT}} - 1) - I_{sc}$, but the convention is to invert the current axis.

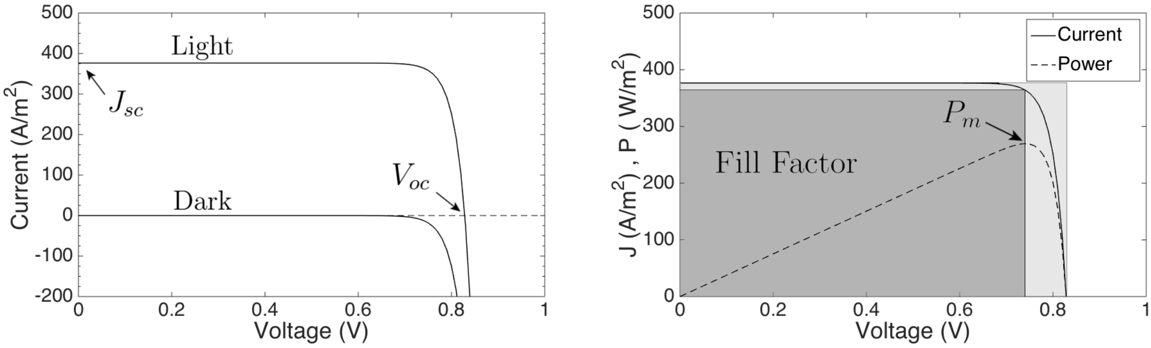
expressed as current densities[2] (J, J_0, J_{sc}), and I follow the same convention throughout this text. Figure 4.12a shows the I-V characteristic of a solar cell under illuminated and dark conditions, whereas Fig. 4.12b demonstrates the definition of the *Fill Factor* (FF) [17]:

$$FF = \frac{P_m}{V_{oc}J_{sc}}, \tag{4.9}$$

where $P_m = J_m V_m$ is the maximum power density and V_{oc} is the *open circuit* voltage, which is the voltage of the diode at $J = 0$:

$$J_{sc} - J_0(e^{\frac{eV_{oc}}{k_B T}} - 1) = 0 \quad \Leftrightarrow \quad V_{oc} = \frac{k_B T}{q} \ln(\frac{J_{sc}}{J_0} + 1). \tag{4.10}$$

The fill factor is a measure for how close a given characteristic is to obtaining the ideal power density $J_{sc}V_{oc}$, i.e. operating at the short-circuit current and open-circuit voltage.



(a) (b)

Figure 4.12: (a) I-V curve of a solar cell under dark and illuminated (light) conditions. (b) Power maximization and the corresponding fill factor, given by the dark area (P_m) divided by the light area ($V_{oc}J_{sc}$).

4.3 Selection Metric

Materials play a central role in the effort to produce cheaper and more efficient solar cells. The discovery of improved absorber materials has the potential to significantly increase the cost-effectiveness of photovoltaic devices, but experimental trial and error methods are often slow and expensive. Computational high throughput screening can offer a quick and relatively cheap approach for refining the selection of materials which warrant experimental investigation. In order to do so, however, a suitable selection metric is required that can determine the potential of a material to function as an absorber layer in a solar cell.

The *Spectroscopic Limited Maximum Efficiency* (SLME) is a calculable selection metric that tries to improve upon its predecessors by including the absorption spectrum, calculated from first-principles methods, in its estimation of the efficiency of materials as absorber layers. This section is dedicated to explaining the SLME metric, as well as its predecessor, the Shockley-Queisser (SQ) limit.

---

[2] Note that these current densities are not defined in the conventional way. Rather, they are considered as currents per surface area of the solar cell. This allows us to ignore the surface area of the solar cell in our discussion.

4.3.1 Solar Cell Efficiency

In order to compare the performance of solar cells, we need to have a gauge for their efficiency. A sensible way of defining the efficiency of a photovoltaic device is as the ratio of the maximum power density P_m produced by the solar cell, and the total incident power density from the solar spectrum P_{in} [17]:

$$\eta = \frac{P_m}{P_{in}}. \tag{4.11}$$

When determining the efficiency, it is important to use agreed upon conditions for the measurement and calculation of both power densities. For the solar spectrum, researchers normalize the AM1.5G spectrum to produce a total incident power density of 1 $\mathrm{kW\,m^{-2}}$, and generate it in a lab to find experimental values for the output power density of the solar cell. The reference temperature of the device is usually 25 °C. Figure 4.13 shows the evolution of the maximum efficiency found for different types of solar cells. We can see that the current highest performance is achieved by the multi-junction PV devices, which use a combination of P-N junctions with a different band gap for each semiconductor compound [21]. The materials are chosen with decreasing band gaps, so that the higher energy photons are absorbed first. In this way, the multi-junction solar cell absorbs photons over a large range of frequencies, and loses less energy to thermal relaxation of electrons in the conduction band. Other promising results are shown for the *thin film* technologies [22], which focus on reducing the material consumption in an attempt to make PV energy generation more financially competitive. Finally, there is an increasing interest in developing organic cells, which have the potential of being very cost effective and having far less impact on the environment.

4.3.2 Shockley-Queisser Limit

In 1961, Shockley and Queisser proposed a theoretical upper limit for the efficiency of PV devices with a single P-N junction, which solely depends on the band gap of the absorber material [24]. Their derivation is based on the principle of *detailed balance*, which states that [25]: "*transitions between any two states take place with equal frequency in either direction at equilibrium*". They also make the assumption that every incident photon with an energy above the band gap ($h\nu > E_g$), no matter how high its energy, produces a single electron-hole pair[3]. This corresponds to setting the absorptivity $a(E)$ - the chance that an incoming photon is captured by the absorber layer - to a step function:

$$a(E) = \begin{cases} 0, & \text{if } E < E_g \\ 1, & \text{if } E \geq E_g \end{cases} \tag{4.12}$$

To calculate P_m, the power density $P = JV$ is maximized versus the voltage V, where the current density J is derived from the ideal $J - V$ characteristic of an illuminated solar cell (Eq. (4.13)):

$$J = J_{sc} - J_0 \left(e^{\frac{eV}{k_B T}} - 1 \right). \tag{4.13}$$

---

[3] As Shockley and Queisser note in their original paper, it is possible to have a higher absorber layer efficiency than the SQ limit, for example if we include the possibility that one photon produces multiple electron-hole pairs.

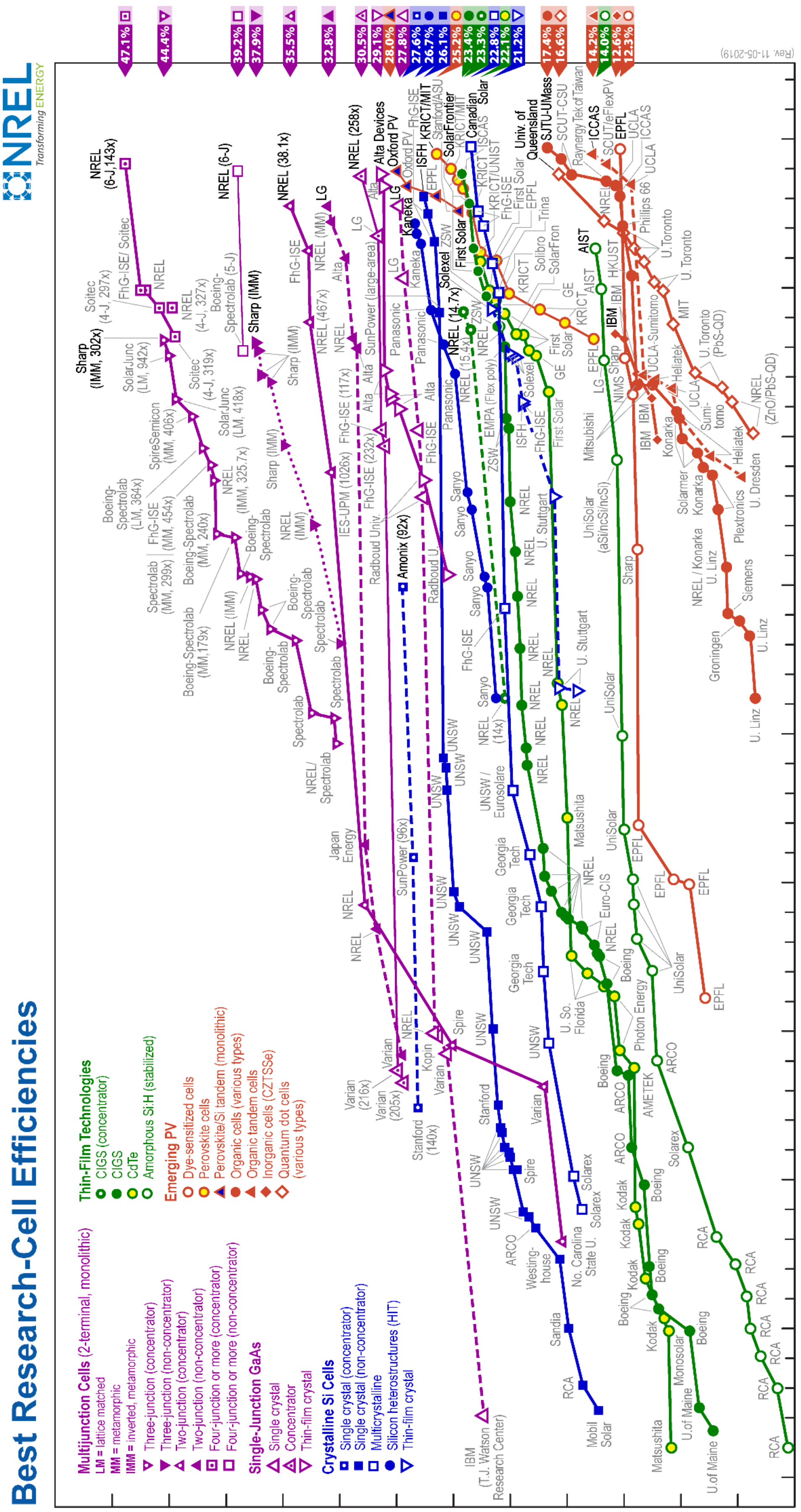


Figure 4.13: Overview of the evolution of solar cell efficiencies over the past four decades. Provided by the National Renewable Energy Laboratory [23].

The short-circuit current density J_{sc}, also known as the photo-generated current or the illuminated current, is calculated from the number of photons of the solar spectrum that are absorbed by the solar cell:

$$
\begin{aligned}
J_{sc} &= e\int_0^\infty a(E)\Phi_s(E)dE && (4.14)\\
&\overset{(4.12)}{=} e\int_{E_g}^\infty \Phi_s(E)dE && (4.15)
\end{aligned}
$$

where $\Phi_s(E)$ is the photon flux density of the solar spectrum. In their original paper, Shockley and Queisser used a black-body spectrum of $T_s = 6000$ K, but the current convention is to use the AM1.5G solar spectrum [9]. The reverse saturation current density J_0 is calculated by considering the principle of detailed balance, i.e. in equilibrium conditions the rate of photon emission from radiative recombination must be equal to the photon absorption from the surrounding medium. Because the cell is considered to be attached to an ideal heat sink, the ambient temperature is the same as that of the solar cell. Hence, the spectrum of the surrounding medium is that of a black body at cell temperature T:

$$
\begin{aligned}
J_0 &= e\pi\int_0^\infty a(E)\Phi_{bb}(E)dE \\
&= e\pi\int_0^\infty a(E)\frac{2E^2}{h^3c^2}\frac{dE}{e^{\frac{E}{k_BT}}-1} && (4.16)\\
&\overset{(4.12)}{=} e\pi\int_{E_g}^\infty \frac{2E^2}{h^3c^2}\frac{dE}{e^{\frac{E}{k_BT}}-1} && (4.17)
\end{aligned}
$$

where h is Planck's constant and c is the speed of light. As mentioned previously, because of its connection with the recombination of electron-hole pairs at equilibrium, J_0 is also referred to as the recombination current density [16]. This is the convention used in the sections that follow.

The resulting function for the efficiency only depends on the band gap E_g and is called the Shockley-Queisser or detailed balance limit. Figure 4.14 shows the curve for the SQ limit for band gaps with an efficiency above 5%. The maximum conversion of the incoming power density is 33.7%, found for a band gap of 1.34 eV.

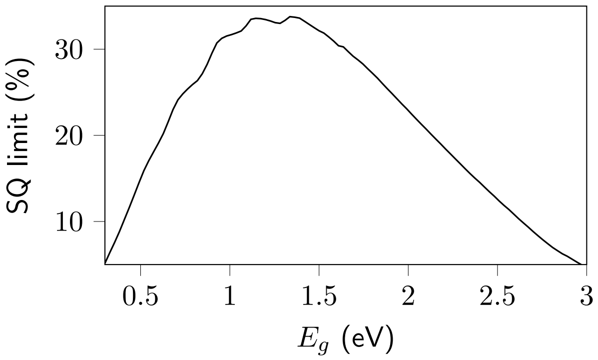


Figure 4.14: The Shockley-Queisser limit, based on the AM1.5G spectrum.

4.3.3 Spectroscopic Limited Maximum Efficiency

Although the work of Shockley and Queisser was an important step forward, only relying on the band gap as the only piece of information to determine the efficiency of a compound is too limited. Modern quantum mechanical models allow us to determine the response of a material to an incident photon in much more detail, based on first-principles calculations. The spectroscopic limited maximum efficiency (SLME), introduced by Yu and Zunger [26], attempts to use these advances in computational materials science to get a better picture of which semiconductors have the greatest potential as absorber materials. The SLME also includes the thickness of the absorber layer in the calculation of the maximum efficiency, and is therefore particularly suited to study the thin-film technologies presented in Fig. 4.13. Since its conception, the SLME has been successfully applied to perovskites [27–30], chalcogenides [31, 32], direct band gap silicon crystals [33, 34] and other materials [35–38].

The SLME tries to improve upon the Shockley-Queisser limit in two ways:

I. Instead of assuming that every photon with an energy greater than the band gap is absorbed with absolute certainty, the SLME incorporates the absorption coefficient $\alpha(E)$ (Sec. 4.2.2) of the material in order to more accurately model the generation of electron-hole pairs. After calculating the absorption coefficient from first-principles, it is included in both the calculation of the short circuit current and the recombination current. This is done through the *absorptivity* $a(E)$, which is defined as the fraction of sunlight that is absorbed when the photons pass twice through a layer of thickness L:

$$a(E) = 1 - e^{-2\alpha(E)L}. \tag{4.18}$$

II. Similar to the SQ limit, the SLME models radiative recombination using the absorptivity in combination with the black-body spectrum at the temperature T of the device. In order to include the non-radiative recombination mechanisms, Yu and Zunger make the following consideration. The total recombination current is the sum of the radiative and non-radiative parts $J_0 = J_0^r + J_0^{nr}$. If one defines f_r as the fraction of radiative recombination[4], the total recombination current density can be written as:

$$J_0^r = f_r J_0 \quad \Leftrightarrow \quad J_0 = \frac{J_0^r}{f_r}. \tag{4.19}$$

The fraction of radiative recombination is expressed as a Boltzmann factor:

$$f_r = e^{-\frac{\Delta}{kT}}, \tag{4.20}$$

where Δ is the difference in energy between the direct allowed and fundamental band gap: $\Delta = E_g^{da} - E_g$. Using this approximation, the recombination of direct band gap absorber materials is entirely radiative in nature ($f_r = 1$). However, for indirect band gap semiconductors ($\Delta \neq 0$), the non-radiative recombination quickly becomes the dominant mechanism. This is in accordance with what is observed experimentally for direct and indirect band gap semiconductors.

---

[4] Actually, Shockley and Queisser also considered the fraction of radiative recombination in their original paper [24]. They did not, however, provide a model to calculate it, simply observing that the maximum efficiency is significantly reduced for small fractions f_r.

The process for determining the SLME is schematically presented in Fig. 4.15. First, the absorptivity is computed from the absorption coefficient, derived from *ab initio* calculations. The absorptivity is then used to determine the short circuit and radiative recombination current densities:

$$\begin{aligned} J_{sc} &= e \int_0^\infty a(E)\Phi_s(E)dE, \\ J_0^r &= e\pi \int_0^\infty a(E)\Phi_{bb}(E,T)dE, \end{aligned} \tag{4.21}$$

where $\Phi_s(E)$ and $\Phi_{bb}(E,T)$ are once again the solar spectrum AM1.5G and the black-body spectrum at device temperature T. The optical and fundamental band gaps, also retrieved from first-principles calculations, are then used to derive the radiative fraction f_r (Eq. (4.20)). From here, it is a simple matter of dividing J_0^r by f_r to find the total recombination current density J_0. Once both J_{sc} and J_0 have been determined, the J-V characteristic of the illuminated P-N junction (Eq. (4.13)) is used to calculate the total current density versus a range of voltages V up to the open circuit voltage V_{oc}. Finally, the product $J \cdot V$ is maximized versus V to find the maximum output power density P_m. The SLME efficiency is the result of P_m divided by the total incident power density P_{in} = 1 kW/m^2.

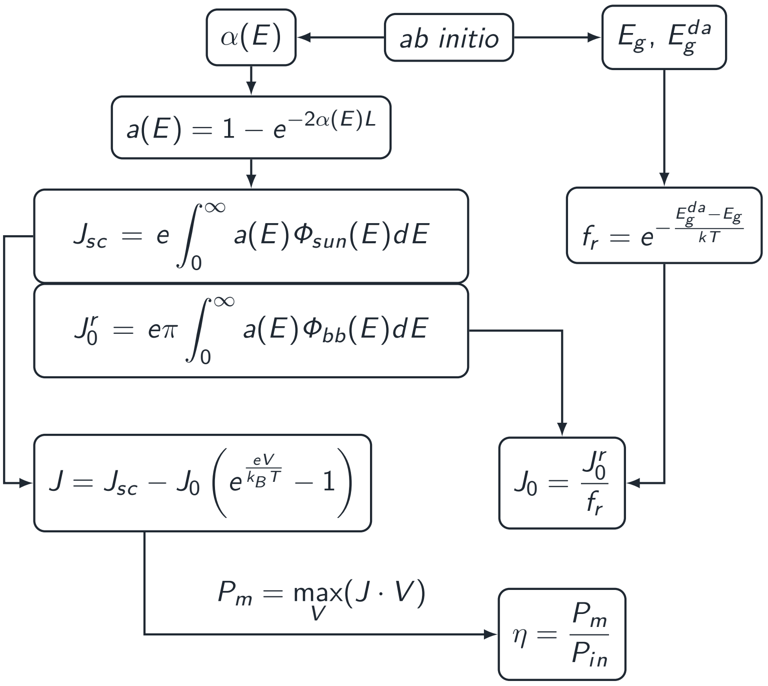


Figure 4.15: Schematic representation of the calculation of the SLME metric.

4.4 CuAu-likes

Ternary I-III-VI$_2$ semiconductors, such as the well known Cu(In, Ga)(S, Se)$_2$ compounds, are commonly used as absorber materials to produce highly flexible and lightweight solar cells. The high absorption coefficient of these compounds allows for cost-efficient absorber layers that are particularly suited for deposition on flexible substrates [39]. Laboratory values for the efficiency of CuIn(S,Se)$_2$ thin film solar cells have recently reached a record value of 22.3%. Furthermore,

$CuIn(S,Se)_2$ is also considered a suitable material for the top cell in tandem structures [40] and quantum dot based luminescent solar concentrators [41]. The rapid succession of new record efficiencies indicates that there is still room for improvement in these applications.

The most common phase of I-III-VI_2 class materials is chalcopyrite (CH). When growing films of these compounds, however, they are often found to contain CuAu-like (CA) domains, a metastable phase of chalcopyrite [42]. Moreover, it has been reported that for $CuInS_2$, the presence of the CuAu-like phase improves the short circuit current of the chalcopyrite-based photovoltaic cell. In this section I present a first-principles investigation of the efficiency of the CA phase for a selection of compounds. Section 4.4.1 presents the structure of the CH and CA phases, as well as an analysis of the thermodynamic stability of CH versus CA, in order to determine the likelihood of the presence of CA domains within a CH-based solar cell. Section 4.4.2 continues by presenting the optoelectronic properties of the CA phase materials. Finally, these results are used to calculate the SLME and the section concludes with a discussion the obtained efficiencies of specific compounds.

4.4.1 Structure and formation energy

I-III-VI_2 compounds are stable at room temperature in the chalcopyrite (CH) structure (space group $I\bar{4}2d$). However, Su and Wei [42] have demonstrated the presence of CuAu-like (CA) orderings (space group $P\bar{4}2m$) in thin films of $CuIn(S,Se)_2$, grown by vapor-phase epitaxy on Si and GaAs substrates. Alvarez et al. [43] also analyzed films of $CuInS_2$, using X-ray diffraction to estimate the relative amount of phase domains. They found that the total amount of CA ordered phase in samples grown under Cu-poor conditions was between 8% and 25%. By growing films of $CuInS_2$ on various Si substrates, Su et al. [44] discovered that although the CA phase is always present, the amount of CA domains is influenced by the substrate orientation. Moreover, Hahn et al. [45] found that by using a Si(001) substrate, the CA phase will dominate the orderings of the cation sublattice. Recently, Moreau et al. [46] have stated that for the $CuInS_2$ compound, introducing domains of CA phase can lead to a reduction of strain in the absorber layer, resulting in an increased carrier mobility and reduced recombination. Despite the fact that this phase is often found together with CH in thin films, little research has been done to determine its properties. Figure 4.16 shows the CH and CA structure of the ternary I-III-VI_2 materials.

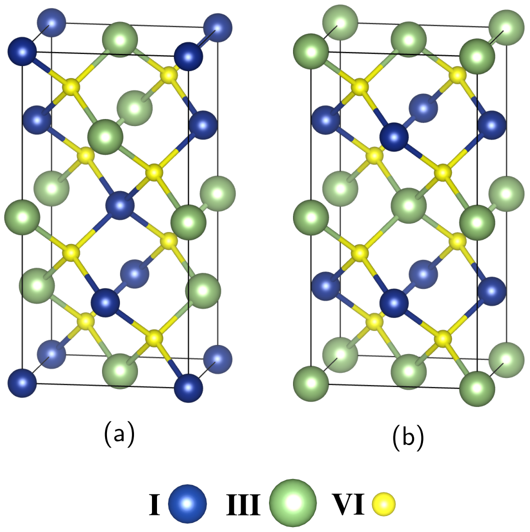


Figure 4.16: Chalcopyrite (a) and CuAu-like (b) structure of ternary I-III-VI_2 compounds.

Table 4.1: Calculated lattice parameters of the CuAu-like (CA) and chalcopyrite (CH) phase of the considered compounds, compared with the experimental results from Hahn et al. [47].

Material	CA			CH			CH (Ref [47])		
	a (Å)	c (Å)	c/a	a (Å)	c (Å)	c/a	a (Å)	c (Å)	c/a
$\mathrm{AgGaSe_2}$	5.702	12.663	2.221	6.045	11.267	1.864	5.973	10.88	1.823
$\mathrm{AgGaTe_2}$	6.220	13.060	2.100	6.403	12.327	1.925	6.283	11.94	1.897
$\mathrm{AgInS_2}$	5.780	12.132	2.100	5.925	11.554	1.950	5.816	11.17	1.920
$\mathrm{AgInTe_2}$	6.511	13.224	2.031	6.570	13.000	1.979	6.406	12.56	1.962
$\mathrm{CuGaS_2}$	5.341	10.861	2.033	5.384	10.669	1.982	5.349	10.47	1.958
$\mathrm{CuGaSe_2}$	5.662	11.436	2.020	5.683	11.277	1.984	5.607	10.99	1.960
$\mathrm{CuGaTe_2}$	6.109	12.170	1.992	6.091	12.160	1.996	5.994	11.91	1.987
$\mathrm{CuInS_2}$	5.636	11.129	1.975	5.598	11.274	2.014	5.517	11.06	2.005
$\mathrm{CuInSe_2}$	5.914	11.710	1.980	5.881	11.840	2.013	5.773	11.55	2.001
$\mathrm{CuInTe_2}$	6.323	12.590	1.991	6.313	12.681	2.009	6.167	12.34	2.000

To estimate the likelihood of finding a significant amount of CA domains in $\mathrm{CuInSe_2}$, Wei et al. [48] used first-principles calculations to determine the difference in formation energy $\Delta E_f = E_{tot}^{CA} - E_{tot}^{CH}$ between the CH and CA phases of the compound. They found a very small energy difference of 2 meV/atom, which led them to predict the coexistence of the CH and CA structures in $\mathrm{CuInSe_2}$. This was confirmed experimentally by Su and Wei [42], supporting the idea that the presence of CA domains is a result of bulk thermodynamics. In order to determine the formation energy difference, we first optimize the structure of the CA and CH phase for each compound.

Table 4.1 presents the calculated lattice parameters and c/a ratio, as well as the corresponding experimental values for the CH phase of the compounds[3]. We can see that the calculated c/a ratios match well with those obtained from experiment, with a slight overestimation of the calculated results compared to experiment. For the CA phase, replacing the cations Ag by Cu or Ga by In decreases the c/a ratio of the unit cell. This trend is reversed for the CH phase. Comparing the c/a ratio of the CA and CH phase, we find a large difference in the c/a ratio for the $\mathrm{AgGa{-}VI_2}$ compounds. Next, Table 4.2 presents the difference in formation energy for the selected list of compounds. Our first-principles results for $\mathrm{CuInS_2}$, $\mathrm{CuInSe_2}$ and $\mathrm{CuGaSe_2}$ correspond within 1 meV with those of Su et al. [42]. Similar to the results for the c/a ratio, the choice of cations has a large influence on the difference in formation energy. From Table 4.2, it is clear that substituting either In by Ga or Cu by Ag increases the difference in formation energy of the two phases. This means that if we consider the existence of the CA phase to be controlled by bulk thermodynamics, we expect CA domains to be common in the $\mathrm{CuIn{-}VI_2}$ compounds, and less likely in the $\mathrm{AgGa{-}VI_2}$ ones.

---

[3] No experimental values were found for the CA phase in the literature.

Table 4.2: Difference in formation energy between the chalcopyrite and CuAu-like structure of the considered ternary I-III-VI$_2$ compounds. All energy differences are expressed in meV/atom. The results of Su et al. [42] for CuInS$_2$, CuInSe$_2$ and CuGaSe$_2$ are also tabulated for comparison.

Material	ΔE_f	ΔE_f [42]
AgGaSe$_2$	31.3	-
AgGaTe$_2$	27.8	-
AgInS$_2$	8.9	-
AgInTe$_2$	8.5	-
CuGaS$_2$	8.8	-
CuGaSe$_2$	9.9	9
CuGaTe$_2$	7.0	-
CuInS$_2$	1.6	2
CuInSe$_2$	2.2	2
CuInTe$_2$	2.9	-

4.4.2 Absorber layer efficiency

For all of the investigated compounds, we find a direct band gap at the Γ-point. Table 4.3 presents a comparison between the G_0W_0@HSE06 band gaps calculated for the CA and CH structures[4]. The G_0W_0 calculated band gaps for the CA phase are lower than those of the CH phase for all compounds besides CuInS$_2$. Furthermore, the difference is smaller for the I-III-S$_2$ structures compared to the I-III-(Se,Te)$_2$ compounds. Table 4.3 also contains the experimental band gaps of the CH phase of the compounds. Although the G_0W_0@HSE06 band gaps correspond quite well to the experimental values for some compounds, there are clear discrepancies for others. This could be a result of the sensitivity of chalcogenide band gaps to the anion displacement u [49]. As an example of the dielectric function, we show the result for CA-CuInSe$_2$ in Fig. 4.17. The results for the other compounds can be found in the Jupyter notebooks corresponding to this section.

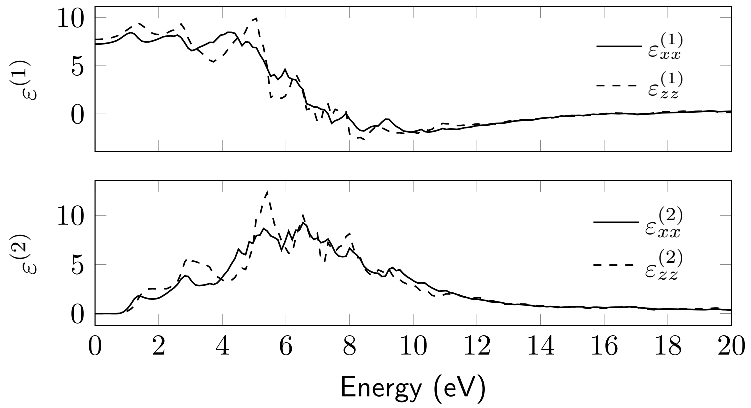


Figure 4.17: Real (upper figure) and imaginary (lower figure) parts of the HSE06 calculated dielectric function of CA-CuInSe$_2$.

---

[4] We did not take all of the CH phase band gaps from Yu and Zunger [26], because of inconsistencies between the tabulated and plotted values for some compounds in this paper.

Table 4.3: Experimental and calculated band gaps of the CuAu-like(CA) and chalcopyrite (CH) phase of the considered compounds.

Material	CA		CH	
	E_g^{HSE}(eV)	$E_g^{G_0W_0}$(eV)	$E_g^{G_0W_0}$(eV)	E_g^{exp}(eV)
$AgGaSe_2$	0.84	1.41	1.80[a]	1.83[b]
$AgGaTe_2$	0.46	0.95	1.54[a]	1.1-1.3[b]
$AgInS_2$	1.20	1.69	1.74[a]	1.87[b]
$AgInTe_2$	0.53	0.92	1.23[a]	0.96-1.04[b]
$CuGaS_2$	1.77	1.94	1.99[a]	2.41[c]
$CuGaSe_2$	0.96	1.19	1.65[a]	1.64[c]
$CuGaTe_2$	0.77	1.06	1.47[a]	1.23[b]
$CuInS_2$	1.14	1.13	1.54[a]	1.53[c]
$CuInSe_2$	0.59	0.58	1.33[a]	1.04[c]
$CuInTe_2$	0.76	0.94	1.03[a]	0.96[b]

[a] Ref. [26], [b] Ref. [50], [c] Ref. [51]

Table 4.4: Calculated SLME for both the CuAu-like and chalcopyrite [26] structures at $L = 500$ nm. The SQ limit of the corresponding band gap of the CA compounds is also given as a reference.

Material	SLME(%) (CA)	SQ(%) (CA)	SLME(%) (CH)
$AgGaSe_2$	27.0	33.3	15.8
$AgGaTe_2$	28.9	31.1	21.8
$AgInS_2$	23.1	29.1	19.7
$AgInTe_2$	28.2	30.5	26.4
$CuGaS_2$	16.4	24.1	16.5
$CuGaSe_2$	27.8	33.4	26.6
$CuGaTe_2$	28.9	32.0	24.8
$CuInS_2$	29.0	33.5	23.1
$CuInSe_2$	20.7	18.3	22.1
$CuInTe_2$	27.9	30.9	28.0

Once the band gap and dielectric function for the CA phase of the selected list of compounds has been determined, we have all the required information to calculate their SLME. Because the CA phase of all of the compounds has a direct allowed fundamental band gap ($E_g^{da} = E_g$), the non-radiative recombination is considered to be negligible for the SLME metric ($f = 1$, see Eq. (4.20)). Table 4.4 presents the calculated efficiency values. In order to compare our results with those of Yu and Zunger, all efficiencies are calculated using thickness $L = 500$ nm and device temperature $T = 300$ K. First, note that several CH structures that are known to have high device efficiencies, such as $CuIn(S,Se)_2$, also have a high SLME. Moreover, it is clear that although the band gap has a large influence on the efficiency, some materials, such as CA- and CH-$AgInS_2$, have a very similar band gap but a significantly different calculated efficiency. This demonstrates the ability of the SLME to provide a more refined selection metric in comparison with the SQ limit. Finally, for several compounds the CA phase has a higher efficiency than the corresponding CH phase. This is consistent with the findings of Moreau et al. [46], who discovered that the presence of CA domains have a positive influence on the efficiency of $CuInS_2$. We suggest that the efficiency of these devices may have benefited from the presence of the CA phase directly through the optical properties of the material.

Figure 4.18a shows the SLME as a function of the film thickness for the CA phase of the I-III-VI_2 compounds. It is clear that for most compounds, the efficiency of the CA phase rises quickly for an increasing thickness. In fact, for $L = 500$ nm, the SLME is already within 5-6% for all compounds save $CuGaS_2$, which demonstrates the potential of the CA phase materials as absorber layers in thin-film solar cells. Fig. 4.18b plots the SLME for $L = 500$ nm of the CA phase of the various compounds versus their band gap, as well as the SQ limit. Observe that the SLME value for CA-$CuInSe_2$ is higher than the corresponding SQ limit. In Section 4.5 I return to this result and discuss it in detail. Finally, note that the influence on the efficiency of a discrepancy between the calculated and experimental band gaps depends on the band gap of the compound. Looking at Fig. 4.18b, one can expect the influence of the band gap to be small in the 1-1.5 eV interval. In case the calculated and experimental band gap are not in this region, however, any discrepancy between the calculated and experimental band gap is likely to significantly influence the SLME.

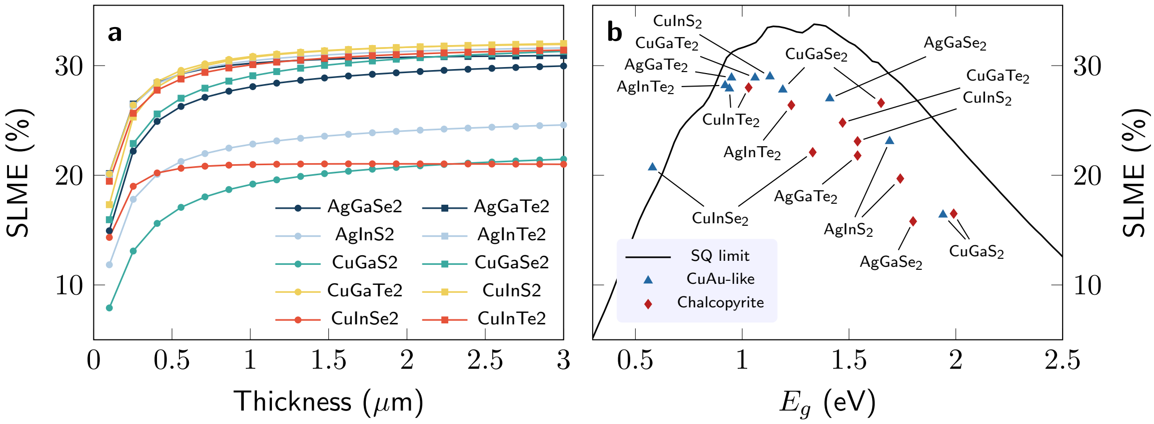


Figure 4.18: (**a**) Calculated maximum efficiencies of the CuAu-like phase materials as a function of film thickness. (**b**) SLME of the CuAu-like compounds versus the band gap, calculated using thickness $L = 500$ nm and device temperature $T = 300$ K. The black line represents the Shockley-Queisser limit.

4.5 SLME Analysis

The Shockley-Queisser limit is one of the most fundamental results in the field of photovoltaics. Based on the principle of detailed balance, it defines an upper limit for a single junction solar cell that uses an absorber material with a specific band gap. Since its conception, numerous methods have been proposed to exceed the Shockley-Queisser limiting efficiency [52]. Examples include multi-junction [22, 53] and hot carrier solar cells [54], as well as concepts that use multiple exciton generation [55]. None of these concepts, however, are implemented in the SLME, certainly not when we consider the radiative limit ($f_r = 1$). However, in the previous section the SLME value of CA-$CuInSe_2$ turned out to be higher than the corresponding SQ limit.

In this section, I analyze this surprising result in more detail by taking a closer look at how the calculated efficiency depends on the thickness and band gap of the material, as well as the temperature of the device. The analysis presented in Section 4.5.1 demonstrates that the detailed balance approach for the radiative recombination current allows for higher open circuit voltages at lower thicknesses, producing a higher SLME than the corresponding SQ limit. That is, simply by dropping the assumption of an infinite absorber layer, i.e. by replacing the Heaviside step function for the absorptivity by a calculated spectrum, it is possible to obtain efficiencies above the Shockley-Queisser limit. Next, Section 4.5.2 discusses this phenomenon in more detail by replacing the absorptivity by a parameterized sigmoid function, and analyzes for which band gap range a material's efficiency is more likely to exceed the Shockley-Queisser limit. Finally, Section 4.5.3 broadens our analysis to indirect band gap materials by studying the absorption and efficiency of silicon. From these results, it is clear that in the SLME model, the fraction of non-radiative recombination can be so high that many indirect band gap absorber layers have a very high recombination current, resulting in an unreasonably low calculated efficiency.

4.5.1 The curious case of CA-CuInSe$_2$

During the discussion of the SLME results of the CA phase, I noted that CA-CuInSe$_2$ has an SLME value above the SQ limit. This result is surprising because the SQ limit is widely considered to be a theoretical maximum efficiency of a single junction absorber layer, and the SLME is based on the same detailed balance approach as the SQ limit. Due to the definition of the SLME, the calculated efficiency returns to the SQ limit for $L \to \infty$, since for an infinitely thick absorption layer the absorptivity becomes a step function. However, looking at the thickness dependence of the SLME for CA-CuInSe$_2$ and CA-CuGaSe$_2$ (Fig. 4.19a), it is apparent that the way they approach the SQ value is different. Most importantly, the SLME of the compound CuInSe$_2$ crosses the SQ limit, whereas that of CuGaSe$_2$ does not.

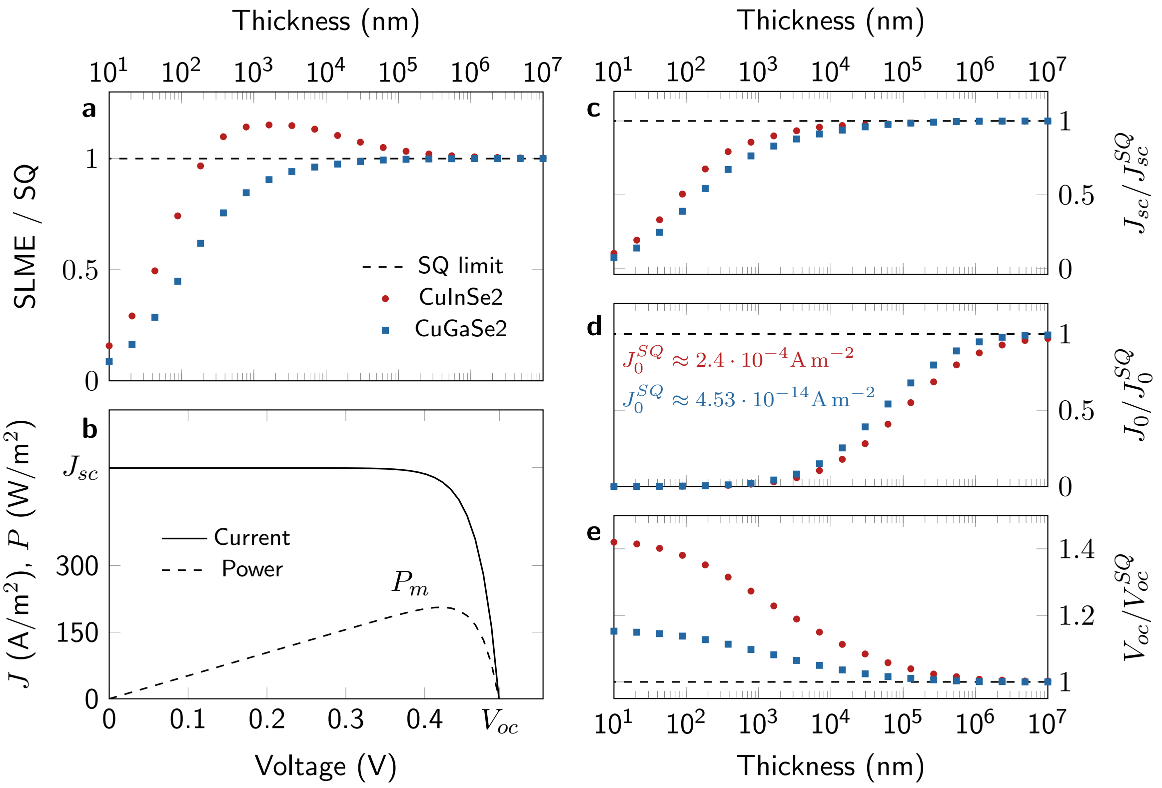


Figure 4.19: (**a**) Thickness dependence of the SLME of CA-CuInSe$_2$ and CA-CuGaSe$_2$ at 300 K versus their SQ limit. (**b**) Calculated J-V characteristic of CuInSe$_2$ at $T = 300$ K and $L = 500$ nm (full line), as well as the corresponding power density (dashed line). (**c-d**) Thickness dependence of the current densities of CuInSe$_2$ and CuGaSe$_2$, as well as the open circuit voltage V_{oc} (**e**) versus their respective SQ values.

The origin of this behavior can be understood by considering how the short-circuit current density J_{sc} and recombination current density J_0 are used to calculate the power density of the absorber layer $P = JV$. Figure 4.19b shows the calculated J-V characteristic of CuInSe$_2$. The total current density J remains close to J_{sc} up to a certain voltage. The value of this voltage, and hence the value of the open circuit voltage V_{oc} as well as the voltage that maximizes the power density V_m, depends strongly on J_0. Looking at both current densities as a function of the thickness, presented in Figs. 4.19c-d, it is clear that for both compounds J_{sc} converges to the corresponding SQ value at far lower thicknesses than J_0. The relatively low value for J_0 at certain thicknesses allows for a higher open circuit voltage V_{oc} (Fig. 4.19e). This is the case for both CuInSe$_2$ and CuGaSe$_2$. However, because the order of magnitude of J_0 is much larger for CuInSe$_2$ than for CuGaSe$_2$, the open circuit voltage V_{oc} at low thicknesses is significantly larger than the corresponding Shockley-Queisser value for CuInSe$_2$ than for CuGaSe$_2$.

To quantify this a little more: the SLME crosses the SQ limit when its maximum power density is higher than the one calculated using the SQ values for J_{sc} and J_0:

$$\begin{aligned} & J_m V_m = P_m > P_m^{SQ} = J_m^{SQ} V_m^{SQ} \\ \Leftrightarrow \qquad & \frac{V_m}{V_m^{SQ}} > \frac{J_m^{SQ}}{J_m}. \end{aligned} \tag{4.22}$$

Because the order of magnitude of J_0 is much larger for CuInSe$_2$, the value of the fraction V_m/V_m^{SQ} at low thicknesses is higher for CuInSe$_2$ when compared to that for CuGaSe$_2$ (Fig. 4.20a). In comparison, the convergence of the fraction J_m^{SQ}/J_m is similar for both compounds. From Eq. (4.22), it is clear that when V_m/V_m^{SQ} is larger than J_m^{SQ}/J_m, the maximized power density is higher than its SQ value, which means that the SLME will be higher than the Shockley-Queisser limit for that thickness. Looking at Fig. 4.20a, we can see that at $T = 300$ K, this happens for CuInSe$_2$.

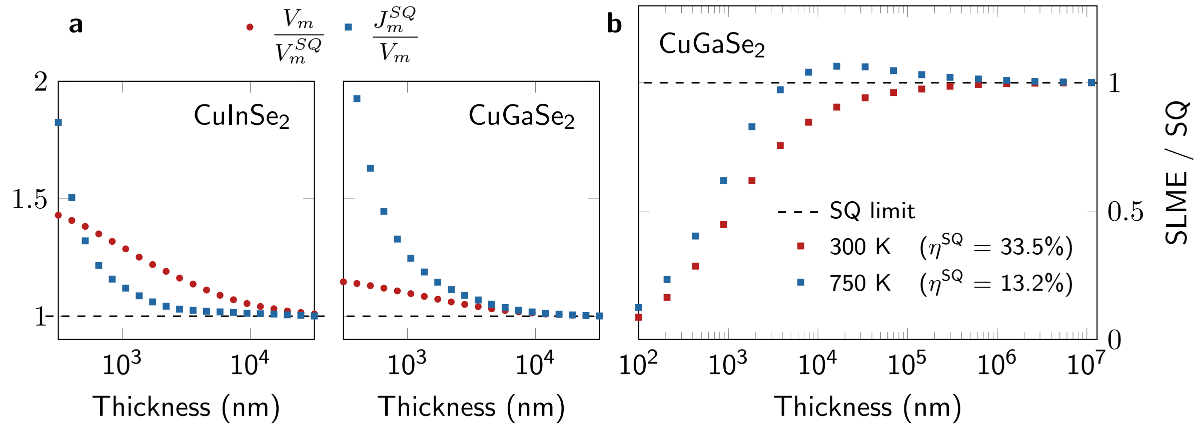


Figure 4.20: (**a**) Comparison of the relative increase of the voltage that maximizes the power density (V_{max}/V_{max}^{SQ}) with the relative decrease of the corresponding current density (J_{max}^{SQ}/J_{max}) for both CA-CuInSe$_2$ (left) and CA-CuGaSe$_2$ (right). (**b**) Thickness dependence of the SLME of CA-CuGaSe$_2$ at $T = 300$ K and $T = 750$ K.

For direct band gap absorbers, $f_r = 1$, and J_0 is calculated from the overlap of the black-body spectrum $I_{bb}(E, T)$ and the absorptivity $a(E)$ of the material. From Eq. (4.21), it should be clear that lowering the band gap increases J_0. As a result, materials with a low band gap are more likely to have an SLME value above the SQ limit at a specific thickness. It is also the case, however, that J_0 increases at higher temperatures. This raises the relative increase of V_m at lower thicknesses, potentially producing calculated efficiencies above the SQ limit. For example, looking at the thickness dependence of the SLME of CA-CuInS$_2$ at $T = 750$ K (Fig. 4.20b), we see that at this temperature the calculated efficiency also crosses the Shockley-Queisser limit.

Since the calculation of the SLME only deviates from the SQ limit by the introduction of an *ab initio* calculated absorption spectrum, these results show that the SQ limit is not a theoretical upper limit within the assumptions of the detailed balance approach. This is because considering an infinite thickness for the solar cell, i.e. taking a step function for $a(E)$, overestimates J_0 as it is calculated in the detailed balance framework. As a result, it is possible that for a material with a certain band gap and absorptivity, J_0 is very low compared to its SQ value, which allows for a high V_{oc}. In case V_{oc} is increased sufficiently, the total power density can go above that of the SQ limit, even though the calculated J_{sc} is lower than its SQ value. In other words, when considering all of the assumptions made in the Shockley-Queisser approach and introducing an absorption spectrum of an absorption layer with a finite thickness, it is possible to obtain efficiencies above the SQ limit.

4.5.2 Logistic function model

The efficiency for CA-CuInSe$_2$ is not the only one to exceed the SQ limit at 500 nm. Figure 4.21 shows a selection of calculated SLME values of direct band gap materials from previous work [26, 32, 56], compared with the Shockley-Queisser limit. Among the presented materials, a significant amount have a calculated SLME above the Shockley-Queisser limit. Since the calculation of the SLME does not introduce any of the concepts that would typically allow its value to exceed the Shockley-Queisser limit, these results further indicate that for thin-film materials with lower band gaps, the Shockley-Queisser limit does not necessarily represent an upper limit for the efficiency.

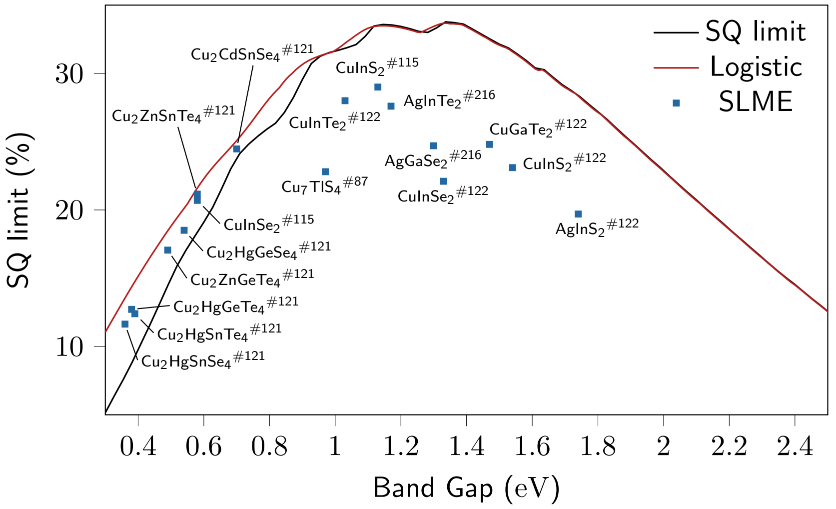


Figure 4.21: Collection of calculated SLME values from Yu and Zunger [26], as well as our results for CuAu-like [56] and Stannite [32] structures. The space group number of the material structure is added as a superscript. The efficiency values were calculated for a thickness of 500 nm. The orange curve represents the maximum efficiencies obtained using the logistic model explained in the text.

Shockley and Queisser considered their metric as the detailed balance *limit* because of the assumption that since the step function represents the highest possible absorption spectrum for a material with a specific direct band gap, the resulting efficiency must represent an upper limit. However, as we demonstrated in Section 4.5.1, this also means that the recombination current density J_0 (Eq. (4.17)) will be maximal. Since electron-hole recombination results in a loss of electrons contributing to the external current, this has a negative effect on the photovoltaic conversion efficiency. Hence, it is possible that there is an absorptivity function that would result in a higher efficiency than the Shockley-Queisser limit. As we can see in Fig. 4.21, this is exactly what happens for the presented smaller band gap materials.

The next questions are how far we can exceed the Shockley-Queisser limit, and at which band gap a material is more likely to do so. Clearly, this depends on the shape of the absorptivity function. Fig. 4.22, shows the calculated absorptivity of Cu_2ZnGeS_4 for various thicknesses, derived from the absorption coefficient calculated from first principles (For computational details, we refer the reader to [32]). The absorptivity has a shape reminiscent of a sigmoid function at all thicknesses. In order to analyze the maximum efficiency for materials with a direct band gap in the range 0.3-3 eV, we model $a(E)$ using a generalized logistic function:

$$a(E) = f(E) = \frac{1}{(1 + e^{-\delta(E-E_g)})^{\beta}}, \tag{4.23}$$

where E_g is the band gap of the material, and β, δ are parameters that determine the shape of the function. In this model for the absorptivity, the parameter δ can be related to the thickness of the material, as for $\delta \to \infty$, $f(E)$ approaches the Heaviside step function (Fig. 4.22). The second parameter (β) is important to make sure that the model function "starts" at the band gap, i.e. that its value for $E < E_g$ is suitably small, so that it can be approximated to zero. Since $f(E_g) = \frac{1}{2^{\beta}}$, and $f(E) < f(E_g)$ for $E < E_g$, increasing β to a suitably large value gives us this desired function trait. Here, we choose $\beta = 10$ and set $f(E) = 0$ for $E \leq E_g$. As is clear from Fig. 4.22, this model function describes the shape of the calculated absorptivity spectra quite well, and allows us to use it as a test function to analyze how much we can exceed the SQ limit.

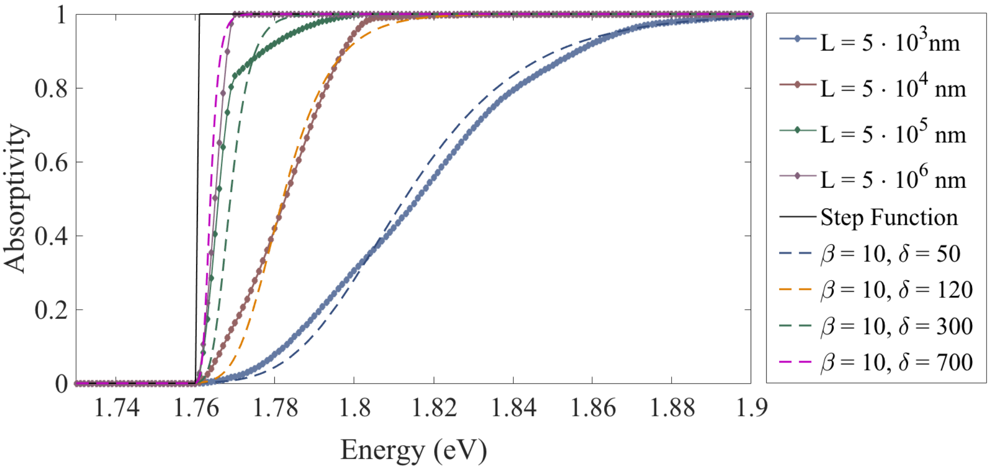


Figure 4.22: Comparison of the model function with calculated absorptivity spectra for Cu_2ZnGeS_4 at different thicknesses L. The model function shape matches that of the calculated absorptivity quite well as $L, \delta \to \infty$.

To study the influence of the band gap on the likelihood of the efficiency exceeding the Shockley-Queisser limit, we calculate the efficiency for $\delta \in [1, 10^4]$ over the band gap range $E_g \in [0.3, 3]$ eV. Figure 4.23 presents the δ-dependency of the efficiency for a selection of band gaps. For low band gaps, the calculated efficiency crosses the detailed balance limit of the corresponding band gap, in order to return to the limit value for $\delta \rightarrow \infty$. Since δ can be related to the thickness of the material, this implies that for lower band gap materials, there is a thickness that is optimal for the efficiency. Moreover, a clear trend is visible, with the efficiency exceeding the Shockley-Queisser limit more as the band gap is decreased. This is also clear when looking at the plot for the maximum efficiency values in Fig. 4.21.

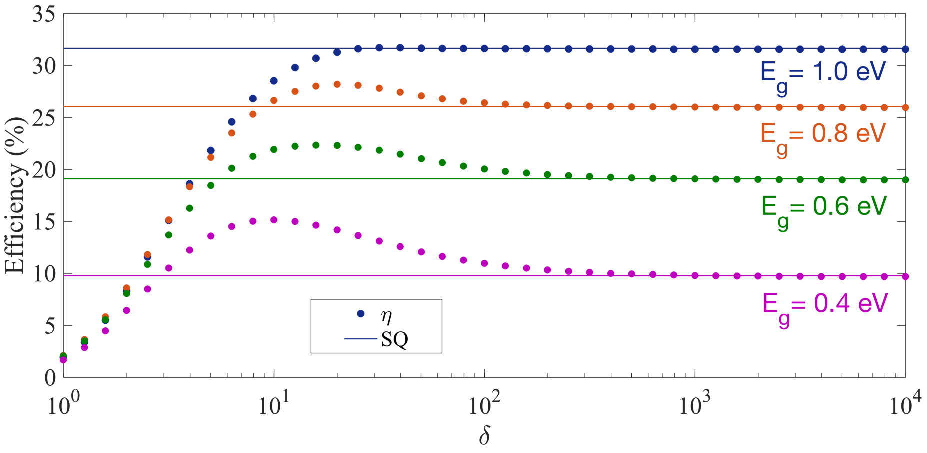


Figure 4.23: Calculated efficiencies for a range of δ values and a selection of band gaps, compared with the corresponding Shockley-Queisser limit.

Finally, it is interesting to note that the SLME values of the materials that exceed the Shockley-Queisser limit are still below the maximum efficiency for the model absorptivity functions of the corresponding band gap in Fig. 4.21. However, this does not imply that the logistic function maxima curve represents a new upper limit. It is entirely possible that there is another function profile that would allow for higher efficiencies. Using the logistic function approach, we are simply able to observe for which band gap range the Shockley-Queisser limit does not provide a theoretical upper limit within the detailed balance approach.

4.5.3 Indirect band gap absorbers

So far, I have only discussed materials which have a direct band gap. In order to test the application of the SLME to indirect band gap absorbers, we decided to calculate the SLME of silicon, which is still one of the most popular materials for the production of solar cells. Fig. 4.24a shows the experimental[5] absorption coefficient of crystalline silicon [57]. Notice the onset of the indirect and direct absorption at $E_g = 1.17$ eV and $E_g^{da} = 3.4$ eV, respectively.

Calculating the SLME using this optical spectrum produces an efficiency of zero for any thickness L and temperature T. The origin of this troubling result is rooted in the fraction of radiative recombination expressed in Eq. (4.20). Because of the large difference between the direct allowed and fundamental band gap of silicon ($\Delta = E_g^{da} - E_g = 2.23$ eV), the radiative fraction is of the

---

[5] We choose to use an experimental spectrum in order to include the phonon-mediated contributions to the absorption coefficient.

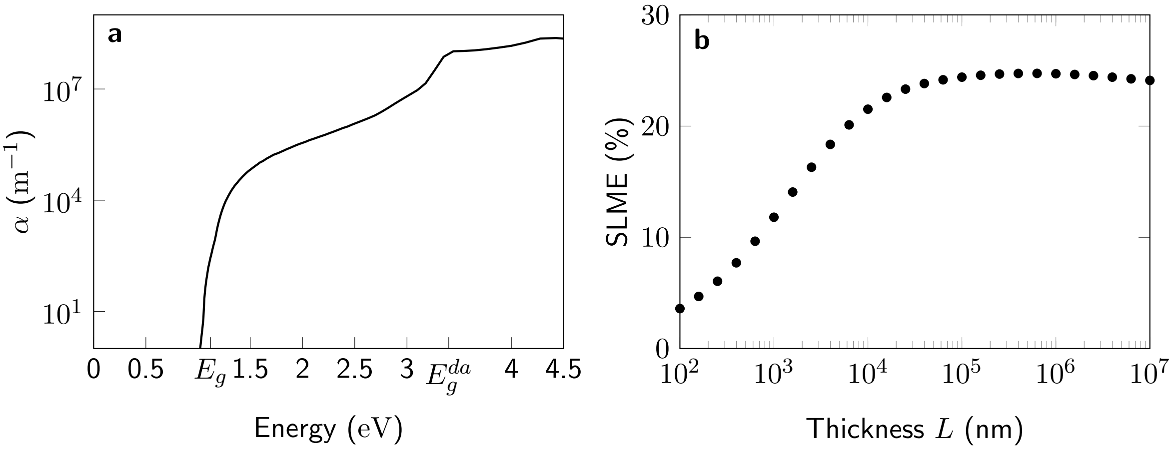


Figure 4.24: (**a**) Experimental absorption coefficient at $T = 300$ K of crystalline silicon based on data taken from [57]. (**b**) Thickness dependence of the SLME of silicon, based on the experimental spectrum in (**a**) and fraction of recombination $f_r 10^{-3}$.

order 10^{-38}. Since this fraction is used to calculate the recombination current (see Eq. (4.21)), this results in a J_0 that is unreasonably large. As discussed in Section 4.5.1, J_0 has a significant influence on the open circuit voltage V_{oc}. In this case, the high value of J_0 leads to a V_{oc} that is too small to produce any significant power density. However, in case the fraction of radiative recombination is set $f_r = 10^{-3}$ – a more reasonable value for silicon [12, 58, 59] – the results shown in Fig. 4.24b are obtained.

One could argue that silicon is a special case, and that generally efficient indirect absorbers do not have such a large band gap difference $\Delta = E_g^{da} - E_g$. For thin-film solar cells, indirect absorption also contributes significantly less to the power density. Consequently, indirect band gap materials with a large fundamental band gap are not suitable for these applications in any case. However, even for materials with a small Δ, the modeled fraction of radiative recombination quickly becomes minute. For example, consider the compound $\mathrm{Cu_3TlSe_2}$, which has been investigated by Yu and Zunger [26]. The reported difference between the fundamental and direct allowed band gap is 0.24 eV. At 300 K, the fraction of radiative recombination then becomes $f_r = 10^{-4}$. This means that although 99.99 % of the recombination is non-radiative in nature, the recombination current is still derived from an entirely radiative principle, based on the black-body spectrum in Eq. (4.21). Furthermore, it is clear that because of the exponential function in Eq. (4.20), the fraction of radiative recombination drops very rapidly with increasing Δ. This indicates that even for materials with a relatively low Δ, the recombination current will rise significantly, which is detrimental for the calculated efficiency. Hence, it is entirely possible that the recombination model of the SLME metric does not judge indirect band gap absorbers fairly, potentially eliminating good materials during the selection procedure.

4.6 Conclusions and Outlook

We have compared the structural and thermodynamic properties of the CA and CH phase of the compounds. By analyzing the difference in formation energy of the CH and CA phase, we conclude that CA domains are most likely to be present in CuIn-VI_2 compounds, which is in good agreement with experimental results. From the calculated optoelectronic properties of the materials, we have determined their potential as absorbers for solar cells by applying the SLME selection metric. We identify several compounds with a high theoretical efficiency in the CA phase, most notably CA-$CuInS_2$, which has a significantly higher efficiency than the corresponding CH phase.

After observing an SLME value above the Shockley-Queisser limit for CA-$CuInSe_2$, we have performed a detailed analysis to find the origin of this result. We find that, within the details balance approach, the reverse saturation current J_0 approaches its SQ value very slowly for an increasing thickness L. This causes the SLME to cross the SQ limit for materials with a J_0 that is relatively high, i.e. materials with a low band gap or at higher temperatures. In their 1961 paper, Shockley and Queisser characterized their calculated efficiency as an upper limit, because of the assumption that if every photon with an energy above the band gap is absorbed, the obtained efficiency must be maximal. Although this assumption may seem entirely sensible at first glance, it does not consider the fact that it also maximizes the recombination current, which is calculated using the detailed balance principle. Because an increased recombination current results in a lower efficiency, this means that lowering the absorptivity can produce higher efficiencies than the Shockley-Queisser limit under the right conditions. By using a model absorptivity function, which closely resembles absorptivity spectra calculated from first principles, we have shown that this can occur for low band gaps. This means that one must take care when dismissing low band gap materials based on their Shockley-Queisser limit, for their actual efficiency at certain thicknesses might still make them suitable for thin film photovoltaic applications. Finally, we have shown that the model that introduces non-radiative recombination to the SLME quickly undercuts the efficiency of indirect band gap absorbers, as the Boltzmann factor used increases the recombination current drastically once the difference between the direct and fundamental band gap becomes larger.

Although the SLME has shown promise as a metric for computational materials design of photovoltaic absorber layers, there are still some issues which need to be resolved. The SLME could benefit from an improved description of the fraction of radiative recombination, especially if it is to be applied to indirect band gap absorbers. Moreover, other important effects can be introduced, such as multiple exciton generation and photon recycling. Finally, we also note that as the efficiency of a solar cell depends heavily on the band gap of a material, calculating the SLME based on results from density functional theory calculations still poses the risk of serious inaccuracy, especially when the band gap is outside the 1-1.5 eV range.

Bibliography

[1] J. Marsden. “Distributed generation systems: A new paradigm for sustainable energy”. In: *2011 IEEE Green Technol. Conf. Green 2011*. 2011. ISBN: 9781612847146. DOI: `10.1109/GREEN.2011.5754858`.

[2] K. P. Bhandari et al. “Energy payback time (EPBT) and energy return on energy invested (EROI) of solar photovoltaic systems: A systematic review and meta-analysis”. *Renewable Sustainable Energy Rev.* 47 (2015), pp. 133–141. DOI: `10.1016/j.rser.2015.02.057`.

[3] M. S. Chowdhury et al. “An overview of solar photovoltaic panels’ end-of-life material recycling”. *Energy Strategy Rev.* 27 (2020), p. 100431. DOI: `10.1016/j.esr.2019.100431`.

[4] NREL. *News Release: Tenth Edition of Data Book Reveals Trends in U.S. and Global Renewable Energy Growth.* URL: `https://www.nrel.gov/news/press/2019/tenth-edition-of-data-book-reveals-trends-in-us-and-global-renewable-energy-growth.html` (visited on 11/28/2019).

[5] M. Planck. “On the Law of Distribution of Energy in the Normal Spectrum”. *Ann. Phys.* 4 (1901), p. 533.

[6] R. E. Bird and C. Riordan. “Simple solar spectral model for direct and diffuse irradiance on horizontal and tilted planes at the earth’s surface for cloudless atmospheres.” *J. Clim. Appl. Meteorol.* 25 (1986), pp. 87–97. DOI: `10.1175/1520-0450(1986)025⟨0087:SSSMFD⟩2.0.CO;2`.

[7] M. A. Green. *Solar Cells: Operating Principles, Technology, and System Applications.* Prentice-Hall series in solid state physical electronics. Prentice Hall, 1981. ISBN: 0138222703.

[8] Newport. *Introduction to Solar Radiation.* URL: `https://www.newport.com/t/introduction-to-solar-radiation` (visited on 07/20/2015).

[9] *ASTM G173-03. Standard Tables for Reference Solar Spectral Irradiances: Direct Normal and Hemispherical on 37° Tilted Surface.* ASTM International, West Conshohocken, PA, 2012, `www.astm.org`. DOI: `10.1520/G0173-03R12`.

[10] P Auger. “Sur les rayons beta secondaries produits dans gaz par des rayons X”. *C.R.A.S.* 177 (1923), pp. 169–171.

[11] N. G. Nilsson. “Band-to-band auger recombination in silicon and germanium”. *Phys. Scr.* 8 (1973), pp. 165–176. DOI: `10.1088/0031-8949/8/4/007`.

[12] W. Shockley and W. T. Read. “Statistics of the recombinations of holes and electrons”. *Phys. Rev.* 87 (1952), pp. 835–842. DOI: `10.1103/PhysRev.87.835`.

[13] R. N. Hall. “Electron-hole recombination in germanium”. *Phys. Rev.* 87 (1952), p. 387. DOI: `10.1103/PhysRev.87.387`.

[14] W. Shockley. “The Theory of p-n Junctions in Semiconductors and p-n Junction Transistors”. *Bell Syst. Tech. J.* 28 (1949), pp. 435–489. DOI: `10.1002/j.1538-7305.1949.tb03645.x`.

[15] *PN junction.* URL: `https://en.wikipedia.org/wiki/File:Pn-junction-equilibrium-graphs.png` (visited on 11/14/2019).

[16] A. Cuevas. “The recombination parameter J_0”. In: *Energy Procedia*. Vol. 55. Elsevier Ltd, 2014, pp. 53–62. DOI: `10.1016/j.egypro.2014.08.073`.

[17] S. J. Fonash. *Solar Cell Device Physics*. Elsevier Inc., 2010. ISBN: 9780123747747. DOI: `10.1016/C2009-0-19749-0`.

[18] B. Swatowska et al. "The role of antireflective coatings in silicon solar cells - The influence on their electrical parameters". In: *Opt. Appl.* Vol. 41. 2. 2011, pp. 487–492.

[19] K. Grattan. "Optoelectronics of solar cells". *Opt. Laser Technol.* 36 (2004), p. 515. DOI: `10.1016/j.optlastec.2004.01.014`.

[20] F. A. Lindholm, J. G. Fossum, and E. L. Burgess. "Application of the Superposition Principle to Solar-Cell Analysis". *IEEE Trans. Electron Devices* 26 (1979), pp. 165–171. DOI: `10.1109/T-ED.1979.19400`.

[21] F. Dimroth and S. Kurtz. "High-efficiency multijunction solar cells". *MRS Bull.* 32 (2007), pp. 230–235. DOI: `10.1557/mrs2007.27`.

[22] A. V. Shah et al. "Thin-film silicon solar cell technology". *Prog. Photovoltaics Res. Appl.* 12 (2004), pp. 113–142. DOI: `10.1002/pip.533`.

[23] N. R. E. Laboratory. *Best Research-Cell Efficiency Chart.* URL: `https://www.nrel.gov/pv/cell-efficiency.html` (visited on 11/02/2019).

[24] W. Shockley and H. J. Queisser. "Detailed balance limit of efficiency of p-n junction solar cells". *J. Appl. Phys.* 32 (1961), pp. 510–519. DOI: `10.1063/1.1736034`.

[25] M. J. Klein. "Principle of detailed balance". *Phys. Rev.* 97 (1955), pp. 1446–1447. DOI: `10.1103/PhysRev.97.1446`.

[26] L. Yu and A. Zunger. "Identification of potential photovoltaic absorbers based on first-principles spectroscopic screening of materials". *Phys. Rev. Lett.* 108 (2012), p. 068701. DOI: `10.1103/PhysRevLett.108.068701`.

[27] W.-J. Yin, T. Shi, and Y. Yan. "Unique Properties of Halide Perovskites as Possible Origins of the Superior Solar Cell Performance". *Adv. Mater.* 26 (2014), pp. 4653–4658. DOI: `10.1002/adma.201306281`.

[28] W.-J. Yin et al. "Halide perovskite materials for solar cells: a theoretical review". *J. Mater. Chem. A* 3 (2015), pp. 8926–8942. DOI: `10.1039/c4ta05033a`.

[29] W.-J. Yin, T. Shi, and Y. Yan. "Superior Photovoltaic Properties of Lead Halide Perovskites: Insights from First-Principles Theory". *J. Phys. Chem. C* 119 (2015), pp. 5253–5264. DOI: `10.1021/jp512077m`.

[30] W. Meng et al. "Alloying and Defect Control within Chalcogenide Perovskites for Optimized Photovoltaic Application". *Chem. Mater.* 28 (2016), pp. 821–829. DOI: `10.1021/acs.chemmater.5b04213`.

[31] F. Hong et al. "Trigonal Cu_2-II-Sn-VI_4 (II = Ba, Sr and VI = S, Se) quaternary compounds for earth-abundant photovoltaics". *Phys. Chem. Chem. Phys.* 18 (2016), pp. 4828–4834. DOI: `10.1039/c5cp06977g`.

[32] N. Sarmadian et al. "First-principles study of the optoelectronic properties and photovoltaic absorber layer efficiency of Cu-based chalcogenides". *J. Appl. Phys.* 120 (2016), p. 085707. DOI: `10.1063/1.4961562`.

[33] I. H. Lee et al. "Computational search for direct band gap silicon crystals". *Phys. Rev. B - Condens. Matter Mater. Phys.* 90 (2014), p. 115209. DOI: `10.1103/PhysRevB.90.115209`.

[34] Y. J. Oh et al. "Dipole-allowed direct band gap silicon superlattices". *Sci. Rep.* 5 (2015), p. 18086. DOI: `10.1038/srep18086`.

[35] L. Yu et al. "Inverse Design of High Absorption Thin-Film Photovoltaic Materials". *Adv. Energy Mater.* 3 (2012), pp. 43–48. DOI: `10.1002/aenm.201200538`.

[36] T. Yokoyama et al. “Theoretical Photovoltaic Conversion Efficiencies of $ZnSnP_2$, $CdSnP_2$, and $Zn_{1-x}Cd_xSnP_2$ Alloys”. *Appl. Phys. Express* 6 (2013), p. 61201. DOI: `10.7567/apex.6.061201`.

[37] J. Heo et al. “Design meets nature: Tetrahedrite solar absorbers”. *Adv. Energy Mater.* 5 (2015), p. 1401506. DOI: `10.1002/aenm.201401506`.

[38] X. Huang et al. “Hexagonal rare-earth manganites as promising photovoltaics and light polarizers”. *Phys. Rev. B* 92 (2015), p. 125201. DOI: `10.1103/physrevb.92.125201`.

[39] P Reinhard et al. “Review of Progress Toward 20% Efficiency Flexible CIGS Solar Cells and Manufacturing Issues of Solar Modules”. *IEEE J. Photovolt.* 3 (2013), pp. 572–580. DOI: `10.1109/jphotov.2012.2226869`.

[40] G. Cheek, F. Yang, and H. Lee. “Thin film PV: Moving at the speed of solar”. In: *2013 IEEE 39th Photovolt. Spec. Conf.* Institute of Electrical & Electronics Engineers IEEE, 2013. DOI: `10.1109/pvsc.2013.6745180`.

[41] X. Hu et al. “Ray-trace simulation of $CuInS(Se)_2$ quantum dot based luminescent solar concentrators”. *Opt. Express* 23 (2015), A858. DOI: `10.1364/oe.23.00a858`.

[42] D. S. Su and S.-H. Wei. “Transmission electron microscopy investigation and first-principles calculation of the phase stability in epitaxial $CuInS_2$ and $CuGaSe_2$ films”. *Appl. Phys. Lett.* 74 (1999), p. 2483. DOI: `10.1063/1.123014`.

[43] J Alvarez-García et al. “Polymorphism in $CuInS_2$ epilayers: Origin of additional Raman modes”. *Appl. Phys. Lett.* 80 (2002), p. 562. DOI: `10.1063/1.1435800`.

[44] D. S. Su, W Neumann, and M Giersig. “Structure modifications in chalcopyrite semiconductors”. *Thin Solid Films* 361-362 (2000), pp. 218–222. DOI: `10.1016/s0040-6090(99)00837-8`.

[45] T Hahn et al. “Order and disorder in epitaxially grown $CuInS_2$”. *Thin Solid Films* 387 (2001), pp. 83–85. DOI: `10.1016/s0040-6090(01)00790-8`.

[46] A. Moreau et al. “Impact of Cu-Au type domains in high current density $CuInS_2$ solar cells”. *Sol. Energy Mater. Sol. Cells* 139 (2015), pp. 101–107. DOI: `10.1016/j.solmat.2015.03.008`.

[47] H. Hahn et al. “Untersuchungen Öber ternäre Chalkogenide. V. öber einige ternäre Chalkogenide mit Chalkopyritstruktur”. *Z. Anorg. Allg. Chem.* 271 (1953), pp. 153–170. DOI: `10.1002/zaac.19532710307`.

[48] S.-H. Wei, S. B. Zhang, and A. Zunger. “Band structure and stability of zinc-blende-based semiconductor polytypes”. *Phys. Rev. B* 59 (1999), R2478–R2481. DOI: `10.1103/physrevb.59.r2478`.

[49] J. E. Jaffe and A. Zunger. “Anion displacements and the band-gap anomaly in ternaryABC2chalcopyrite semiconductors”. *Phys. Rev. B* 27 (1983), pp. 5176–5179. DOI: `10.1103/physrevb.27.5176`.

[50] J. E. Jaffe and A. Zunger. “Theory of the band-gap anomaly in ABC_2 chalcopyrite semiconductors”. *Phys. Rev. B* 29 (1984), pp. 1882–1906. DOI: `10.1103/physrevb.29.1882`.

[51] M. I. Alonso et al. “Optical functions and electronic structure of $CuInSe_2$, $CuGaSe_2$, $CuInS_2$, and $CuGaS_2$”. *Phys. Rev. B* 63 (2001). DOI: `10.1103/physrevb.63.075203`.

[52] C. A. Nelson, N. R. Monahan, and X. Y. Zhu. *Exceeding the Shockley-Queisser limit in solar energy conversion.* 2013. DOI: `10.1039/c3ee42098a`.

[53] H. Paul, C. David, and B. P. Rand. “Strategies for increasing the efficiency of heterojunction organic solar cells: Material selection and device architecture”. *Acc. Chem. Res.* 42 (2009), pp. 1740–1747. DOI: `10.1021/ar9000923`.

[54] D. Knig et al. "Hot carrier solar cells: Principles, materials and design". In: *Phys. E Low-Dimensional Syst. Nanostructures.* Vol. 42. 10. 2010, pp. 2862–2866. DOI: `10.1016/j.physe.2009.12.032`.

[55] M. C. Hanna and A. J. Nozik. "Solar conversion efficiency of photovoltaic and photoelectrolysis cells with carrier multiplication absorbers". *J. Appl. Phys.* 100 (2006), p. 074510. DOI: `10.1063/1.2356795`.

[56] M. Bercx et al. "First-principles analysis of the spectroscopic limited maximum efficiency of photovoltaic absorber layers for CuAu-like chalcogenides and silicon". *Phys. Chem. Chem. Phys.* 18 (2016), pp. 20542–20549. DOI: `10.1039/c6cp03468c`.

[57] M. A. Green. "Self-consistent optical parameters of intrinsic silicon at 300K including temperature coefficients". *Sol. Energy Mater. Sol. Cells* 92 (2008), pp. 1305–1310. DOI: `10.1016/j.solmat.2008.06.009`.

[58] T Trupke et al. "Temperature dependence of the radiative recombination coefficient of intrinsic crystalline silicon". *J. Appl. Phys.* 94 (2003), p. 4930. DOI: `10.1063/1.1610231`.

[59] A. Richter et al. "Improved quantitative description of Auger recombination in crystalline silicon". *Phys. Rev. B* 86 (2012), p. 165202. DOI: `10.1103/physrevb.86.165202`.

Chapter 5

Li-ion Batteries

> *"Solar and battery go together like peanut butter and jelly."*
>
> Elon Musk

Since their commercialization in 1991 by Sony, Lithium-ion batteries have become one of the most popular type of rechargeable batteries used in portable electronics and electric vehicles. Despite their success, Li-ion batteries still can benefit significantly from improvements in order to expand their use in automotive and grid storage applications. One of the main elements of the Li-ion battery with room for improvement is the cathode, where research is primarily focused on layered transition metal (TM) oxides and polyanionic materials.

In this chapter I present an overview of my work on Li-ion batteries, which has largely focused on a class of materials called Li-rich battery cathodes. The chapter starts with a brief introduction on the topic of Li-ion batteries in Section 5.1. Next, Section 5.2 introduces Li-rich materials and presents an analysis of the structure and Li-configuration (Sec. 5.2.1), the redox processes (Sec. 5.2.2) and dimer formation (Sec. 5.2.3). Section 5.3 concerns the solubility of Sn in Li_2MnO_3 (Sec. 5.3.1) as well as the influence of a local substitution of Sn, V and Mo on the stability of the oxygen framework (Sec. 5.3.2). Finally, Section 5.4 briefly introduces solid electrolytes, as well as a promising class of materials called polyborane salts (Sec. 5.4.1) and the calculation of their local cation energy landscape (Sec. 5.4.2).

5.1 Introduction

Energy storage is one of the most important topics of our contemporary society. From portable electronics to the transportation sector, the storage of electrical energy plays a vital role in many aspects of our daily lives. Moreover, it is an essential component of the transition to renewable energy, as many renewable sources of energy are intermittent, requiring the storage of excess energy to stabilize the grid. Ranging from something as simple as pumping water to a higher elevation[1] to storing superconducting magnetic energy storage, there are many methods for storing energy. Which storage solution is optimal depends largely on the application. In portable electronics and the electric vehicle (EV) market, Li-ion batteries have been used extensively as the energy storage medium of choice.

Li-ion batteries consist of two electrodes with different chemical potentials for Li^+, separated by an electrolyte (Fig. 5.1). When the battery is charged or discharged, Li^+ ions move between the two electrodes through the electrolyte, while electrons flow in the same direction via an external circuit in order to maintain charge balance. Although which electrode functions as a cathode or anode depends on whether the battery is being charged or discharged, it is the convention to stick to the standard terminology used during the charging process. That is, the cathode delivers electrons to the external circuit as the battery is charged. At the same time, Li^+ ions move to the anode, which means that for a fully charged Li-ion battery, as much Li as possible should be stored in the anode. Using this convention, the chemical potential for Li^+ should be higher for the anode than the cathode to ensure that Li^+ ions and electrons spontaneously move to the cathode when the battery is discharged.

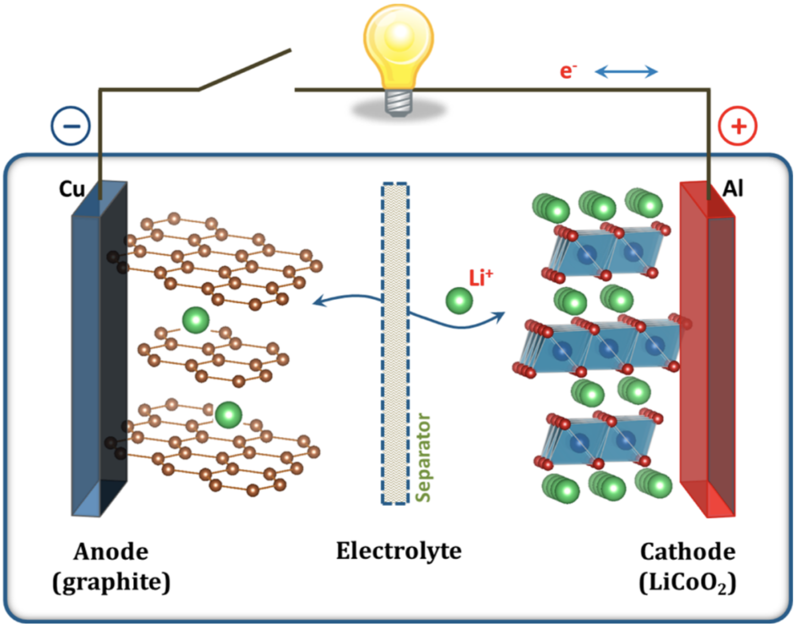


Figure 5.1: Basic composition of a Li-ion battery [2].

For the anode, most commercially available Li-ion batteries use graphite due to its low price, weight and relatively large specific capacity of 372 mAh/g [3]. Moreover, its layered structure is remarkably stable, leading to a high reversibility and cyclability. For the cathode, a more

---

[1] Although simple, this is arguably still one of the best ways to store energy for grid applications. A good example is our very own hydro storage facility at Coo, which professes to have an energy conversion efficiency of 75% [1].

diverse set of materials is being considered. The conventional layered oxides $LiMO_2$, where M is a (combination of) transition metals, is still one of the most popular chemistries due to their high energy density and rate capacity [4]. Spinel-type oxygen-based cathodes [5] (e.g. $LiMn_2O_4$) have a lower capacity compared to the layered oxides, but are also receiving a fair bit of attention because of their excellent safety. Similarly, ordered olivine compounds [6] (e.g. $LiFePO_4$) are lauded for their high safety and structural stability, but suffer from a reduced specific capacity due to their high weight. The electrolyte separating the electrodes can be either a liquid or a solid, with the liquid being the conventional choice owing to its high ionic conductivity. However, there are certain safety hazards associated with the use of liquid electrolytes, which has prompted an increased research effort for developing functional solid electrolytes (See Section 5.4).

5.2 Li-Rich Battery Cathodes

Li-ion batteries are currently the primary method of energy storage for many important applications, however many potential gains in energy density can still be made by improving the cathode capacity. Layered $LiMO_2$ compounds, where M is a transition metal, allow for fast two dimensional lithium diffusion through a divacancy mechanism [7], and high voltages versus the battery anode. Among this group, $LiCoO_2$ has long been the favored cathode in commercial applications. Cobalt is expensive and toxic, however, and $LiCoO_2$ suffers from safety problems due to its low thermal stability [8]. Moreover, the capacity of $LiCoO_2$ is limited to 130 mAh/g, because only about half of the lithium can be extracted without causing severe electrode degradation [9].

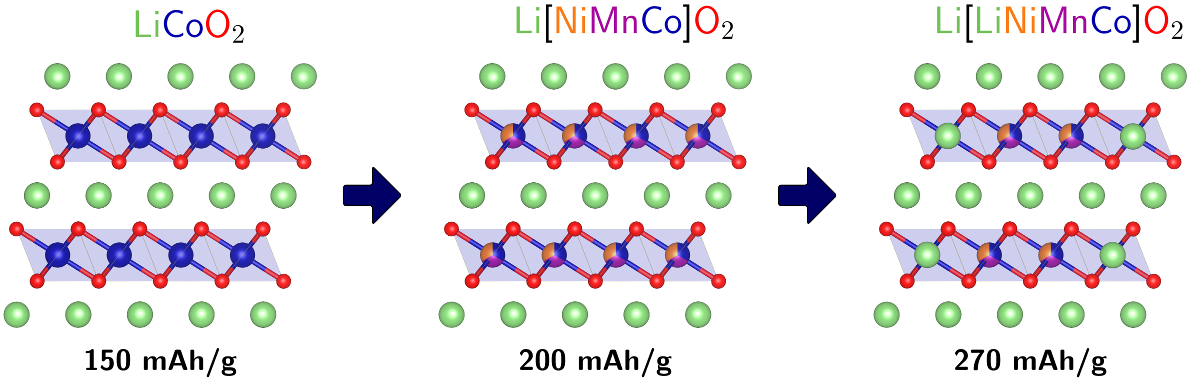


Figure 5.2: Transition from $LiCoO_2$ to NMC to Li-rich layered oxides for battery cathodes.

In order to improve upon these deficiencies, material scientists have attempted chemical substitution of Co by other transition metals such as Mn and Ni (Fig. 5.2). Since $LiNiO_2$ is deemed unsafe because of its low thermal stability [10] and $LiMnO_2$ suffers from poor electrochemical performance [11], researchers use partial substitution of Mn and Ni in $LiCoO_2$ to fine-tune the qualities of the cathode material [12]. The resulting $Li[Ni_{1-x-y}Mn_xCo_y]O_2$ (NMC) compounds show improved capacities (200 mAh/g) and safety characteristics, without significantly changing the operating voltage [13].

More recently, further explorations on layered oxide structures have led to Li-rich materials, which have an excess of Li in the material composition [14] (Fig. 5.2). These compounds can attain even higher capacities. The origin of this extra capacity is believed to be anionic reversible redox processes ($O^{2-} \rightarrow {O_2}^{2-}$) [15], which changes the fundamental minimum of transition metal content that was considered necessary in layered oxides for decades, and could lead to the next generation of high energy density Li-ion batteries. However, these materials still suffer from structural degradation as the battery is cycled, reducing the average voltage and capacity of the cell. The voltage fade is believed to be related to the migration of transition metals into the lithium layer, linked to the formation of O-O dimers with a short bond length, which in turn is driven by the presence of oxygen holes due to the participation of oxygen in the redox process. Finally, the Li-rich cathodes have also demonstrated oxygen evolution from the structure as the battery is charged, which is detrimental for the safety of battery.

This section presents an investigation into the connection between oxygen redox and the stability of the oxygen framework for Li-rich materials, based on Li_2MnO_3 and Li_2IrO_3. These two Li-rich cathode materials have demonstrated significantly different cycling properties. Li_2MnO_3, a well studied Li-rich material, suffers from a substantial amount of voltage fade as the battery is cycled [16], whereas Li_2IrO_3 does not [17]. Studying the differences in oxidation and structural stability between these two compounds can offer insight as to why their cycling properties are so different.

5.2.1 Structure and Li configuration

In order to compare the structural stability of the oxygen framework for the Li_2MnO_3 and Li_2IrO_3 compounds, we have to calculate the chemical reaction energy for the formation of O-O dimers for both cathode materials in a charged state, i.e. after the removal of a certain fraction of lithium. However, the fully charged structure for Li_2MnO_3 is found to be highly unstable, i.e. lead to the spontaneous formation of several oxygen dimers, especially when any local changes to the structure are made. Moreover, the cathode is unlikely to ever be fully delithiated in a practical battery, rendering an investigation of the fully charged state less relevant.

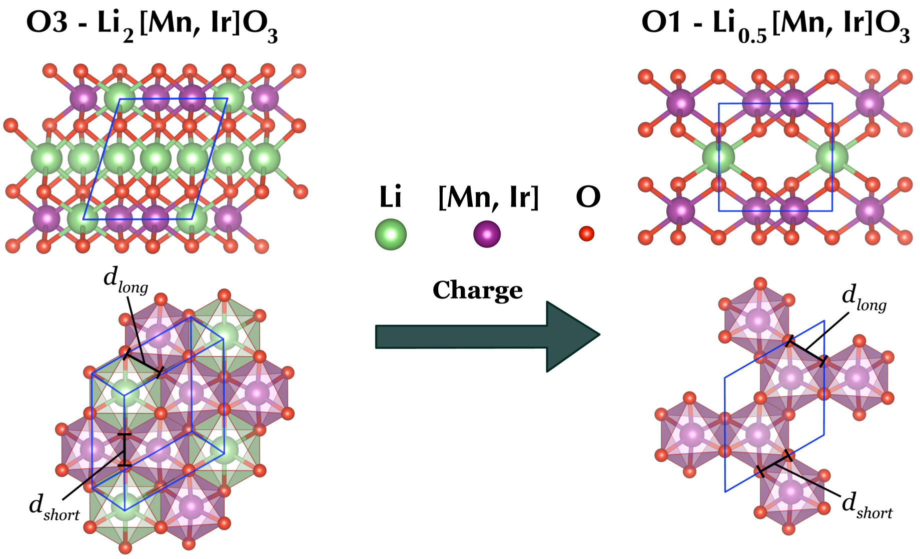


Figure 5.3: Transformation from the O3 to O1 stacking for both Li_2MnO_3 and Li_2IrO_3 as the cathode is charged. The primitive unit cell is drawn in blue. The top figures represent the structure shown in the [100] projection, whereas the lower figures represent a single octahedral layer of the layered structure, viewed top down. d_{short} and d_{long} both represent O-O distances across the Li/TM layer, corresponding to an octahedral edge bordering two TM's and a TM and Li/Vacancy, respectively.

Hence, the structural stability should preferably be studied in a partially charged cathode material. This requires knowledge about the location of the lithium for each state of charge, as there are many possible Li-Vacancy configurations to consider. We investigate the lithium configuration for Li_2MnO_3 by calculating the energy of all symmetrically non-equivalent configurations in the conventional unit cell. This is done based on the workflow described in Section 3.3.2, resulting in 94 non-equivalent configurations. Similar to previous work [18], we find that for several lithium configurations, the Li_2MnO_3 structure spontaneously shifts from an O3 stacking to the O1 stacking[2] at higher charge state of the battery (see Fig. 5.3). In order to verify this transition from the O3 to the O1 stacking, we once again use the configuration workflow to calculate the energy of all Li configurations in unit cells up to two times the size of the conventional unit cell of the O1 stacking, which results in 220 non-equivalent configurations.

For the O1 stacking, we find that several configurations at lower states of charge switch to the O3 stacking, i.e. the opposite transformation occurs compared to that at higher states of charge. To be able to compare the energies of the O1 and O3 stackings fairly, we remove the configurations which change stacking. As the stacking of the oxygen octahedra is closely connected to the angles between the lattice vectors, we remove all configurations for which any lattice angle has changed more than 12°. This leaves 84 and 181 configurations for the O3 and O1 stacking, respectively. Finally, we calculate the formation energy for all configurations versus the fully charged and discharged state of the O3 stacking:

$$E_f(x) = E(\text{Li}_x\text{MnO}_3) - (1 - \frac{x}{2})E(\text{O3-MnO}_3) - \frac{x}{2}E(\text{O3-Li}_2\text{MnO}_3) \tag{5.1}$$

The corresponding formation energies are plotted for both stackings in Fig. 5.4. It is clear that as Li is removed from Li_2MnO_3, the O1 stacking becomes thermodynamically favorable.

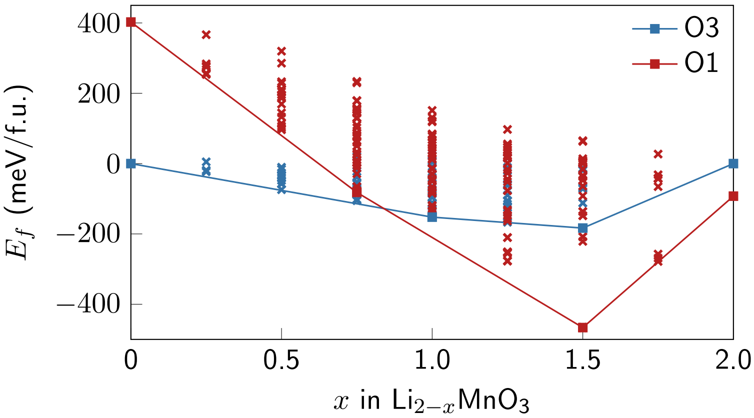


Figure 5.4: Formation energies of all configurations in both the O3 and O1 stacking of Li_xMnO_3. Each mark represents the formation energy of one non-equivalent Li configuration. The full lines correspond to the convex hull of the corresponding stacking.

A similar transformation from the O3 to O1 stacking is found to occur for $Li_{0.5}IrO_3$, which has been experimentally verified and leveraged in order to study the deformation of the oxygen framework by McCalla et al. [17]. This means that structures of the discharged and charged

---

[2] The various stackings of layered oxides were first classified by Delmas et al. [19]. O refers to the octahedral coordination of O atoms around the alkali (Li, Na, ...) ions. The number is related to the stacking of the O atoms, i.e. for O3 the stacking is AB CA BC, so after 3 layers of alkali ions, the oxygen environment returns to its original stacking. For O1, the stacking is AB AB, so the stacking for each alkali layer is the same.

structures are the same, save for a difference in the lattice parameters, which facilitates the comparison of the changes in geometry and oxidation state between Li_2MnO_3 and Li_2IrO_3. We choose to focus on the 75% charged structures for several reasons. First, the optimal lithium configuration for $Li_{0.5}MnO_3$ is on the convex hull of the O1 stacking, indicating that this structure is quite stable and hence easier to work with once we start introducing O-O dimers (Sec. 5.2.3). Second, oxygen gas is only released from the Li_2IrO_3 cathode once it is charged beyond 75% [17], whereas Li_2MnO_3 has already lost oxygen at this state of charge [20]. Hence, studying the stability for this lithium content is most interesting, as it may show discrepancies between the stability of the oxygen frameworks. Finally, the oxygen framework of $Li_{0.5}IrO_3$ was studied by McCalla et al. [17], which allows for a direct comparison of our calculated O-O distances with experiment.

Table 5.1: O-O distances for the discharged and charged $Li_2[Mn, Ir]O_3$ structures, all expressed in Å. The distances for Li_2IrO_3 are compared with the neutron powder diffraction results of McCalla et al. [17].

		$Li_2[Mn, Ir]O_3$		$Li_{0.5}[Mn, Ir]O_3$	
		DFT	Neutron	DFT	Neutron
Mn	d_{short}	2.52	-	2.31	-
	d_{long}	2.75	-	2.62	-
Ir	d_{short}	2.75	2.77	2.51	2.45
	d_{long}	2.87	2.84	2.74	2.73

As was noted previously by McCalla et al. [17], the oxygen framework is distorted as lithium is removed from the structure. In order to quantify this, we calculate the distances between the various oxygen pairs, connected in the tetrahedral environment of Mn or Ir. The distances between oxygen pairs which are part of the same oxygen layer change little. However, for the interlayer oxygen pairs, denoted as d_{short} and d_{long} in Fig. 5.3, the change in bond length is more pronounced (Table 5.1). Moreover, the shorter bonds for an oxygen pair sharing two [Mn, Ir] neighbors, shrink more than the long bonds, which share a transition metal and Li or vacancy. This leads to a distortion of the octahedral environment around the transition metals, resulting in short O-O bonds which McCalla et al. refer to as dimers. In Section 5.2.3, we will return to this topic, focusing our attention on the formation of a peroxo species with a bond length closer to that of the oxygen molecule, as this formation has been derived theoretically for Li_2MnO_3, and believed to be related to the migration of Mn and the resulting voltage fade [21].

5.2.2 Oxidation

Sathiya et al. [15] have discussed that removing lithium from Li-rich cathodes leads to the formation of holes on the oxygen, i.e. the oxidation of oxygen. Seo et al. [22] have proposed that the formation of localized holes relies on the presence of labile oxygen states, which are found for oxygen with Li atoms on opposite sites of its octahedral environment. Moreover, they explain that because of the honeycomb structure of Li_2MnO_3, all oxygen environments have such a Li-O-Li configuration, leading to a high participation of oxygen in the redox processes. Although Li_2IrO_3 has a similar structure, Hong et al. [23] assert that because Ir^{4+} can be more

Table 5.2: Calculated absolute values of the magnetic moments for the discharged and charged $Li_2[Mn, Ir]O_3$ structures, all expressed in Bohr magnetons μ_B. Note that for the Ir structures, non-collinear calculations were performed to include spin-orbit coupling, and the norm of the local magnetization vector was calculated in order to express the local magnetic moment as a scalar.

		$Li_2[Mn, Ir]O_3$	$Li_{0.5}[Mn, Ir]O_3$
Mn	$\|\mu\|$ (Mn)	2.918	2.949
	$\|\mu\|$ (O)	0.001	0.445
Ir	$\|\boldsymbol{\mu}\|$ (Ir)	0.374	1.025
	$\|\boldsymbol{\mu}\|$ (O)	0.025	0.314

easily oxidized, these labile oxygen states are not depleted to the same extent, stabilizing the oxygen framework.

To study the change in the oxidation state of the atoms, we compare the calculated local magnetic moments of the various elements for both structures in the discharged and 75% charged state in Table 5.2. Note that as Ir exhibits strong spin-orbit coupling effects, non-collinear calculations were performed for both Li_2IrO_3 and $Li_{0.5}IrO_3$. We can see that for Li_2MnO_3, the magnetic moment of Mn remains largely the same, whereas the magnetic moment on oxygen increases significantly. For oxygen, the increase in magnetic moment corresponds to an oxidation from its O^{2-} state, as its p-orbitals are no longer fully occupied, leading to an increased local density of unpaired electrons. The results for Mn indicate that it does not participate much in the redox processes that occur when the battery is charged. Instead, its magnetic moment remains close to 3 μ_B, which corresponds to the initial oxidation state Mn^{4+} in the discharged cathode structure. For Li_2IrO_3, removing lithium from the discharged structure results in a significant change of the local magnetic moment for both Ir and O, implying a more mixed redox process during the charging of the cathode. This mixed redox for Li_2IrO_3 is in agreement with the X-ray photoelectron spectroscopy results of McCalla et al. [17], where they explain that this is in part due to the covalent character of the Ir-O bond. This covalency could also explain why oxygen is oxidized less when charging Li_2IrO_3, as valence electrons are removed from both Ir and O. Mn^{4+} could in principle also be oxidized further, but the Mn^{5+} oxidation state is rare, and generally not octahedrally coordinated [24]. The fact that the change in magnetic moment on O is smaller for Li_2IrO_3, than for Li_2MnO_3 indicates that the mixed redox process in the charging of Li_2IrO_3 results in a lower state of oxidation for the oxygen of the structure in the charged state.

This conclusion is supported by the projected density of states (PDOS), plotted in Fig. 5.5. For Li_2MnO_3, the electronic states close to the Fermi level correspond largely to the O-2p states, which indicates that as the battery is charged, electrons are removed from oxygen rather than Mn. In contrast, looking at the PDOS for Li_2IrO_3 reveals that the states near the Fermi level are more evenly distributed between O-2p and Ir-5d, which corresponds well to the picture of a more mixed redox activity for this material. When comparing the PDOS of the charged structures with the discharged ones, we note that in both cases the number of O-2p states near the Fermi level has decreased. The difference is much more substantial for Li_2MnO_3 than for Li_2IrO_3, once again implying a larger oxidation of oxygen for Li_2MnO_3.

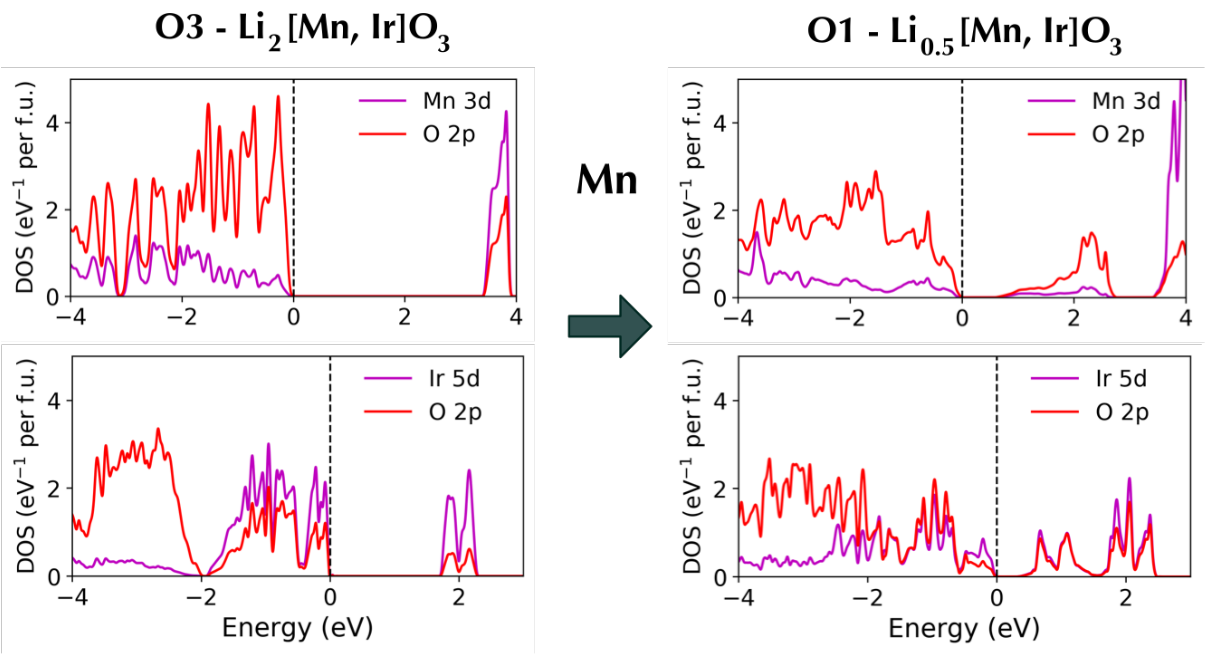


Figure 5.5: The projected density of states of the Mn-3d, Ir-5d and O-2p orbitals for the discharged (left) and charged (right) structures, where we have aligned the Fermi level to zero. In order to allow for a reasonable comparison between the pristine and charged state, we have consistently plotted the number of states per electronvolt per formula unit.

5.2.3 Dimer Analysis

Once the oxygen atoms develop holes on their p-orbitals, they can be subsequently stabilized by a reorganization of the oxygen framework, forming a peroxo-like species of oxygen pairs with shortened O-O bonds. McCalla et al. [17] were able to demonstrate the shortening of such bonds for Li_2IrO_3, which they referred to as an O-O peroxo-like dimer. More recently, other authors [21, 25] have asserted that the presence of unstable holes on the oxygen can also lead to the formation of a true oxygen dimer, finding O-O bonds with distances closer to that of molecular oxygen (1.3-1.5 Å). Such short O-O distances have also been reported recently by Li et al. [26], who found shifts in their Raman spectra that correspond to similar bond lengths. Both Saubanière et al. [25] and Chen and Islam [21] discuss that the dimerization of oxygen can trigger the migration of Mn in fully charged Li_2MnO_3, which is considered to be the mechanism by which the structure transforms into a spinel-type phase. This structural change results in a reduced average voltage, which is detrimental for the energy density of the battery [9]. Moreover, Chen and Islam contend that the O-O dimer is eventually released from the structure as O_2. Such oxygen evolution has been observed for several Li-rich materials [27, 28].

So far the study of dimer formation in Li_2MnO_3 has been limited to the O3 stacking and the fully charged structure. However, as we have seen in Section 5.2.1, both Li_2MnO_3 and Li_2IrO_3 are believed to transform into an O1 stacking, which changes the possible migration pathways for the transition metal. Here, we compare the stability of the oxygen framework of Li-rich Li_2MnO_3 and Li_2IrO_3 by calculating the thermodynamic driving force of the dimer formation, as well as the kinetic barrier. The goal is to check if there is a connection between the formation

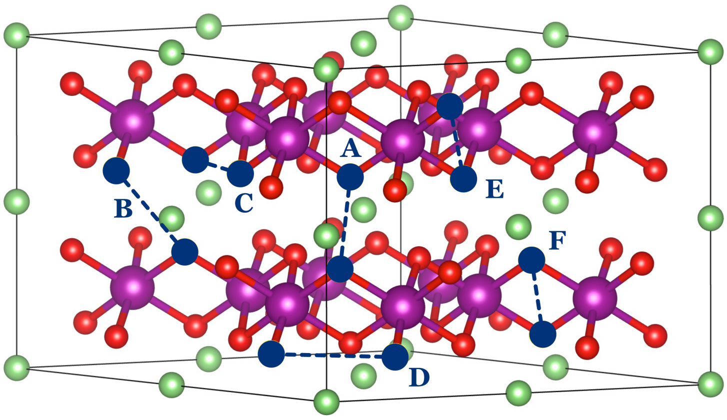


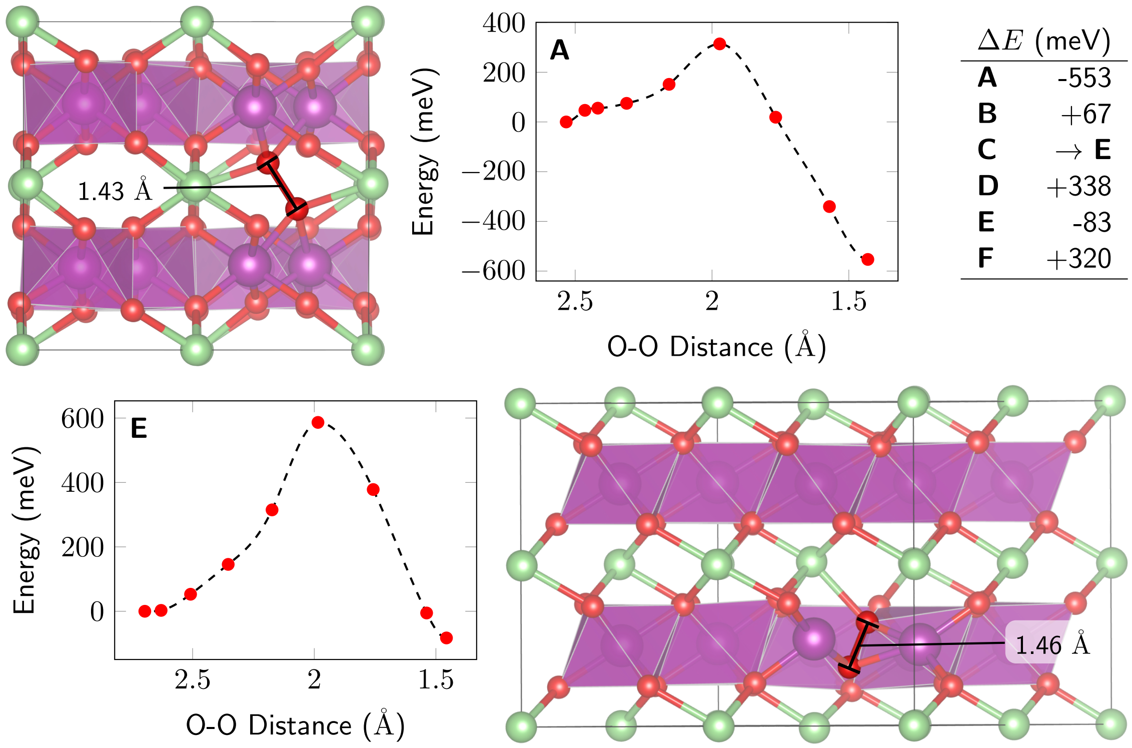
Figure 5.6: Potential oxygen dimers in O1-$Li_{0.5}[Mn, Ir]O_3$.

of dimers and the oxidation of oxygen. Instead of investigating the fully charged structures, we apply our methodology to 75% delithiated $Li_{0.5}MnO_3$ and $Li_{0.5}IrO_3$.

To calculate the chemical reaction energy of the dimer formation, we construct a 2×2×2 supercell of the primitive unit cell, for the O1 stacking of 75% charged $Li_{0.5}MnO_3$ and $Li_{0.5}IrO_3$ (Fig. 5.6). In order to rigorously study the dimer formation, we need to consider all non-equivalent oxygen pairs that have the potential to form a dimer for each material. We use the workflow described in Section 3.3.3 to calculate the reaction energy of all non-equivalent potential dimers in the structure. In short, all potential oxygen dimers in the structure are found using a voronoi decomposition to find the neighbors of the various atoms in the unit cell. Once all oxygen dimers have been found, we set up a list of all non-equivalent potential dimers based on the symmetry operations of the structure. In the charged O1-$Li_{0.5}MnO_3$ and O1-$Li_{0.5}IrO_3$ structures, we find a total of 6 non-equivalent dimers, shown in Fig. 5.6. For each structure and each potential dimer, we reduce the distance between the oxygen atoms in the dimer pair to 1.4 Å and once again optimize all atomic positions as described in the methods section. To make sure the interaction between the dimers in the periodic boundary conditions approach of VASP is sufficiently small, we have also performed similar calculations in a 3×3×3 supercell, and found the differences between the reaction energies to be smaller than 50 meV for all potential dimers [29].

Figure 5.7 shows the results for the reaction energy and final bond length. Even though all perturbed structures produce a stable oxygen dimer after optimization, only two dimers have a negative reaction energy, which are labeled as **A** and **E** in Fig. 5.6. One dimer (**C**) results in a geometry similar to the formation of the **E** dimer, with comparable energies. The two dimers that have a reduced energy in the final state are formed by two oxygen atoms from different layers, with dimer **A** being the most energetically favorable by far. The final structures of both dimers are shown in Fig. 5.7, along with the kinetic barrier, calculated using the NEB method. For the **A** dimer, the kinetic barrier is equal to 314 meV, which is smaller than the typical kinetic barrier for lithium migration in layered structures [30]. This implies that the formation of the **A** dimer is very likely to occur during the charging process. The kinetic barrier for the **E** dimer is significantly higher at 586 meV, but is by no means insurmountable. Hence, we would expect to find both dimers to play a significant role in the structural changes that occur for Li_2MnO_3 as it is cycled.

In stark contrast with the results of O1-$Li_{0.5}MnO_3$, none of the dimer optimizations for O1-$Li_{0.5}IrO_3$ result in a new geometry with a lower energy as the unperturbed structure. In fact, out of all the dimers, all but one return to the original oxygen framework. The only dimer that is stable after optimization has an increased energy of +2.2 eV, and is hence unlikely



	ΔE (meV)
A	-553
B	+67
C	→ **E**
D	+338
E	-83
F	+320

Figure 5.7: Reaction energies ΔE, final O-O bond length d_{O-O} and kinetic barriers for the dimers in O1-$Li_{0.5}MnO_3$. The final geometry and kinetic barrier is only shown for dimer **A** and **E**, which have a negative reaction energy.

to ever be formed in practice. In our view, the enhanced stability of the oxygen framework can be in part explained by the reduced participation of oxygen in the redox process as the battery is charged, which is believed to be the primary driver for dimer formation.

5.3 Substitutions

Based on the discussion from Section 5.2, it is clear that the stability of the oxygen lattice plays an important role in maintaining the structure of Li-rich cathodes as the battery is cycled. One suggested strategy for stabilizing the oxygen framework is the (partial) substitution of Mn^{4+} by other elements. Based on our results, as well as the results of McCalla et al. [17], Ir^{4+} seems like a natural choice, as this leads to an improved stability of the structure and hence better cycling properties. However, designing a Ir-based cathode material is not a practical approach due to the weight and price of Ir.

Another element that has been suggested in order to improve the cycling behavior of Li-rich materials is Sn^{4+}, for several reasons. First, the bonding energy of $Sn-O$ is higher than that of $Mn-O$, which can enhance the structural stability [31]. Second, Sn^{4+} does not tend to adopt a tetrahedral coordination [15] and has a much larger ionic radius compared to Mn^{4+}, implying that the migration of Sn^{4+} to the Li layers is less likely to occur as the battery is charged[3]. For these reasons, Sn^{4+} was chosen as the first element by our experimental collaborators from Hasselt University[4] to substitute in Li-rich NMC samples in an attempt to improve the cycling properties of the cathode.

In this section, I start by briefly discussing their powder X-ray diffraction (PXRD) results, as well as the energy-dispersive X-ray analysis (EDX) results of our collaborators within EMAT[5], as a motivation for the calculations we have performed on the solubility of Sn in a Li-rich/Mn-rich material. Next, I perform a similar analysis as in the previous section for investigating the influence of Sn-substitution on the stability of the oxygen framework. Finally, I extend this analysis to the substitution of several other elements that have the potential to oxidize further (Co, V and Mo), in order to investigate if they have a stabilizing effect similar to that of Ir.

5.3.1 Thermodynamic Stability of Sn substitution

Although the increased ionic radius of Sn^{4+} compared to Mn^{4+} is believed to inhibit its migration into the Li-layer, it can potentially also lead to a reduced solubility of Sn in the NMC structure. In order to investigate the solubility of Sn in Li-rich NMC, several samples with increased Sn-substitution were prepared for a Li-rich/Mn-rich NMC structure, leading to stoichiometries $Li_{1.2}Ni_{0.13}Co_{0.13}Mn_{0.54-x}Sn_xO_2$ for $x \approx 0, 0.027, 0.054, 0.108$ and 0.54. Looking at the change in the PXRD pattern of Sn-substituted NMC samples in Fig. 5.8a, a second phase appears as the amount of Sn is increased, starting from $x = 0.054$. This phase becomes dominant for $x = 0.54$ and was identified as Li_2SnO_3 based on an indexation of the PXRD results. The EDX results for a particle with $x = 0.108$ (Fig. 5.8b) indicate that the exsolution of this Sn-rich phase already occurs at lower levels of Sn-substitution.

To determine the Sn solubility computationally, we investigate the thermodynamic stability of Sn-substituted structures versus their decomposition into an Mn-rich and Sn-rich phase. However, some approximations must be made in order to make the problem computationally feasible. First, including four different elements (Ni, Mn, Co and Sn) in our calculations would make both the number of configurations and reaction products prohibitively large, so we limit our study to the solubility of Sn in $Li_{1.2}Mn_{0.8}O_2$. Considering the similar ionic radii of Mn^{4+}, Co^{4+} and Ni^{4+}, this should not affect our conclusion significantly. Using this approximation, we

---

[3] Note that the migration path for Sn^{4+} does not have to pass a tetrahedral site in case the stacking is changed from O3 to O1.

[4] Andreas Paulus, Marlies van Bael and An Hardy, Inorganic and Physical Chemistry, Hasselt University.

[5] Mylène Hendrickx, Olesia Karakulina, Artem Abakumov and Joke Hadermann, Electron Microscopy for Materials Science, Antwerp University.

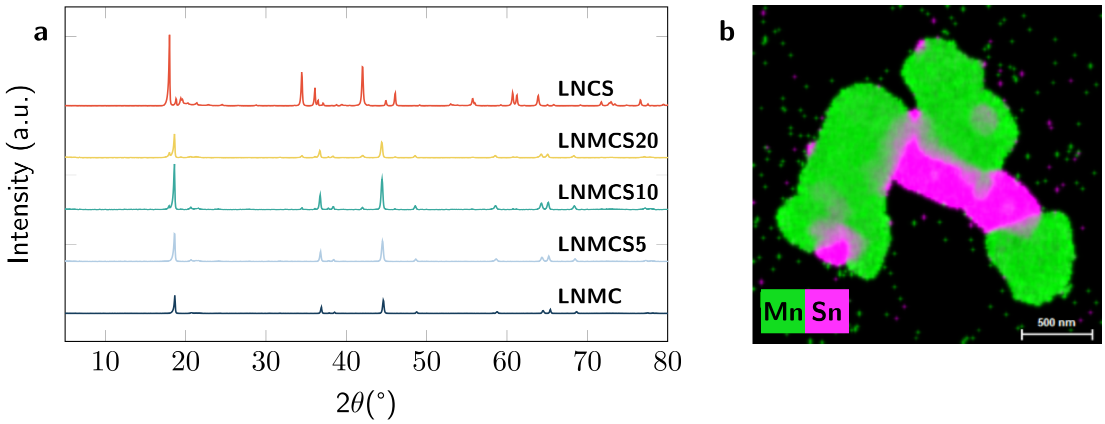


Figure 5.8: (**a**) PXRD patterns of $Li_{1.2}Ni_{0.13}Co_{0.13}Mn_{0.54-x}Sn_xO_2$ for increasing values of Sn-substitution x. (**b**) Mixed (Mn, Sn) elemental EDX map of representative LNMCS20 (x=0.108) particles, which demonstrates the presence of two different regions, a Mn-rich and Sn-rich one. Courtesy of Andreas Paulus (**a**) and Mylène Hendrickx (**b**).

first consider the following exsolution reaction:

$$\mathrm{Li_{1.2}Mn_{0.8-\mathit{x}}Sn_{\mathit{x}}O_2} \rightarrow \left(1 - \frac{x}{0.8}\right) \mathrm{Li_{1.2}Mn_{0.8}O_2} + \frac{x}{0.8} \mathrm{Li_{1.2}Sn_{0.8}O_2} \tag{5.2}$$

The PXRD results indicate, however, that the final Sn-rich phase most likely corresponds to Li_2SnO_3 and, to a lesser extent, SnO. This means that the following exsolution reaction is more probable:

$$\mathrm{Li_{1.2}Mn_{0.8-\mathit{x}}Sn_{\mathit{x}}O_2} \rightarrow \left(1 - \frac{x}{0.8}\right) \mathrm{Li_{1.2}Mn_{0.8}O_2} + \frac{x}{0.8} \left[0.6 \cdot \mathrm{Li_2SnO_3} + 0.2 \cdot \mathrm{SnO}\right] \tag{5.3}$$

Finally, there is another restriction to which our structures should adhere: based on the HAADF-STEM results for several of the Sn-substituted samples, the honeycomb ordering of the Li-TM/Sn cations is largely maintained[6]. This leads to the following construction of the $Li_{1.2}Mn_{0.8}O_2$ structures:

- Start from the pristine primitive structure of O3-Li_2MnO_3, with space group $C2/m$.
- Replace the Li in the Li/Mn layer by a placeholder element, e.g. Lr. This is just to keep track of which sites correspond to the Li sites in the honeycomb layer.
- Make a 2×2×2 supercell. This size is chosen to get as close as possible to the experimental composition of Li, without having to consider a unit cell size that is prohibitively large.
- Use the `Cathode.get_cation_configurations()` method to generate honeycomb-like structures by substituting the Lr by Li and Mn, restricting the Li concentration of the final configurations to closely match that of the experimental samples.

Because of the restrictions of the honeycomb pattern and Li concentration, this method only results in 5 configurations, each with a composition that closely matches the experimental one: $Li_{1.2083}Mn_{0.7917}O_2 \approx Li_{1.21}Mn_{0.79}O_2$. To generate the Sn-substituted structures, we consider each of the Mn configurations, and once again use the `Cathode.get_cation_configurations()`

---

[6] Note that as the ratio of Li over TM/Sn elements is smaller than 2, it is no longer possible to have a complete honeycomb ordering in the Li-TM/Sn layer, as was the case for Li_2MnO_3.

method, this time partially substituting the Mn elements by Sn. For a single Sn substitution in the supercell ($x = 0.042$), there are 47 possible Li-Mn-Sn, which is still a manageable amount to handle with our configuration workflow. However, increasing the Sn content further leads to 361/1867/7202 configurations for $x = 0.083/0.125/0.167$ respectively. Optimizing the geometry of all these configurations is clearly not possible, so we have to limit the number of configurations to a more manageable number.

Before analyzing the reaction in Eq. 5.3, it is important to confirm that the Sn-rich phase is more likely to be a combination of Li_2SnO_3 and SnO than $Li_{1.2}Sn_{0.8}O_2$. For this purpose, we first generate all Li-Sn configurations of $Li_{1.2}Sn_{0.8}O_2$, similar to the procedure described in the previous paragraph. Next, we optimize the geometry and calculate the energy of all 5 configurations, along with the energies of Li_2SnO_3 and SnO. Note that, similar to the Li-rich Mn structure, the composition of the configurations is closer to $Li_{1.2083}Sn_{0.7917}O_2$.This means that the effective decomposition reaction of interest is:

$$\mathrm{Li}_{1.2083}\mathrm{Sn}_{0.7917}\mathrm{O}_2 \rightarrow A \cdot \mathrm{Li}_2\mathrm{SnO}_3 + B \cdot \mathrm{SnO}, \tag{5.4}$$

where $A \approx 0.604$ and $B = 0.1875$. The corresponding formation energy is:

$$E_f = E(\mathrm{Li}_{1.2083}\mathrm{Sn}_{0.7917}\mathrm{O}_2) - A \cdot E(\mathrm{Li}_2\mathrm{SnO}_3) - B \cdot E(\mathrm{SnO}), \tag{5.5}$$

which for the lowest energy configuration of $Li_{1.2083}Sn_{0.7917}O_2$ is equal to 459 meV. Considering the significantly larger energy of $Li_{1.2083}Sn_{0.7917}O_2$ compared to Li_2SnO_3 and SnO, it is reasonable to suggest that the Sn-rich phase corresponds more closely to a combination of these end products.

Finally, the exsolution reaction becomes:

$$\mathrm{Li}_{1.21}\mathrm{Mn}_{0.79-x}\mathrm{Sn}_x\mathrm{O}_2 \rightarrow \left(1 - \frac{x}{0.79}\right) \mathrm{Li}_{1.21}\mathrm{Mn}_{0.79}\mathrm{O}_2 + \frac{x}{0.79}\left[0.604 \cdot \mathrm{Li}_2\mathrm{SnO}_3 + 0.1875 \cdot \mathrm{SnO}\right], \tag{5.6}$$

with formation energy

$$\begin{aligned} E_f(x) = E(\mathrm{Li}_{1.21}\mathrm{Mn}_{0.79-x}\mathrm{Sn}_x\mathrm{O}_2) - \left(1 - \frac{x}{0.79}\right) E(\mathrm{Li}_{1.21}\mathrm{Mn}_{0.79}\mathrm{O}_2) \\ - \frac{x}{0.79}\left[0.604 \cdot E(\mathrm{Li}_2\mathrm{SnO}_3) + 0.1875 \cdot E(\mathrm{SnO})\right], \end{aligned} \tag{5.7}$$

Figure 5.9 shows the calculated formation energies for 40 configurations of the Sn substituted structures for each $x = \{i/24 | i = 1, 2, 3, 4, 5\}$, compared to their decomposition in $Li_{1.21}Mn_{0.79}O_2$, Li_2SnO_3 and SnO. For the lowest Sn concentration, x = 0.042, the formation energy of the lowest energy configuration is only +6.5 meV/atom. This structure can reasonably be considered as metastable [32] and as such the formation of a single phase is feasible at low Sn concentrations. However, as the Sn-concentration x is increased, the $Li_{1.2}Mn_{0.8-x}Sn_xO_2$ configurations become more unstable, increasing the likelihood of a decomposition in $Li_{1.2}Mn_{0.8}O_2$, Li_2SnO_3 and SnO phases, as observed in the PXRD results for the high Sn concentration samples. Note that if the Sn substituted orderings are generated randomly, i.e. without respecting the honeycomb pattern, the energies are significantly higher compared to the honeycomb structures at each Sn concentration. This matches the preservation of the honeycomb ordering for the Sn substituted structure found for the HAADF-STEM results.

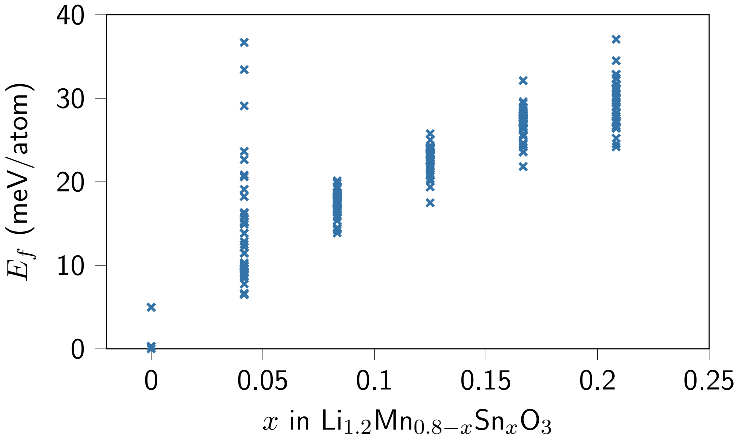


Figure 5.9: Calculated formation energies using Eq. (5.7) for all configurations with $x = \{i/24 | i = 1, 2, 3, 4, 5\}$.

5.3.2 Influence of Mn^{4+} substitution on oxygen stability

The next question is whether the stability of the oxygen framework of Li_2MnO_3 can be improved by a local substitution of Mn^{4+} by Sn^{4+}. In order to make a fair comparison with the results presented in Sections 5.2.2 and 5.2.3, we start from the 2×2×2 supercell of the charged O1-$Li_{0.5}MnO_3$ structure. Based on the results of the previous section, only a limited amount of Sn can be substituted in Li_2MnO_3 before we expect the cathode to separate in several phases, and hence we simply substitute a single Mn atom by Sn. In light of the discussion of Section 5.2, however, we would not expect Sn to be very effective in stabilizing the oxygen framework, as it is unable to oxidize beyond +4. Hence, we expand our search of suitable substitutions to V and Mo, two elements that permit higher states of oxidation and have shown promise in Li-rich materials [33, 34]. Moreover, in order to study the influence of the exchange-correlation functional, we make a comparison between the PBE+U [35, 36] results and the recently introduced SCAN [37]. In contrast to PBE+U, SCAN does not rely on an element-dependent parameter, so it would be interesting to see if it produces a similar trend for the barrier of the various substituted elements. This discussion, however, is left for the end of this section.

Table 5.3 contains the magnetic moments and Fig. 5.10 shows the projected density of states near the Fermi level for the Sn/V/Mo-substituted structures, both in the discharged and charged state. For Sn, there are practically no states in near the Fermi level, which is not surprising considering that in a +4 oxidation state, Sn has donated its 5s and 5p valence electrons to the surrounding oxygen. This is also clear from the magnetic moments, which are close to zero for Sn in both states of charge. Because of this inability of Sn to oxidize further, the oxygen redox is similar to that of undoped Li_2MnO_3. In light of this, it is unsurprising that the kinetic barrier for the A dimer in the 75% charged Sn-doped structure is similar, even when one of the oxygen atoms neighbors the substituted Sn (Fig. 5.11).

Table 5.3: Calculated absolute values of the magnetic moments for the discharged and charged for the Sn/V/Mo-substituted structures, all expressed in Bohr magnetons μ_B. For the oxygen, we make a distinction between the neighbors of the substituted element O_n and other oxygen elements in the unit cell O_o.

		PBE+U		SCAN	
		discharged	charged	discharged	charged
Sn	$\|\mu\|$ (Sn)	0.018	0.045	0.017	0.064
	$\|\mu\|$ (O_n)	0.021	0.434	0.020	0.355
	$\|\mu\|$ (O_o)	0.000	0.463	0.000	0.333
V	$\|\mu\|$ (V)	0.953	0.270	0.875	0.260
	$\|\mu\|$ (O_n)	0.012	0.329	0.032	0.219
	$\|\mu\|$ (O_o)	0.000	0.459	0.000	0.322
Mo	$\|\mu\|$ (Mo)	1.792	0.172	1.293	0.110
	$\|\mu\|$ (O_n)	0.028	0.208	0.001	0.128
	$\|\mu\|$ (O_o)	0.001	0.446	0.000	0.310

For V and Mo, the magnetic moments and projected density of states are also in the line of expectations. Both substituted elements show a clear change in their magnetic moment, indicating that they have oxidized further as the battery is charged. The neighboring oxygens O_n also have a significantly lower magnetic moments compared to other oxygen atoms O_o for the charged structure, confirming the decreased oxidation of the oxygen framework around the substituted element. Looking at the projected density of states in Fig. 5.10, the V and Mo states are both right below the Fermi level, basically corresponding to donor levels in the band gap of the discharged Li_2MnO_3 structure. As Li is removed from the structure, these states will be the first to be depopulated, which matches well with the picture provided by the magnetic moments.

However, in contrast to Ir, the increased propensity of V and Mo to oxidize does not seem to increase the stability of the surrounding oxygen framework. The kinetic barriers for the formation of the **A** dimer, shown in Fig. 5.11, is easily surmountable for both elements, and is in fact even lower compared to that of Li_2MnO_3. Considering this, it would appear that neither the substitution of V nor Mo improves the stability of the oxygen framework, despite the fact that they are more likely to oxidize before the oxygen.

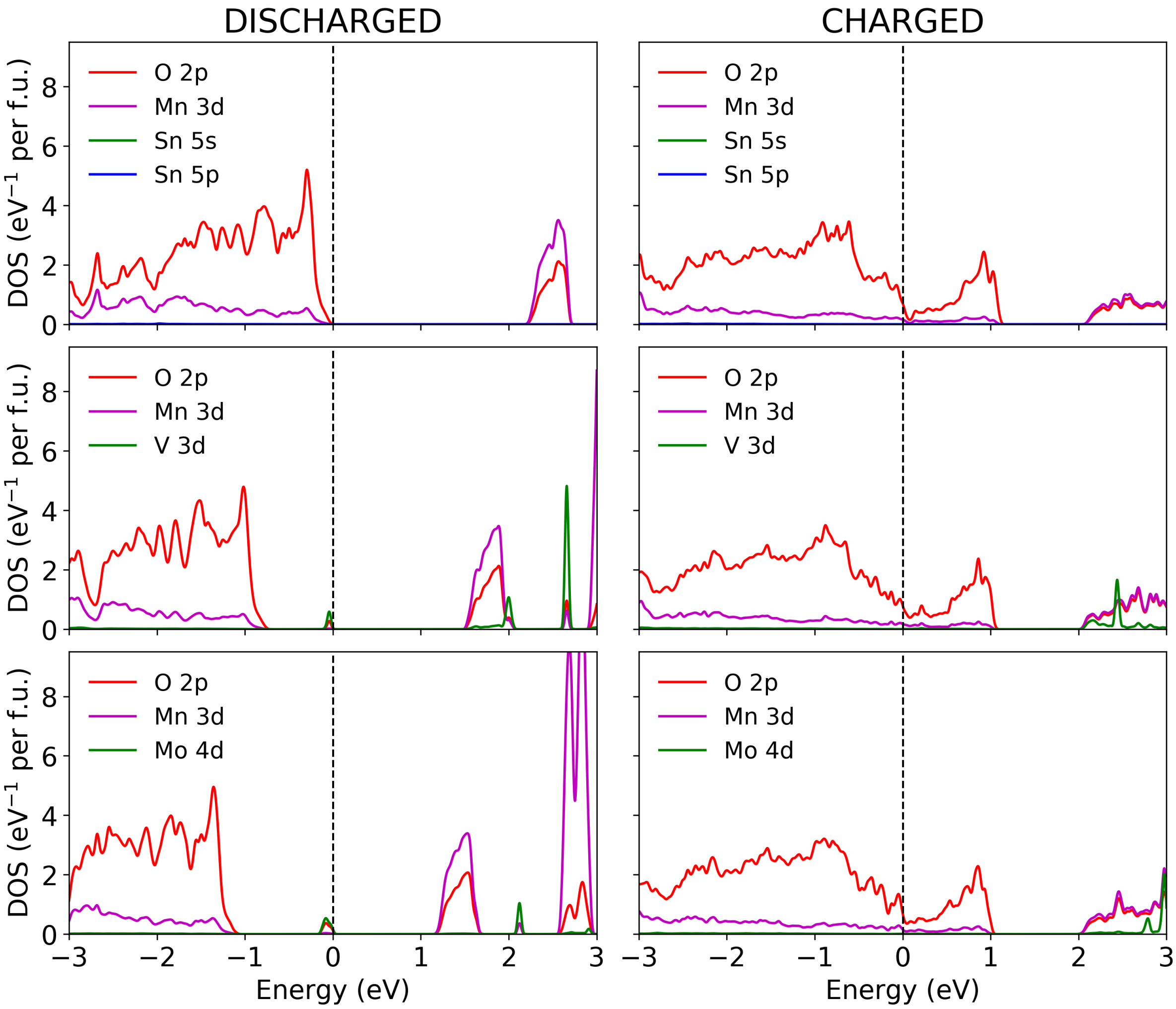


Figure 5.10: Projected density of states for the discharged (left) and charged (right) structures, for the 2×2×2 supercell of Li_2MnO_3 with a single substitution of Sn, V and Mo.

Finally, we have also performed all of the calculations in this section with the recently published SCAN [37], in order to compare the results with those of PBE+U [35, 36]. As SCAN does not rely on specifying a parameter for each transition metal that can influence the oxidation state of the element, this provides an unbiased set of data to compare with. Looking at the magnetic moments in Table 5.3, the discussion from the previous paragraphs remains largely intact. However, the magnetic moments on the oxygen atoms is decidedly lower compared to the PBE+U results. One could argue that this indicates that our choice of Hubbard-U correction might have excessively localized the electrons around the transition metals, but a similar reduction in magnetic moments is found for the oxygen neighboring Sn, for which we have applied no Hubbard-U correction. Moreover, our chosen U value for Mn has been carefully benchmarked versus our previous results for HSE06, which has demonstrated a good ability[7] for correcting the self-interaction error for battery cathodes [39]. Looking at the kinetic barriers in Fig. 5.11, SCAN predicts a slightly increased barrier for each of the substituted elements, which could be linked to the reduced magnetic moment on *all* oxygen atoms. Even this increased

---

[7] Note that by tuning the fraction of exact exchange, it is possible to improve the accuracy of the HSE functional further, but HSE06 [38] (a = 0.25) does a fairly good job of reproducing experimental results.

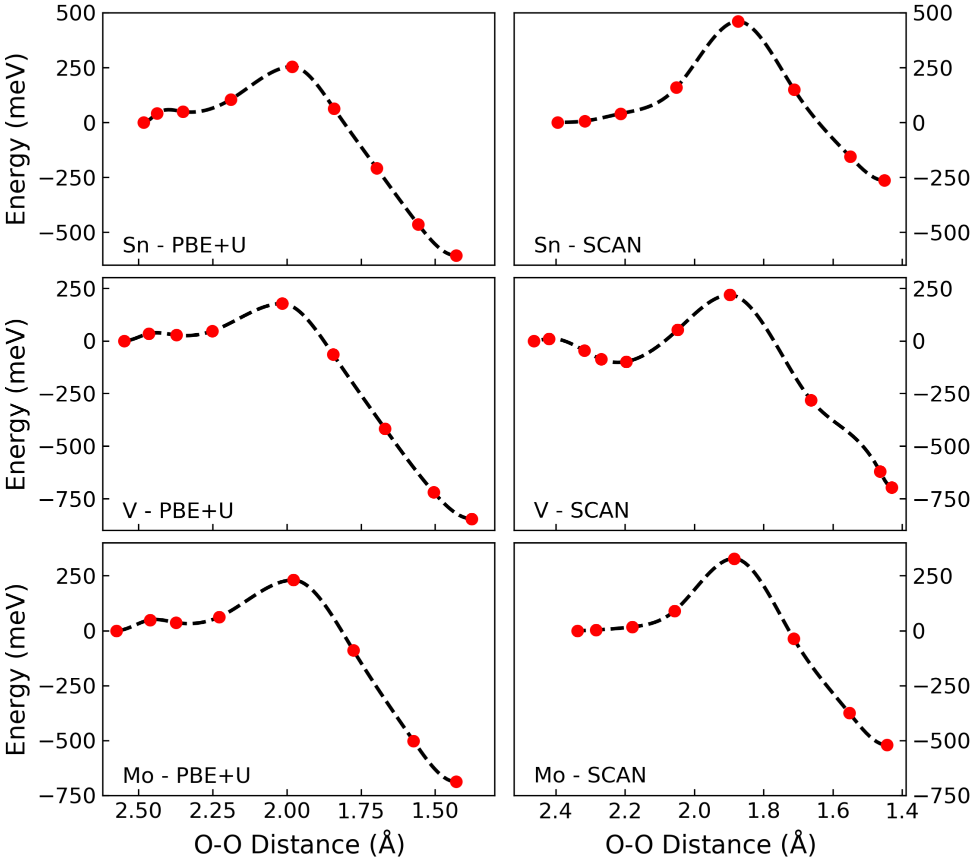


Figure 5.11: Kinetic barriers for the dimerization of oxygen neighboring Sn, V and Mo, both for the PBE+U and SCAN functional.

barrier is still relatively low, however, and hence not much of our analysis would if we would base it on the SCAN results.

5.4 Polyborane solid electrolytes

Many of the current safety issues that plague Li-ion batteries, such as thermal runaway [40] and electrolyte decomposition [41], are related to the use of a flammable liquid electrolyte [42, 43]. To prevent hazardous incidents, complex packaging design is required at the cell, module and pack level [44] which increases the dead weight of the battery, reducing the energy density. A promising strategy for dealing with these issues is replacing the liquid electrolyte by a solid state ionic conductor. Besides improving the safety, solid state electrolytes also offer improved stability, which significantly increases the lifetime of the battery [45]. Moreover, the electrochemical window of the solid electrolyte is typically larger, which allows for larger operating voltages [46] and hence significant increases in the energy density. Finally, a solid electrolyte could also enable the development of Li-metal and Li-air batteries, as well as the miniaturization [47, 48] and three dimensional battery architectures [49, 50].

A good solid electrolyte must demonstrate a high ionic conductivity and negligible electronic conductivity at the range of lithium activity and operating temperature of the battery [51]. Other important properties include the chemical stability versus reactions at the electrode interfaces, and good mechanical properties in order to accommodate for the change in volume of the electrodes during the cycling of the battery [52]. Different classes being considered as solid electrolytes include perovskite (e.g. LLTO [53]), NASICON (e.g. $Na_{1+x}Zr_2Si_xP_{3-x}O_{12}$ $(0 \leq x \leq 3)$ [54]) and garnet types (e.g. $Li_7La_3Zr_2O_{12}$ [55]). For a recent overview, we refer the reader to the review paper of Zheng et al. [56].

Here I present my contribution to the investigation of the theoretical principles behind the superionic conductivity of polyborane salts, a class of materials that has recently demonstrated significant potential as a solid electrolyte. This work was performed during a three month research stay at Lawrence Livermore National Laboratory, under the supervision of Dr. Brandon Wood and his group at the Materials Science Division. My work focused on setting up a toolbox for calculating such landscapes quickly, as described in the sections that follow and Section 3.3.5. As such, the analysis presented in the following sections has been heavily inspired by the work of my collaborators, and largely corresponds to that of Dimitrievska et al. [57].

5.4.1 Polyborane salts

Polyborane salts have a rich chemistry which has been investigated for 60 years since dodecahydro-*closo*-dodecaborate $[B_{12}H_{12}]^{2-}$ was synthesized by Pitochelli and Hawthorne [58]. The first proposition of using polyborane salts as a solid electrolyte was made by Johnson and Whittingham [59], an idea that has been revived recently due to the increased interest in solid-state batteries by Udovic et al. [60]. They found that above 529 K, $Na_2B_{12}H_{12}$ undergoes a order-disorder phase transition which increases its ionic conductivity to $> 0.1\ \mathrm{S\,cm^{-1}}$, orders of magnitude larger than at room temperature. A similar phase transition was found to occur for $Li_2B_{12}H_{12}$ at 600 K [61]. Subsequently, Tang et al. [62] found that by substituting one of the B atoms by C, the temperature of the superionic transition is reduced drastically to 400 K and 380 K for $LiCB_{11}H_{12}$ and $NaCB_{11}H_{12}$, respectively. Figure 5.12 shows the structure of $Li_2B_{12}H_{12}$ at room temperature, alongside the $[B_{12}H_{12}]^{2-}$ and $[CB_{11}H_{12}]^{-}$ anions.

The high ionic conductivity of polyborane salts is believed to be connected to rapid reorientations of the anions [63, 64], as well as the frustration between crystal symmetry and the local anion geometry as a result of long-range coulombic and short-range covalent-like interactions [65]. Moreover, the lattice stacking of the large anions introduces spacious interstitial channels which facilitate cation conduction [62], and because there are many more cation sites than cations, the structure can be interpreted as intrinsically high-vacancy, reducing the chance of migration

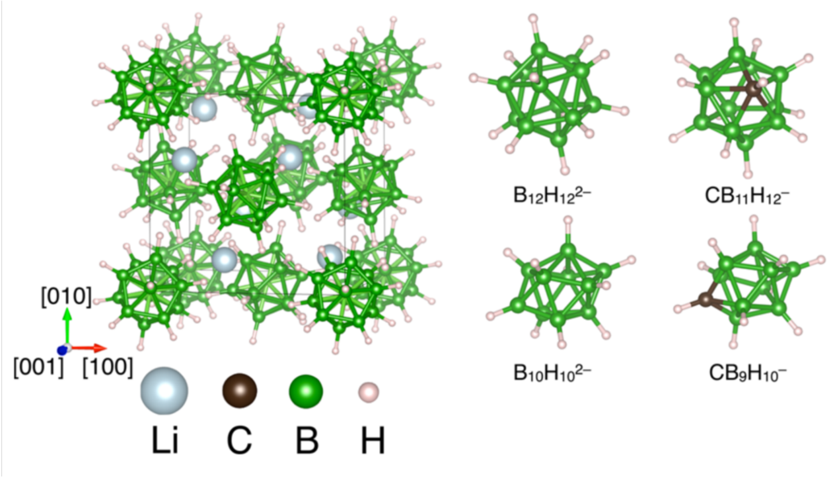


Figure 5.12: Crystalline *fcc* structure of $Li_2B_{12}H_{12}$ at room temperature.

channels being blocked. The interplay between the anion dynamics and cation mobility is complex, and I do not aim to provide a detailed explanation here. Instead, I will simply focus on my contribution to this line of research and its relation to other computational and experimental results. For more details, I refer the reader to the excellent analysis presented in the work of my collaborators [57, 64, 65].

5.4.2 Energy landscape of $[CB_{11}H_{12}]^-$

In order to understand the local interaction between the anion and cation, I have calculated the energy landscape of the cation along a chain of "wedges", i.e. curved 2D landscapes that connect the inequivalent facets of the anion (see Fig. 5.13a, as well as Fig. 3.7). This involves calculating the energy of the anion-cation system with a static calculation for many different cation positions around the anion. Moreover, in order to be able to reasonably compare the energy landscapes of Li^+ versus Na^+, I have calculated a reference energy based on the average of a spherical landscape with a radius of 8 Å. Besides giving a better idea of the binding energy of the cation-anion pair, this also provides a measure for how easily the cation is able to hop back to an interstitial site, as the *fcc* lattices of $LiCB_{11}H_{12}$ and $NaCB_{11}H_{12}$ have similar lattice constants (9.936 Å and 10.066 Å, respectively [62]). All of the landscapes presented in this section are compared with respect to this reference energy. The workflow used to calculate the landscapes is described in Section 3.3.5, the computational details can be found in corresponding section in Appendix A.2.

The resulting energy landscapes are compared with the ab initio molecular dynamics (AIMD) results from Dimitrievska et al. [57] in Fig. 5.13. The landscapes of Li^+ and Na^+ are qualitatively similar, and show a preference for the cations to bind at all-boron facets, where the depth of the energy wells is progressively larger for sites further removed from the C atom. This result is in good agreement with the angle distributions obtained from the AIMD, where we can see that the probability density is larger for angles corresponding to the all-boron docking sites (Ω_2, Ω_3 and Ω_4), especially at smaller distances. At these distances, the likelihood of finding the cation near the Ω_4 site is largest, which matches nicely with the increasing depth of the energy wells for sites further removed from C.

Moreover, the difference in the energy landscape between the C facet (Ω_1) and the lowest energy binding site (Ω_4) is significant (> 0.6 eV). As the anions undergo rapid reorientation in the superionic phase of both $LiCB_{11}H_{12}$ and $NaCB_{11}H_{12}$ [57], this leads to a strongly fluctuating

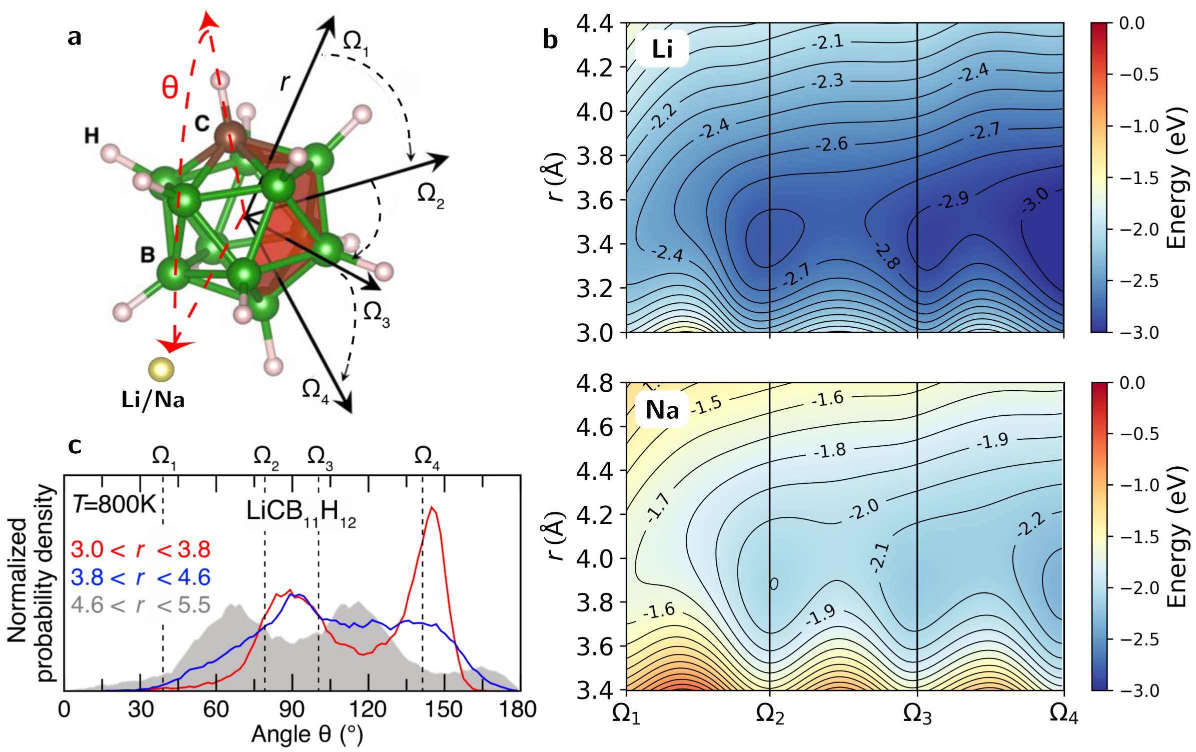


Figure 5.13: (**a**) $[CB_{11}H_{12}]^-$ anion, with the symmetrically inequivalent facets colored red. Each of the facets corresponds to a binding site with direction Ω_i with respect to the center of the anion. (**b**) Angular distributions of the Li^+ cations derived from AIMD simulations (Taken from [57]), where the angle θ is defined versus the C_5 axis connecting the C atom with the opposite B atom. (**c**) Calculated energy landscapes for Li^+ (top) and Na^+ (bottom) along wedges connecting the binding sites Ω_i, relative to the spherical average at 8 Å.

cation energy landscape close to the anions, which can push the cation back into interstitial sites. Hence, the substitution of B by C introduces a dipole in the anion, which in combination with its high rotational mobility can improve the ionic conductivity. This “paddle wheel” effect was already described by Lunder et al. in the context of lithium sulphate materials [66], and is further supported by the by the AIMD and quasielastic neutron scattering results of my collaborators [57].

Comparing the results for Li^+ and Na^+, it is clear that the cation is bound less strongly for Na^+, as the wells corresponding to the binding sites are much higher in energy compared to the reference at 8 Å. Moreover, the wells are also broader and located at a larger distance from the anion, which further indicates that cation can more easily be detached from the anion. This, in combination with the fact that the energy difference upon reorientation of the anion is similar to that for Li^+, can explain the lower transition temperature to the superionic phase for $NaCB_{11}H_{12}$ (380 K) compared to $LiCB_{11}H_{12}$ (400 K).

5.5 Conclusions and Outlook

In this chapter, I have made a comparison of the stability of the oxygen framework of two layered oxide materials which are being investigated for use as a cathode in Li-ion batteries. An extensive study of the optimal lithium configuration at different states of charge shows that when the Li_2MnO_3 cathode is charged by 75%, the stacking changes from O3 to O1. Based on the charged structure, a comparison of the stability of the oxygen framework indicates that the formation of O-O dimers is both thermodynamically and kinetically viable for O1-$Li_{0.5}MnO_3$. For O1-$Li_{0.5}IrO_3$, the oxygen lattice is much more stable, either returning to its original state when perturbed, or resulting in a structure with an O-O dimer that is much higher in energy. This can in part be explained by the mixed redox process for Li_2IrO_3, which is also confirmed by the calculated magnetic moments and calculated change in projected density of states.

The lack of O-O dimer formation in O1-$Li_{0.5}IrO_3$ suggests that introducing transition metals in the Li-rich structure which allow for higher states of oxidation is a reasonable path for reducing the likelihood of the formation of O-O dimers, and the corresponding structural changes of the cathode that are tied to the detrimental voltage fade and oxygen evolution. However, other research has also shown that Sn substitution can improve the structural stability. We have studied the solubility of Sn in the $Li_{1.2}Mn_{0.8}O_2$ structure, and find that only a limited substitution is thermodynamically feasible. Based on these results, we decided to study the influence of a single substitution of Mn by Sn, V or Mo on the oxygen oxidation and stability of its framework. Our results indicate that substituting Mn by Sn does little to change the properties of the oxygen framework, most likely due to their similar chemical inactivity during the charging process. For V and Mo, the substitution does reduce the oxidation of the neighboring oxygen atoms, but does not result in an improved stability. Instead, the kinetic barrier for dimerization is decreased further, indicating that the substitution destabilizes the structure instead.

Although our results indicate that the formation of oxygen dimers in O1-$Li_{0.5}MnO_3$ is likely to occur, we have yet to study its connection with the migration of Mn into the lithium layer. Other further investigations that could be interesting are the formation of oxygen dimers at the cathode surface, and subsequent evolution of O_2 from the cathode into the electrolyte. So far, no substitution seems to be successful at stabilizing the structure. However, other approaches have been suggested for increasing the cycling properties of Li-rich materials, such as Ni substitution in the Li layer [67], or the substitution of oxygen by fluor [68, 69]. Both make sense in the context of our results. The most likely dimer according to our analysis is formed across the Li layer, which would be inhibited by the presence of Ni. Fluor, on the other hand, does not oxidize Mn as much, leaving more room for the transition metal to oxidize as Li is removed from the cathode. Further research is necessary to see if these ideas can properly stabilize the Li-rich cathode, opening it up to further development and integration in commercial applications.

Finally, we have calculated the energy landscapes of Li^+ and Na^+ cations around the carbonated polyborane salt anion $[CB_{11}H_{12}]^-$. From the landscapes, it is clear than substituting a single boron by carbon introduces a dipole in the anion molecule, which in combination with the rapid reorientations of the anions results in a paddle wheel mechanism that improves the ionic conductivity of the material. This suggests a novel strategy for improving the properties of these materials for solid-state battery applications.

Bibliography

[1] Engie-Electrabel. *Coo: recognised expertise in energy storage.* 2015. URL: `https://corporate.engie.be/en/coo-recognised-expertise-energy-storage/` (visited on 12/06/2019).

[2] J. B. Goodenough and K.-S. Park. "The Li-Ion Rechargeable Battery: A Perspective". *J. Am. Chem. Soc.* 135 (2013), pp. 1167–1176. DOI: `10.1021/ja3091438`.

[3] C. Mao et al. "Selecting the best graphite for long-life, high-energy li-ion batteries". *J. Electrochem. Soc.* 165 (2018), A1837–A1845. DOI: `10.1149/2.1111809jes`.

[4] D. Bresser, E. Paillard, and S. Passerini. "Lithium-ion batteries (LIBs) for medium- and large-scale energy storage:" in: *Advances in Batteries for Medium and Large-Scale Energy Storage.* Elsevier, 2015, pp. 125–211. DOI: `10.1016/b978-1-78242-013-2.00006-6`.

[5] M. M. Thackeray. "Spinel Electrodes for Lithium Batteries". *J. Am. Ceram. Soc.* 82 (2004), pp. 3347–3354. DOI: `10.1111/j.1151-2916.1999.tb02250.x`.

[6] A. K. Padhi. "Phospho-olivines as Positive-Electrode Materials for Rechargeable Lithium Batteries". *J. Electrochem. Soc.* 144 (1997), p. 1188. DOI: `10.1149/1.1837571`.

[7] A. Van der Ven and G. Ceder. "Lithium diffusion mechanisms in layered intercalation compounds". *J. Power Sources* 97-98 (2001), pp. 529–531. DOI: `10.1016/S0378-7753(01)00638-3`.

[8] D. Larcher and J. M. Tarascon. "Towards greener and more sustainable batteries for electrical energy storage". *Nat. Chem.* 7 (2015), pp. 19–29. DOI: `10.1038/nchem.2085`.

[9] P. Rozier and J. M. Tarascon. "Review-Li-rich layered oxide cathodes for next-generation Li-ion batteries: Chances and challenges". *J. Electrochem. Soc.* 162 (2015), A2490–A2499. DOI: `10.1149/2.0111514jes`.

[10] T. Ohzuku, A. Ueda, and M. Nagayama. "Electrochemistry and Structural Chemistry of $LiNiO_2$ (R3m) for 4 Volt Secondary Lithium Cells". *J. Electrochem. Soc.* 140 (1993), pp. 1862–1870. DOI: `10.1149/1.2220730`.

[11] G. Vitins and K. West. "Lithium intercalation into layered $LiMnO_2$". *J. Electrochem. Soc.* 144 (1997), pp. 2587–2592. DOI: `10.1149/1.1837869`.

[12] Y. Koyama et al. "Crystal and electronic structures of superstructural $Li_{1-x}[Co_{1/3}Ni_{1/3}Mn_{1/3}]O_2$ ($0 \leq x \leq 1$)". *J. Power Sources* 119-121 (2003), pp. 644–648. DOI: `10.1016/s0378-7753(03)00194-0`.

[13] F. Zhou et al. "Comparison of $Li[Li_{1/9}Ni_{1/3}Mn_{5/9}]O_2$, $Li[Li_{1/5}Ni_{1/5}Mn_{3/5}]O_2$, $LiNi_{0.5}Mn_{1.5}O_4$, and $LiNi_{2/3}Mn_{1/3}O_2$ as High Voltage Positive Electrode Materials". *J. Electrochem. Soc.* 158 (2011). DOI: `10.1149/1.3525268`.

[14] M. M. Thackeray et al. "Li_2MnO_3-stabilized $LiMO_2$ (M = Mn, Ni, Co) electrodes for lithium-ion batteries". *J. Mater. Chem.* 17 (2007), pp. 3112–3125. DOI: `10.1039/b702425h`.

[15] M. Sathiya et al. "Reversible anionic redox chemistry in high-capacity layered-oxide electrodes". *Nat. Mater.* 12 (2013), pp. 827–835. DOI: `10.1038/nmat3699`.

[16] J. R. Croy et al. "Quantifying hysteresis and voltage fade in $xLi_2MnO_3 \cdot (1-x)LiMn_{0.5}Ni_{0.5}O_2$ electrodes as a function of Li_2MnO_3 content". *J. Electrochem. Soc.* 161 (2014). DOI: `10.1149/2.049403jes`.

[17] E. McCalla et al. "Visualization of O-O peroxo-like dimers in high-capacity layered oxides for Li-ion batteries". *Science* 350 (2015), pp. 1516–1521. DOI: `10.1126/science.aac8260`.

[18] Y. Koyama et al. "First-principles study on lithium removal from Li_2MnO_3". *J. Power Sources* 189 (2009), pp. 798–801. DOI: `10.1016/j.jpowsour.2008.07.073`.

[19] C. Delmas, C. Fouassier, and P. Hagenmuller. "Structural classification and properties of the layered oxides". *Physica B+C* 99 (1980), pp. 81–85. DOI: `10.1016/0378-4363(80)90214-4`.

[20] E. Castel et al. "Differential Electrochemical Mass Spectrometry Study of the Interface of $xLi_2MnO_3\cdot(1\text{–}x)LiMO_2$ (M = Ni, Co, and Mn) Material as a Positive Electrode in Li-Ion Batteries". *Chem. Mater.* 26 (2014), pp. 5051–5057. DOI: `10.1021/cm502201z`.

[21] H. Chen and M. S. Islam. "Lithium extraction mechanism in Li-rich Li_2MnO_3 involving oxygen hole formation and dimerization". *Chem. Mater.* 28 (2016), pp. 6656–6663. DOI: `10.1021/acs.chemmater.6b02870`.

[22] D. H. Seo et al. "The structural and chemical origin of the oxygen redox activity in layered and cation-disordered Li-excess cathode materials". *Nat. Chem.* 8 (2016), pp. 692–697. DOI: `10.1038/nchem.2524`.

[23] J. Hong et al. "Metal–oxygen decoordination stabilizes anion redox in Li-rich oxides". *Nat. Mater.* 18 (2019), pp. 256–265. DOI: `10.1038/s41563-018-0276-1`.

[24] J. A. Saint, M. M. Doeff, and J. Reed. "Synthesis and electrochemistry of Li_3MnO_4: Mn in the +5 oxidation state". *J. Power Sources* 172 (2007), pp. 189–197. DOI: `10.1016/j.jpowsour.2007.07.027`.

[25] M. Saubanère et al. "The intriguing question of anionic redox in high-energy density cathodes for Li-ion batteries". *Energy Environ. Sci.* 9 (2016), pp. 984–991. DOI: `10.1039/c5ee03048j`.

[26] X. Li et al. "Direct Visualization of the Reversible O^{2-}/O^{-}Redox Process in Li-Rich Cathode Materials". *Adv. Mater.* 30 (2018). DOI: `10.1002/adma.201705197`.

[27] A. R. Armstrong et al. "Demonstrating Oxygen Loss and Associated Structural Reorganization in the Lithium Battery Cathode $Li[Ni_{0.2}Li_{0.2}Mn_{0.6}]O_2$". *J. Am. Chem. Soc.* (2006), pp. 8694–8698. DOI: `10.1021/ja062027`.

[28] K. Luo et al. "Anion Redox Chemistry in the Cobalt Free 3d Transition Metal Oxide Intercalation Electrode $Li[Li_{0.2}Ni_{0.2}Mn_{0.6}]O_2$". *J. Am. Chem. Soc.* 138 (2016), pp. 11211–11218. DOI: `10.1021/jacs.6b05111`.

[29] L. Slap. "De performantie van DFT-functionalen op Li-rijke kathode materialen voor batterijen". MA thesis. University of Antwerp, 2020.

[30] A. Van Der Ven, J. Bhattacharya, and A. A. Belak. "Understanding Li diffusion in Li-intercalation compounds". *Acc. Chem. Res.* 46 (2013), pp. 1216–1225. DOI: `10.1021/ar200329r`.

[31] Q. Q. Qiao et al. "Sn-stabilized Li-rich layered $Li(Li_{0.17}Ni_{0.25}Mn_{0.58})O_2$ oxide as a cathode for advanced lithium-ion batteries". *J. Mater. Chem. A* 3 (2015), pp. 17627–17634. DOI: `10.1039/c5ta03415a`.

[32] W. Sun et al. "The thermodynamic scale of inorganic crystalline metastability". *Sci. Adv.* 2 (2016), e1600225. DOI: `10.1126/sciadv.1600225`.

[33] J. Ma et al. "Molybdenum Substitution for Improving the Charge Compensation and Activity of Li_2MnO_3". *Chem.: Eur. J* 20 (2014), pp. 8723–8730. DOI: `10.1002/chem.201402727`.

[34] R. Xiao, H. Li, and L. Chen. "Density functional investigation on Li_2MnO_3". *Chem. Mater.* 24 (2012), pp. 4242–4251. DOI: `10.1021/cm3027219`.

[35] S. L. Dudarev et al. “Electron-energy-loss spectra and the structural stability of nickel oxide: An LSDA+U study”. *Phys. Rev. B* 57 (1998), pp. 1505–1509. DOI: `10.1103/PhysRevB.57.1505`.

[36] O. Bengone et al. “Implementation of the projector augmented-wave LDA+U method: Application to the electronic structure of NiO”. *Phys. Rev. B* 62 (2000), pp. 16392–16401. DOI: `10.1103/PhysRevB.62.16392`.

[37] J. Sun, A. Ruzsinszky, and J. Perdew. “Strongly Constrained and Appropriately Normed Semilocal Density Functional”. *Phys. Rev. Lett.* 115 (2015), p. 036402. DOI: `10.1103/PhysRevLett.115.036402`.

[38] J. Heyd, G. E. Scuseria, and M. Ernzerhof. “Hybrid functionals based on a screened Coulomb potential”. *J. Chem. Phys.* 118 (2003), pp. 8207–8215. DOI: `10.1063/1.1564060`.

[39] D. H. Seo, A. Urban, and G. Ceder. “Calibrating transition-metal energy levels and oxygen bands in first-principles calculations: Accurate prediction of redox potentials and charge transfer in lithium transition-metal oxides”. *Phys. Rev. B* 92 (2015). DOI: `10.1103/PhysRevB.92.115118`.

[40] Q. Wang et al. “Thermal runaway caused fire and explosion of lithium ion battery”. *Journal of Power Sources* 208 (2012), pp. 210–224. DOI: `10.1016/j.jpowsour.2012.02.038`.

[41] D. Lisbona and T. Snee. “A review of hazards associated with primary lithium and lithium-ion batteries”. *Process Saf. Environ. Prot.* 89 (2011), pp. 434–442. DOI: `10.1016/j.psep.2011.06.022`.

[42] K. Liu et al. “Materials for lithium-ion battery safety”. *Sci. Adv.* 4 (2018), eaas9820. DOI: `10.1126/sciadv.aas9820`.

[43] D. Ouyang et al. “A Review on the Thermal Hazards of the Lithium-Ion Battery and the Corresponding Countermeasures”. *Appl. Sci.* 9 (2019), p. 2483. DOI: `10.3390/app9122483`.

[44] D. Doughty and E. P. Roth. “A General Discussion of Li Ion Battery Safety”. *Electrochem. Soc. Interface* (2012). DOI: `10.1149/2.f03122if`.

[45] Mauger et al. “Building Better Batteries in the Solid State: A Review”. *Materials* 12 (2019), p. 3892. DOI: `10.3390/ma12233892`.

[46] J. Li et al. “Solid Electrolyte: the Key for High-Voltage Lithium Batteries”. *Advanced Energy Materials* 5 (2014), p. 1401408. DOI: `10.1002/aenm.201401408`.

[47] J Bates. “Thin-film lithium and lithium-ion batteries”. *Solid State Ionics* 135 (2000), pp. 33–45. DOI: `10.1016/s0167-2738(00)00327-1`.

[48] H.-D. Um et al. “Monolithically integrated, photo-rechargeable portable power sources based on miniaturized Si solar cells and printed solid-state lithium-ion batteries”. *Energy Environ. Sci* 10 (2017), pp. 931–940. DOI: `10.1039/c6ee03266d`.

[49] J. W. Long et al. “Three-Dimensional Battery Architectures”. *Chem. Rev.* 104 (2004), pp. 4463–4492. DOI: `10.1021/cr020740l`.

[50] A. Pearse et al. “Three-Dimensional Solid-State Lithium-Ion Batteries Fabricated by Conformal Vapor-Phase Chemistry”. *ACS Nano* 12 (2018), pp. 4286–4294. DOI: `10.1021/acsnano.7b08751`.

[51] P. Knauth. “Inorganic solid Li ion conductors: An overview”. *Solid State Ionics* 180 (2009), pp. 911–916. DOI: `10.1016/j.ssi.2009.03.022`.

[52] R. Koerver et al. “Chemo-mechanical expansion of lithium electrode materials – on the route to mechanically optimized all-solid-state batteries”. *Energy Environ. Sci* 11 (2018), pp. 2142–2158. DOI: `10.1039/c8ee00907d`.

[53] Y. Inaguma et al. “High ionic conductivity in lithium lanthanum titanate”. *Solid State Commun.* 86 (1993), pp. 689–693. DOI: `10.1016/0038-1098(93)90841-a`.

[54] L.-O. Hagman et al. “The Crystal Structure of $NaM_2IV(PO_4)_3$; MeIV = Ge, Ti, Zr”. *Acta Chem. Scand.* 22 (1968), pp. 1822–1832. DOI: `10.3891/acta.chem.scand.22-1822`.

[55] R. Murugan, V. Thangadurai, and W. Weppner. “Fast Lithium Ion Conduction in Garnet-Type $Li_7La_3Zr_2O_{12}$”. *Angew. Chem. Int. Ed.* 46 (2007), pp. 7778–7781. DOI: `10.1002/anie.200701144`.

[56] F. Zheng et al. “Review on solid electrolytes for all-solid-state lithium-ion batteries”. *J. Power Sources* 389 (2018), pp. 198–213. DOI: `10.1016/j.jpowsour.2018.04.022`.

[57] M. Dimitrievska et al. “Carbon Incorporation and Anion Dynamics as Synergistic Drivers for Ultrafast Diffusion in Superionic $LiCB_{11}H_{12}$ and $NaCB_{11}H_{12}$”. *Adv. Energy Mater.* 8 (2018), p. 1703422. DOI: `10.1002/aenm.201703422`.

[58] A. R. Pitochelli and F. M. Hawthorne. “The isolation of the icosahedral $B_{12}H_{12}^{-2}$ ion”. *J. Am. Chem. Soc.* 82 (1960), pp. 3228–3229. DOI: `10.1021/ja01497a069`.

[59] J. W. Johnson. “Lithium Closoboranes as Electrolytes in Solid Cathode Lithium Cells”. *J Electrochem Soc* 127 (1980), p. 1653. DOI: `10.1149/1.2129971`.

[60] T. J. Udovic et al. “Sodium superionic conduction in $Na_2B_{12}H_{12}$”. *Chem. Commun.* 50 (2014), pp. 3750–3752. DOI: `10.1039/c3cc49805k`.

[61] N. Verdal et al. “Complex high-temperature phase transitions in $Li_2B_{12}H_{12}$ and $Na_2B_{12}H_{12}$”. *J. Solid State Chem.* 212 (2014), pp. 81–91. DOI: `10.1016/j.jssc.2014.01.006`.

[62] W. S. Tang et al. “Unparalleled lithium and sodium superionic conduction in solid electrolytes with large monovalent cage-like anions”. *Energy Environ. Sci* 8 (2015), pp. 3637–3645. DOI: `10.1039/c5ee02941d`.

[63] A. V. Skripov et al. “Anion Reorientations and Cation Diffusion in LiCB11H12 and NaCB11H12: 1H, 7Li, and 23Na NMR Studies”. *J. Phys. Chem. C* 119 (2015), pp. 26912–26918. DOI: `10.1021/acs.jpcc.5b10055`.

[64] J. B. Varley et al. “Understanding Ionic Conductivity Trends in Polyborane Solid Electrolytes from Ab Initio Molecular Dynamics”. *ACS Energy Lett.* 2 (2016), pp. 250–255. DOI: `10.1021/acsenergylett.6b00620`.

[65] K. E. Kweon et al. “Structural, Chemical, and Dynamical Frustration: Origins of Superionic Conductivity in closo-Borate Solid Electrolytes”. *Chem. Mater.* 29 (2017), pp. 9142–9153. DOI: `10.1021/acs.chemmater.7b02902`.

[66] A. Lundén. “On the Paddle -Wheel Mechanism for Cation Conduction in Lithium Sulphate”. *Zeitschrift für Naturforschung A* 50 (1995). DOI: `10.1515/zna-1995-1114`.

[67] M. Y. Yang et al. “Role of Ordered Ni Atoms in Li Layers for Li-Rich Layered Cathode Materials”. *Adv. Funct. Mater.* 27 (2017), p. 1700982. DOI: `10.1002/adfm.201700982`.

[68] W. D. Richards et al. “Fluorination of Lithium-Excess Transition Metal Oxide Cathode Materials”. *Adv. Energy Mater.* 8 (2017), p. 1701533. DOI: `10.1002/aenm.201701533`.

[69] A. Kapylou et al. “Improved Thermal Stability of Lithium-Rich Layered Oxide by Fluorine Doping”. *ChemPhysChem* 19 (2017), pp. 116–122. DOI: `10.1002/cphc.201700927`.

Chapter 6

Ion Induced Secondary Electron Emission

> *"I have not failed. I've just found thousands of ways that don't work."*
>
> Thomas A. Edison

> *"With four parameters I can fit an elephant, and with five I can make him wiggle his trunk."*
>
> John von Neumann

When slow ions incident on a surface are neutralized, the excess potential energy is passed on to an electron inside the surface, leading to emission of secondary electrons. The microscopic description of this process, as well as the calculation of the secondary electron yield, is a challenging problem due to its complexity as well as its sensitivity to surface properties. One of the first quantitative descriptions was articulated in the 1950s by Hagstrum, who based his calculation on a parameterization of the density of states of the material.

In this chapter, I present a new model for calculating the secondary electron yield, derived from Hagstrum's initial approach. After providing a brief introduction to the topic in Sec. 6.1, Sec. 6.2 continues by discussing Hagstrum's model in the context of semiconductors, as well as our adjustments to the escape function (Sec. 6.2.2) and the introduction of electron cascades (Sec. 6.2.3). In Sec. 6.2.4, I compare the results of the model for He^+ and Ne^+ ions incident on Ge(111) and Si(111) with those obtained from experiment. Finally, I expand the application of the model to metals in Sec. 6.3, introducing the influence of collective excitations (Sec. 6.3.1). The chapter concludes by discussing the results of the final model, first by comparing the calculated yield spectra with experiment (Sec. 6.3.2) and performing a high-throughput analysis for elemental surfaces covering a large part of the periodic table (Sec. 6.3.3).

6.1 Introduction

Secondary electron emission (SEE) is an important phenomenon where electrons of a target material are emitted through the impact of energetic primary particles. Such processes lie at the foundation of several techniques for characterizing surfaces and play an important role in applications such as plasma sputtering deposition [1, 2] and plasma display panels [3, 4]. For the case of incident ions, measurements of the electron yield γ were first performed by Hagstrum [5, 6]. Although there is a considerable need for data on the electron emission from incident ions, e.g. as input for models of micro-plasmas [7, 8], measurements of γ are limited. This lack of experimental data has led to an increased interest in theoretical modeling of ion-induced secondary electron emission.

One of the first quantitative descriptions of the interaction of incident ions with a surface was also developed by Hagstrum [9, 10], who used a model for the density of states in order to derive the yield of secondary electrons emitted from the surface per incoming ion. His approach was relatively successful, but required the use of a substantial amount of fitting parameters in order to make his simulations match experiment. Since then, there have been numerous attempts at improving the quantitative modeling of the SEE process. Propst [11] calculated the Auger matrix elements using a WKB approximation[1] for the tunneling process, and was one of the first to consider electron-electron interactions for the secondary electrons. Several authors [12–14] have used a jellium model to calculate the Auger neutralization rate for aluminum and sodium, for which the electrons are well described by a free electron gas. More recently, Cho et al. [4] have used first-principles DFT calculations in combination with Hagstrum's model to determine the SEE coefficients for MgO. However, they used parameters which were fitted for a specific ion-surface combination ($He^+ \rightarrow Ge$) and applied them as if they were fixed parameters of the model. A nice overview of the various theoretical models used to describe the neutralization of incoming ions on a surface can be found in a recent review paper by Monreal [15].

Our approach is to make several adjustments to Hagstrum's model in order to reduce its dependency on fitting parameters and improve the calculated yield spectra. In this way, we aim to provide a quantitative approach to calculating the secondary electron emission from surfaces bombarded by slow ions. Similar to Cho et al., we use first-principles density functional theory calculations to acquire the necessary input, using the workflow described in Sec. 3.3.4.

6.2 Semiconductors

Secondary electron emission from incident ions can be the result of either kinetic or potential emission mechanisms [16]. For the former, the emission process is driven by the kinetic energy of the incoming ion. Potential electron emission, on the other hand, transfers the potential energy of the incoming ion to an electron that can then be emitted. For slow moving ions, the potential emission mechanism is dominant [17]. Following Hagstrum's description, it is performed using a two-electron Auger process. Here, a distinction must be made between Auger neutralization (AN) and resonance neutralization (RN) followed by Auger de-excitation (AD), see Fig. 6.1. For AN, an electron from the surface material tunnels directly into the lowest unoccupied state of the incoming ion. The energy released in this transition is passed to another electron in the material. This electron, in turn, has a probability of escape if it is excited to a state with an energy above the vacuum level. In the case of RN, an electron from the surface tunnels into an excited state of the incoming ion, after which the atom returns to the ground state by AD. For slow incident He^+ and Ne^+ ions on Ge(111) and Si(111), the resonant neutralization channel is unavailable [10, 13], so we can safely disregard the RN and AD processes in our discussion for now.

---

[1] Named after Wentzel, Kramers and Brillouin.

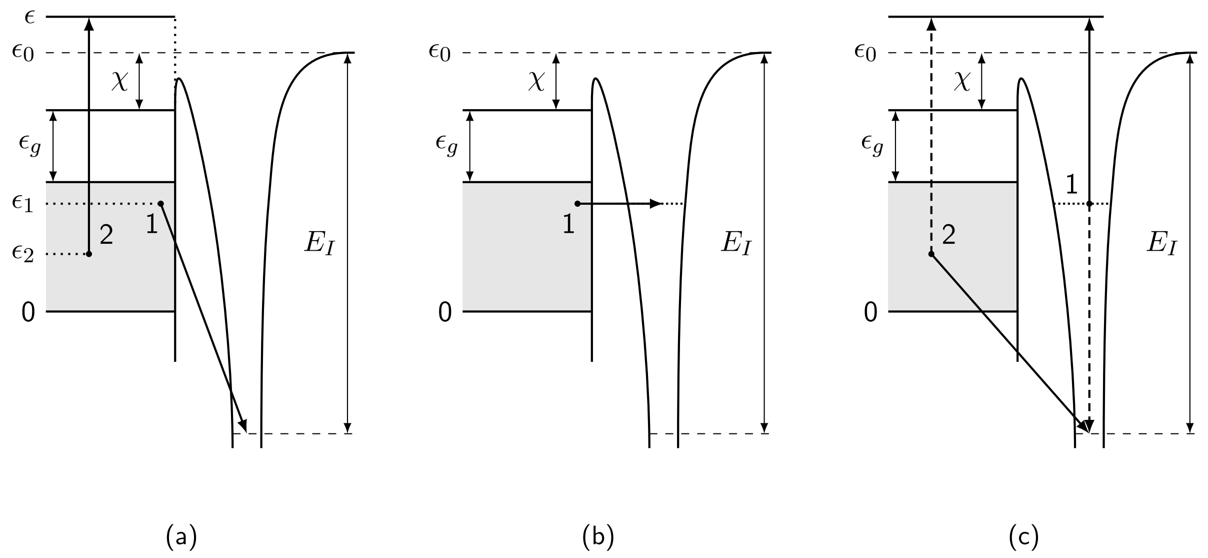


Figure 6.1: Schematic representations of the Auger neutralization (a), resonance neutralization (b) and Auger de-excitation (c) processes. For the de-excitation, the set of full and dashed arrows each represent one possible way via which the de-excitation can occur.

6.2.1 Hagstrum's model

I begin by presenting Hagstrum's quantitative description of ion-based SEE for semiconductor surfaces, before discussing our adjustments in the sections that follow. During the Auger neutralization, an electron tunnels through the surface barrier to the lowest unoccupied ionic state, transferring the excess energy to a secondary electron of the surface material (Fig. 6.1a):

$$\epsilon_1 + \epsilon_2 \rightarrow \epsilon + \epsilon_0 - E_I. \tag{6.1}$$

Here, ϵ_1 and ϵ_2 are the initial energies of the electrons of the surface, ϵ is the energy of the excited Auger electron, ϵ_0 is the vacuum level, and E_I the ionization energy of the incoming ion. Determining the transition rates of the various possible transitions of the Auger neutralization process involves the calculation of the matrix elements in Fermi's golden rule:

$$\Gamma_{i\rightarrow f} = \frac{2\pi}{\hbar} \left| H_{i\rightarrow f} \right|^2 N(\epsilon), \tag{6.2}$$

where $N(\epsilon)$ is the density of states at the final electron energy, $\hbar$ is the reduced Planck constant and $H_{i\rightarrow f}$ is the transition matrix element:

$$H_{i\rightarrow f} = \int\int u_g^*(\mathbf{r}_1) u_e^*(\mathbf{r}_2) V(\mathbf{r}_1, \mathbf{r}_2) u_v'(\mathbf{r}_2) u_v''(\mathbf{r}_1) d\mathbf{r}_1 d\mathbf{r}_2, \tag{6.3}$$

with $u_v'(\mathbf{r}_2)$ and $u_v''(\mathbf{r}_1)$ the initial states of the electrons in the valence band, $u_g(\mathbf{r}_1)$ the ground state of the neutralized ion, and $u_e(\mathbf{r}_2)$ the excited state of the Auger electron, where * denotes the complex conjugate. The interaction potential $V(\mathbf{r}_1, \mathbf{r}_2)$ is the Coulomb potential.

Instead of explicitly calculating the secondary emission yield from the transition matrix elements, as in the earlier work of e.g. Cobas and Lamb [18], Hagstrum introduced a model based on the density of states (DOS) of the surface. In his approach, the matrix element for each transition is considered to be constant, which means that the probability that an electron of energy ϵ will participate in the Auger neutralization is proportional to the density of states $N(\epsilon)$. In this way, we can simply use the density of states of the surface to calculate the probability that a secondary electron with kinetic energy $\epsilon_k = \epsilon - \epsilon_0$ is emitted from the impact of an incoming ion. First, the internal distribution of excited electrons $N_i(\epsilon)$ is calculated by

$$N_i(\epsilon) = \frac{D_c(\epsilon) T \left[\dfrac{\epsilon + \epsilon_0 - E_I}{2} \right]}{\displaystyle\int_{\epsilon_c}^{\infty} D_c(\epsilon) T \left[\dfrac{\epsilon + \epsilon_0 - E_I}{2} \right] d\epsilon}, \tag{6.4}$$

where D_c is the density of the unoccupied states of the surface, ϵ_c the bottom of the conduction band, and T is the Auger transform:

$$\begin{aligned} T\left[\frac{\epsilon + \epsilon_0 - E_I}{2} \right] = \int_{\Delta\epsilon_v} \int_{\Delta\epsilon_v} & D_v(\epsilon_1) D_v(\epsilon_2) \\ & \delta(\epsilon - \epsilon_1 - \epsilon_2 + \epsilon_0 - E_I) d\epsilon_1 d\epsilon_2. \end{aligned} \tag{6.5}$$

Here, $D_v(\epsilon)$ is the density of the valence states, $\Delta\epsilon_v$ is the valence band width, and δ is the Dirac delta function. Note that $N_i(\epsilon)$ is normalized to unity because of the assumption that every incoming ion is neutralized, producing one excited electron inside the surface. The delta function is used to assert energy conservation of the Auger neutralization process (Eq. 6.1).

Next, the distribution of the electrons that can escape from the surface $N_0(\epsilon)$ is calculated by multiplying $N_i(\epsilon)$ with the aptly named escape probability $P_e(\epsilon)$:

$$N_0(\epsilon) = P_e(\epsilon) N_i(\epsilon). \tag{6.6}$$

Hagstrum modeled the escape probability $P_e(\epsilon)$ using a semiclassical approach, where an electron is considered to be able to escape when the projection of its wave vector on the axis perpendicular to the surface is large enough, i.e. when its corresponding energy is larger than the vacuum level. The resulting escape probability is:

$$\begin{aligned} P_e(\epsilon) &= \frac{1}{2} \left[1 - \left(\frac{\epsilon_0}{\epsilon} \right)^{\beta} \right]^{\alpha} \quad &\text{for} \quad \epsilon \geq \epsilon_0, \\ &= 0 &\text{for} \quad \epsilon < \epsilon_0 \end{aligned} \tag{6.7}$$

where α and β represent an anisotropy of the distribution of the initial direction of the wave vector of the electron after excitation. In Hagstrum's approach, these are parameters which are fitted for each ion-surface combination. Finally, the secondary electron yield γ, i.e. the amount of electrons emitted per incoming ion, is calculated by integrating the distribution of escaped electrons:

$$\gamma = \int_{\epsilon_0}^{\infty} N_0(\epsilon) d\epsilon. \tag{6.8}$$

Using his model, Hagstrum was able to fairly accurately reproduce the yield spectra for He^+ and Ne^+ incident on Ge(111) and Si(111). Note that all energies in Eq. 6.7 are defined with respect to the valence band minimum, which was taken as the zero-energy level by Hagstrum. This corresponds to implicitly deciding on a reference level that determines the effective barrier ϵ_0 that the electrons have to pass in order to escape from the surface. Because this reference level affects the probability that an excited electron is emitted, it has a significant influence on the calculated yield.

Although Hagstrum's description of the SEE process was a major step forward, providing an insightful interpretation of his experimental results, he was forced to rely on several fitting parameters in order to be able to reproduce the experimental spectra. Hagstrum had to parameterize the density of states, ionization energy, as well as fit α and β in the escape function (Eq. (6.7)). In our approach, we calculate the density of states from first principles within the DFT framework and remove the parameter dependency of the escape function as described in Sec. 6.2.2. For the ionization energy E_I, we shift the ionization energy of the free atom with the image interaction (2 eV) [19, 20]. Finally, we introduce electron-electron scattering as a cascade process in order to improve the calculated yield spectra (Sec. 6.2.3).

By virtue of using Hagstrum's model, we are able to include the first-principles calculated electronic structure of the surface, as directly calculating the full matrix elements in Eq. 6.2 from the DFT wave functions would not be computationally feasible. Several such calculations have been performed for jellium-model wave functions, but this approach would in turn not provide a good description for a semiconductor, nor does it allow us to include the electronic structure of the chosen surface. Gloebl et al. [21, 22] and Valdès et al. [23] did calculate the Auger Neutralization rate by considering the matrix elements for the neutralization using a linear combination of atomic orbitals (LCAO) and various distances and positions of an incoming He$^+$ ion on several metals and Ge. The electronic response of the surface was modeled using the response function, once again calculated from the jellium model. They found that the Auger rate was not sensitive to the position of the incoming ion at short distances for Ge, which indicates that considering constant matrix elements is a reasonable approximation for Ge. However, they also found that for noble metals, the presence of d electrons in the valence band that can neutralize the incoming ion has a large effect on the AN rate. The efficient contribution of *d* electrons compared to *s* or *p* states will have to be considered in case the model is extended to be applied to noble metals.

6.2.2 Escape Function

An integral part of Hagstrum's model is the escape probability function $P_e(\epsilon)$, which represents the probability that an electron at energy ϵ can escape from the surface after excitation. Initially Hagstrum had derived a parameterless expression for P_e, based on an isotropic angular distribution of the wave vector of the excited electrons. However, the resulting distributions of escaped electrons, or yield spectra, did not have sufficient electrons, especially at lower energies. For this reason, Hagstrum introduced the parameters α and β (see Eq. (6.7)), representing an anisotropy of the initial direction of the wave vector of the electron. By fitting these parameters, Hagstrum was able to adjust the escape function and increase the secondary electron yield. Because α and β are fitted for each ion/surface combination, however, the use of the escape probability of Hagstrum's model is ill suited for any model that aims to determine the secondary electron emission without relying on experimental input. The approach of Motoyama et al. [3] and Cho et al. [4], who simply used the fitted parameters for He$^+$ ions incident on Ge and applied them to other systems, is questionable at best. Finally, Hagstrum's expression for the escape function depends on where we set the zero-energy level. Here, Hagstrum defined all energies with respect to the bottom of the valence band, which results in a low probability of escape due to the relatively large surface barrier.

Instead of Hagstrum's expression for the escape function, we choose to take a quantum mechanical step-barrier approach similar to that of Lorente et al [13]. In this framework, the surface is described as a step function barrier, and the angle-dependent escape probability is derived from the transmission coefficient:

$$P_e(\epsilon, \theta) = T(k_\perp, p_\perp), \tag{6.9}$$

where θ is the angle between the initial wave vector and the surface normal, and $k_\perp$ and $p_\perp$ are the projections of the wave vector on the surface normal inside and outside of the material, i.e. $k_\perp = k\cos(\theta)$. For semiconductors, the barrier is set equal to the electron affinity χ [24–26], which corresponds to setting the reference energy level at the bottom of the conduction band ϵ_c. Using this convention, the wave vector is calculated from

$$k = \sqrt{2m_e(\epsilon - \epsilon_c)/\hbar^2}, \tag{6.10}$$

where ϵ_c is the bottom of the conduction band, $\hbar$ is the reduced Planck constant and m_e is the electron mass. $p_\perp$ is determined using the refraction condition at the surface:

$$\frac{\hbar^2 k_\perp^2}{2m_e} - \chi = \frac{\hbar^2 p_\perp^2}{2m_e}. \tag{6.11}$$

For a step-barrier, the transmission coefficient is given by [13]:

$$T(k_\perp, p_\perp) = \frac{4k_\perp p_\perp}{(k_\perp + p_\perp)^2}. \tag{6.12}$$

Next, in order to determine the escape probability for an excited electron with energy ϵ, we use the following expression [9]:

$$P_e(\epsilon) = \int P_e(\epsilon, \theta) P_\Omega(\epsilon) d\Omega, \tag{6.13}$$

where $P_\Omega(\epsilon)d\Omega$ is the probability that an electron with energy ϵ has a wave vector $\mathbf{k}$ with a direction that is part of the solid angle $d\Omega = \sin\theta d\theta d\varphi$. If, similar to Hagstrum's initial approach, we assume this distribution to be isotropic, i.e. $P_\Omega(\epsilon) = 1/4\pi$, and only consider the half sphere in the direction of the vacuum, we obtain:

$$P_e(\epsilon) = \frac{1}{4\pi}\int_0^{2\pi} d\varphi \int_0^{\frac{\pi}{2}} \sin\theta d\theta \frac{4k_\perp(\theta) p_\perp(\theta)}{(k_\perp(\theta) + p_\perp(\theta))^2}. \tag{6.14}$$

This expression for the escape function has the advantage that it does not depend on the position of the zero point on the energy axis, since it is calculated using a difference of energy values, i.e. $\epsilon - \epsilon_c$ in Eq. (6.10). Moreover, by adopting an isotropic angular distribution for the wave vector of the excited electrons, we avoid the use of the parameters α and β.

6.2.3 Electron Cascades

Using the escape function described in the previous section, we face a similar problem as Hagstrum did when he first considered an isotropic distribution for the wave vectors of the excited electrons: the obtained yield spectra are too low, largely due to an insufficient amount of electrons at lower kinetic energies. This can be seen in Fig. 6.3, where we have plotted the kinetic energy distribution of the electrons emitted by the initial Auger neutralization. However, the low-energy electrons in the experimental yield spectra are often ascribed in the literature to a mechanism that Hagstrum did not consider in his model, i.e. electron cascades via electron-electron interactions. One of the first to implement this idea was Propst [11], who started from the distribution of electrons that did not escape after the Auger excitation, and allowed these electrons to interact with other electrons in the system through scattering events, once again producing electrons that can escape from the surface. He found that 50 % of the final yield was produced by the electron cascade process, making it an important contribution to the total yield. The concept has also been considered more recently by other authors, who first described it as an electron cascade process due to its iterative nature. According to Lorente et al. [27], the cascading electrons can account for 60% of the total secondary electron emission. Moreover, the electron cascades are found to be the source of the low energy electrons often missing in the calculated yield spectra of computational models.

We introduce our own model for the electron cascade process. Similar to the approach Hagstrum used to calculate the distribution of excited electrons from the neutralization of the incoming ion, we base our implementation of the electron cascades on the energy distributions of the interacting electrons. We begin by considering the energy distribution of the electrons that cannot escape:

$$N_c^{(2)}(\epsilon) = (1 - P_e(\epsilon))N_i(\epsilon). \tag{6.15}$$

In Hagstrum's model, these electrons simply do not contribute, and their energy has no influence on the SEE yield. Here, we approximate the scattering process using similar assumptions as Hagstrum made for the Auger neutralization, i.e. by considering the matrix elements of the transition to be constant, which makes the probability of a specific scattering event proportional to the density of states at the energy levels involved. The scattering process can be written as:

$$\epsilon_1 + \epsilon_2 \rightarrow \epsilon + \epsilon'. \tag{6.16}$$

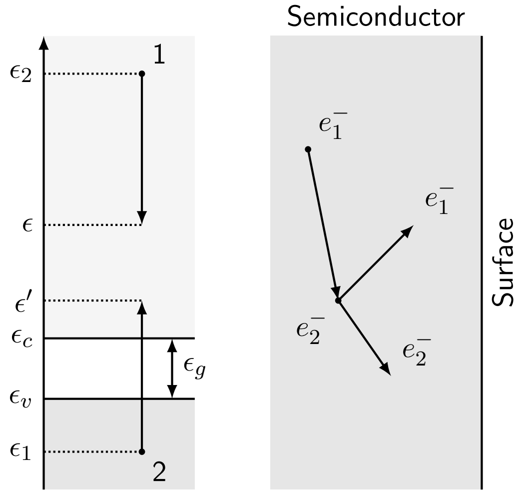


Figure 6.2: Energy diagram of a single step in the electron cascade process. The electron at energy ϵ_1 scatters on an electron in the valence band at energy ϵ_2, transferring sufficient energy for the second electron to excite into an unoccupied state of the conduction band at energy ϵ'.

where ϵ_1 is the energy of the excited electron before scattering, ϵ_2 is the energy of an electron in the valence band and ϵ, ϵ' are the energies of the two electrons after the scattering process (See Fig. 6.2). The distribution of excited electrons after scattering can then be calculated using:

$$\begin{aligned} N_i^{(2)}(\epsilon) &\sim \int_{\epsilon_c}^{\infty} d\epsilon' \int_{\epsilon_0}^{\infty} d\epsilon_1 \int_{\Delta_{\epsilon_v}} d\epsilon_2 D_c(\epsilon) D_c(\epsilon') N_c^{(2)}(\epsilon_1) D_v(\epsilon_2) \delta(\epsilon + \epsilon' - \epsilon_1 - \epsilon_2) \\ &= D_c(\epsilon) \int_{\epsilon_c}^{\infty} d\epsilon' D_c(\epsilon') \int_{\epsilon_0}^{\infty} d\epsilon_1 \int_{-\infty}^{\infty} d\epsilon_2 N_c^{(2)}(\epsilon_1) D_v(\epsilon_2) \delta(\epsilon + \epsilon' - \epsilon_1 - \epsilon_2) \\ &= D_c(\epsilon) \int_{\epsilon_c}^{\infty} d\epsilon' D_c(\epsilon') \int_{\epsilon_0}^{\infty} d\epsilon_1 N_c^{(2)}(\epsilon_1) D_v(\epsilon + \epsilon' - \epsilon_1) \\ &= D_c(\epsilon) \int_{\epsilon_c}^{\infty} d\epsilon' D_c(\epsilon') T_{ee}(\epsilon, \epsilon'). \end{aligned} \tag{6.17}$$

Where we have defined the scattering transform $T_{ee}(\epsilon, \epsilon')$ as

$$T_{ee}(\epsilon, \epsilon') = \int_{\epsilon_0}^{\infty} d\epsilon_1 N_c^{(2)}(\epsilon_1) D_v(\epsilon + \epsilon' - \epsilon_1), \tag{6.18}$$

Note that we only consider the excited electrons of $N_c^{(2)}(\epsilon)$ above the vacuum energy, as electrons with less energy can no longer produce electrons that can contribute to the secondary electron yield. Finally, because every scattering event results in two excited electrons, we normalize the new distribution of excited electrons to two times the number of electrons above the vacuum energy, prior to the scattering event:

$$N_i^{(2)}(\epsilon) = 2n_i \frac{D_c(\epsilon) \int_{\epsilon_c}^{\infty} d\epsilon' D_c(\epsilon') T_{ee}(\epsilon, \epsilon')}{\int_{\epsilon_c}^{\infty} d\epsilon D_c(\epsilon) \int_{\epsilon_c}^{\infty} d\epsilon' D_c(\epsilon') T_{ee}(\epsilon, \epsilon')}, \tag{6.19}$$

where

$$n_i = \int_{\epsilon_0}^{\infty} N_c^{(2)}(\epsilon')d\epsilon'. \tag{6.20}$$

In order to simulate the cascade process, these steps are iterated by once again considering the spectrum of the electrons which cannot escape and calculating the next energy distribution of excited electrons. These iterations continue until the yield difference between two iterations is smaller than 0.001 electrons per ion. At the final iteration step, we add the distributions of the yield that are obtained for each iteration to the original yield from Hagstrum's model. If we apply this approach to the yield calculation of He^+ ions on Ge(111), we obtain the results in Fig. 6.3, where we have plotted the number of emitted electrons N_0 versus their kinetic energy ϵ_k for several iterations. We can see that as we iterate the electron cascade process, the SEE spectrum quickly converges. The final distribution (iteration 3) has significantly more electrons at lower energies, and is a substantial improvement upon the initial yield distribution after Auger neutralization, when compared with the experimental results from Hagstrum [6].

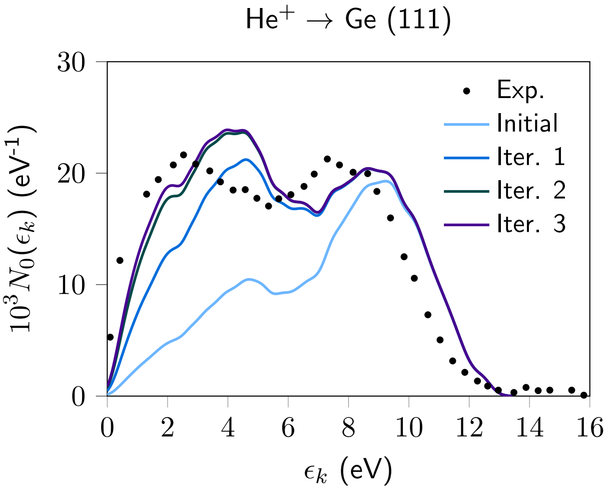


Figure 6.3: Calculated distribution of the kinetic energy ϵ_k of secondary electrons emitted for incident He^+ ions on Ge(111) at various iterations of the electron cascade process.

6.2.4 Comparison with experiment

The resulting yield spectra for incoming He^+ and Ne^+ ions on the reconstructed Ge(111) and Si(111) surfaces are plotted alongside the experimental results of Hagstrum [6] in Fig. 6.4. In his work, Hagstrum measured the total yield and kinetic energy distribution of the emitted electrons for several energies of the incoming ion on atomically clean, annealed (111) surfaces of Ge and Si. For the experimental results, we compare our calculated spectra with the results for low energy (10 eV) ions, for which the contribution of kinetic electron emission is negligible, as it is not considered in our model. Overall, the shape of the calculated yield spectra matches reasonably well with the experimental one[2]. The total SEE yield coefficients γ, obtained s by integrating the yield spectra (Eq. (6.8)), are given in Table 6.2. For both Si and Ge, the yield

---

[2] Note that we have applied a small Gaussian broadening to the final calculated yield spectra. This is largely for visual purposes, i.e. to facilitate the comparison between different spectra.

coefficients are larger for He$^+$ than for Ne$^+$, due to the larger ionization energy of helium. The yield is also found to be lower for Si(111) than Ge(111). The calculated yield coefficients are in fair agreement with the experimental values, with a slight overestimation for He$^+$ on Ge(111) and underestimation for all other ion/surface pairs.

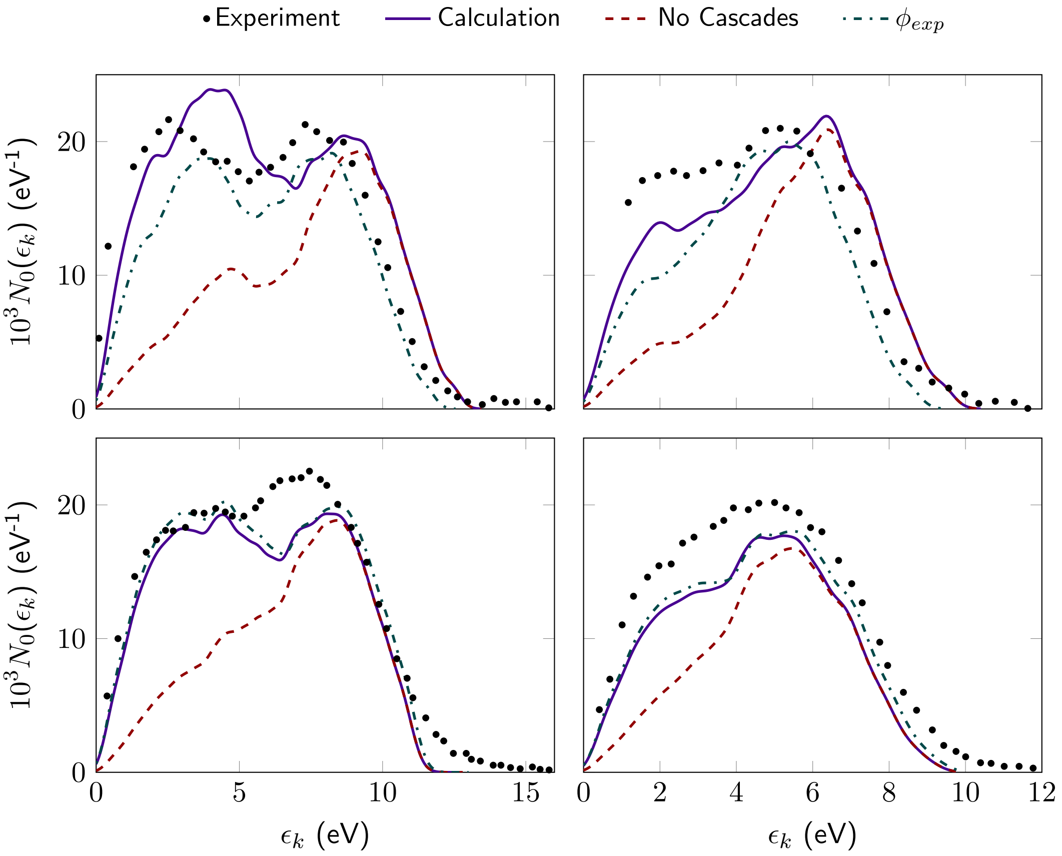


Figure 6.4: Experimental and calculated yield spectra for incoming He$^+$ and Ne$^+$ ions on Ge(111) and Si(111) surfaces. Also shown are the initial yield spectra, i.e. without the addition of the electron cascades process described in Sec. 6.2.3, as well as the calculated yield spectra when the vacuum level is adjusted using the experimental values for the work function in Table 6.1.

For Ge(111), the high energy tail of the calculated spectrum is shifted slightly to higher energies, whereas for Si(111), the tail is shifted in the opposite direction. These discrepancies can be attributed to the error on the calculated work function, which are compared with experiment for the Si(111) and Ge(111) surfaces in Table 6.1. We can see that the calculated work function is lower than the experimental result for Ge(111) and slightly higher for Si(111). Because the vacuum level determines the probability that an excited electron can escape, the work function has an important influence on the calculated yield. This influence becomes clear when we look at the yield spectra calculated when adjusting the vacuum level to the experimental work function, also plotted in Fig. 6.4. It is clear that by increasing the work function of Ge(111) to the experimental result, the tail of the calculated yield spectrum for both He$^+$ and Ne$^+$ shifts to lower energies, slightly below the tail of the experimental spectrum. For Si(111) the difference between the calculated and experimental work function is smaller, leading to only a minor shift in the spectra. It should be noted that although the use of the experimental work function leads to an improvement of the calculated yield spectra for the results presented here, the experimental

Table 6.1: Calculated work functions ϕ, compared with those calculated by De Waele et al. (ϕ_{ref}) [28] and experimental values (ϕ_{exp}).

Surface	ϕ (eV)	ϕ_{ref} (eV)	ϕ_{exp} (eV)
Ge(111)	4.548	4.569	5.00[a]
Si(111)	4.902	4.889	4.79[b]

[a] Polycrystalline sample from Michaelson. [29]

[b] Averaged value from Kawano. [30]

results for the work function can vary significantly, approximately with the same magnitude as the expected inaccuracy on *ab initio* calculated work functions [31]. Because the calculation of the SEE yield is sensitive to the vacuum level, the difficulty in accurately determining the work function adds another challenge to the calculation the ion-induced emission coefficient γ.

Table 6.2: Calculated yield γ, compared with the experimental values of Hagstrum (γ_{exp}) [6]. The fraction f_c of the contribution of the electron cascade process to the total yield, expressed as a percentage, as well as the calculated yield γ_ϕ using the experimental work function are also tabulated.

Surface	Ion	γ	$f_c(\%)$	γ_ϕ	γ_{exp}
Ge(111)	He$^+$	0.205	41	0.156	0.196
	Ne$^+$	0.125	28	0.102	0.138
Si(111)	He$^+$	0.166	32	0.176	0.188
	Ne$^+$	0.100	22	0.106	0.128

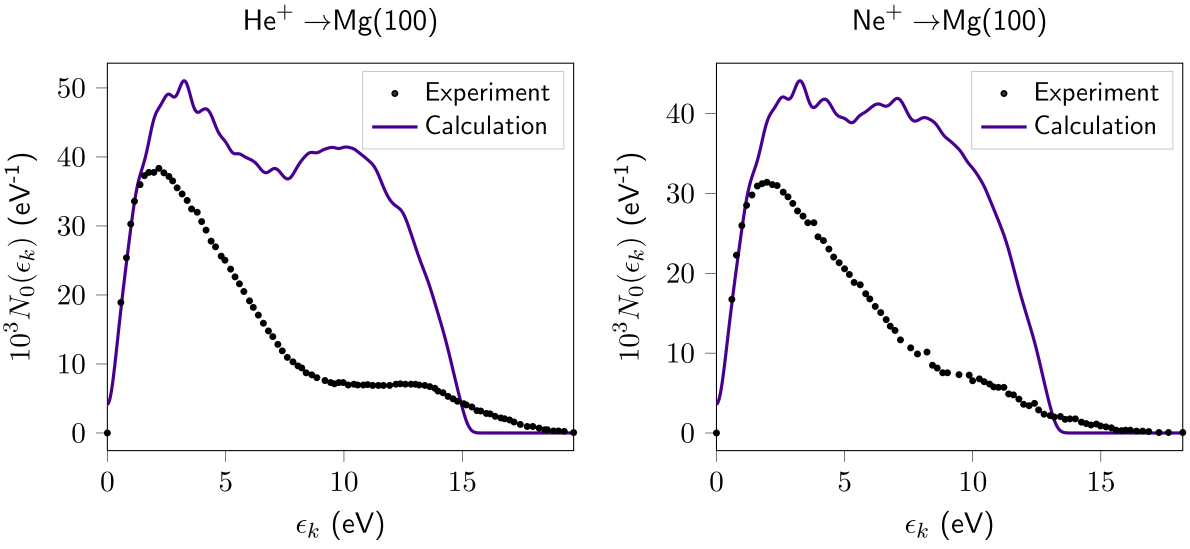
In order to determine the contribution of the electron cascade process, Fig. 6.4 also shows the yield spectrum of the initial neutralization step, i.e. without adding any of the escaped electrons due to electron cascades. For all of the ion-surface combinations, the electron cascades increase the number of low energy electrons substantially. Hence, by considering electron cascades, we can obtain the missing low energy electrons in Hagstrum's original approach, without introducing a large anisotropy in the initial direction of the excited electrons. If we look at the fraction f_c of the contribution of the electron cascades to the final SEE yield (Table 6.2), we can see that on average approximately 31% of the total yield is a result of electron cascade process. Moreover, the contribution is higher for Ge than for Si, most likely due to the larger band gap of Si, which results in a larger minimum energy loss for the scattered electrons. The higher value of f_c for He$^+$ impact compared to Ne$^+$ is because the larger ionization energy of He$^+$ allows for more iterations in the electron cascade process.

However, when the high energy tail of the yield spectrum is reproduced accurately, the model still underestimates the yield spectra. This can in part be explained by the anisotropy of the system caused by the incoming ion, which means that the distribution of the direction of the wave vectors $P_\Omega(\epsilon)$ is most likely not fully isotropic. This was Hagstrum's motivation for introducing the α and β parameters, in order to skew the direction of the excited electrons towards the surface. Our results indicate that Hagstrum most likely overestimated the anisotropy when fitting these parameters, as he did not consider the contribution of the electron cascades to the yield spectrum. However, the fact that we adopt an isotropic distribution could explain the underestimation of the yield in the results. Second, considering the matrix elements to be constant means that we treat all electronic energy levels on an equal footing. It is possible that including the calculation of the matrix elements could result in an increased participation of high energy electrons in the Auger neutralization, producing excited electrons with higher average energies. This extra energy would be passed on in the cascade process and hence result in an increase of the overall yield. The fact that our results match fairly well with experiment, however, is an indication that such effect would likely be small, and that considering the matrix elements to be constant is a reasonable approximation for Ge and Si.

6.3 Metals

So far, I have only considered semiconductors in the discussion of Hagstrum's model, as well as our adjustments. However, it is possible to extend the model to metals without having to significantly change the expressions of Sec. 6.2 by simply redefining the valence states density $D_v(\epsilon)$ as the density of the occupied states of the metal and $D_c(\epsilon)$ as the density of the unoccupied states. For metals then, $\epsilon_c = \epsilon_v = \epsilon_F$, with ϵ_F the Fermi level and χ in Eq. 6.11 becomes the work function ϕ.

However, it is well established that plasmon excitations play an important role in the interaction of ions and metallic surfaces [19, 32, 33], so any attempt to calculate γ for metals has to include a suitable implementation of these collective electron excitations at the surface. This also becomes clear when comparing our results for He$^+$ and Ne$^+$ incident on Mg(100) with the experimental result of Baragiola and Dukes [19] in Fig. 6.5. The experimental yield spectrum is severely overestimated, especially at higher energies. Both experimental spectra also present a rather distinct feature: above $\epsilon_k \approx 7-8$ eV there is a plateau in the yield spectrum for both He$^+$ and Ne$^+$, which is not reproduced by our model.



(a)

(b)

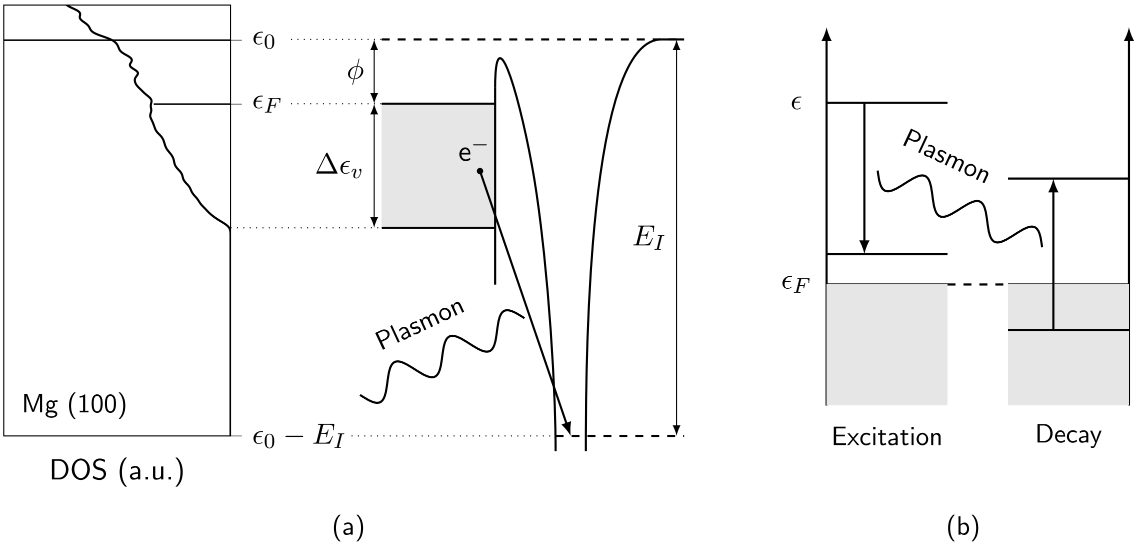
Figure 6.5: Comparison of the experimental secondary electron yield distribution with the calculated ones for He$^+$ (left) and Ne$^+$ (right) incident on Mg(100), without the inclusion of plasmonic effects.

6.3.1 Plasmons

We consider two mechanisms for inducing plasmonic excitations in our model. The first, surface resonance excitation, is introduced as a competing process for the Auger neutralization. The model for semiconductors considers the excess energy of the neutralization of every incoming ion to always result in the excitation of a single electron. There are, however, other processes that are in direct competition with the Auger mechanism. First, the released energy can produce a photon which is subsequently emitted from the material. However, this radiative process is considered negligible for low ion energies [12]. More important are collective charge density oscillations, i.e. plasmon excitations, first considered as a potential competing excitation mechanism by Apell [34]. Other theoretical work also confirms the importance of plasmon mechanisms for the ion neutralization process [12, 14, 35].

However, considering the electronic structure of Mg in Fig. 6.6a, a resonant excitation mechanism alone is not sufficient to explain the experimental electron energy distribution of He$^+$ on Mg in Fig. 6.5. The highest energy electrons in the excited spectrum are produced by electrons near the Fermi energy. The plasmon energy of Mg is approximately 10.6 eV, and the work function $\phi = 3.64$ eV. For a singly charged He$^+$ ion, the ionization energy is $E_I = 22.58$ eV. If the electrons close to the Fermi level in Mg are the ones to neutralize the incoming ion, an energy of up to 18.94 eV is released. Since the width of the density of occupied states is $\Delta\epsilon_v = 6.96$ eV, the lowest energy released by the ion neutralization is $E_I - \Delta\epsilon_v - \phi = 11.98$ eV. If the plasmon activity is only described by a resonant process, we expect very few plasmon excitations, even if the plasmon resonance peak is significantly broadened by the short plasmon lifetime.

There is, however, a second plasmon excitation mechanism which is not resonant in its nature. As the excited electrons travel through the material, they leave behind a wake of electron density fluctuations, leading to the possibility of energy loss through volume plasmon excitations in case the electrons have sufficient energy [36]. This excitation process does not have a resonance condition, i.e. as long as the excess electron energy is above the plasmon energy, the electron can excite a plasmon, losing energy in the process [32].



(a)

(b)

Figure 6.6: Energy diagrams of resonant surface (a) and volume plasmon (b) excitations. (a) also shows the DOS of Mg (110), with the energies aligned to the diagram.

In order to include the plasmonic excitations in our model, we describe them as Poisson point processes [37], i.e. the intervals between events follow an exponential distribution [38]:

$$f_{T_p}(t) = \lambda_p e^{-\lambda_p t}, \tag{6.21}$$

for a process p, i.e. volume plasmon excitation (vp), surface plasmon excitation (sp) or Auger neutralization (aug). The excitation rate of volume plasmons with energy E_{vp} can be related to the dielectric function $\varepsilon(\omega)$ through the volume loss function L [39]:

$$\lambda_{vp}(E_{vp}) \sim L(E_{vp}) = \mathrm{Im}\left[-\frac{1}{\varepsilon(E_{vp})}\right] \tag{6.22}$$

$$\lambda_{vp}(E_{vp}) = c_{vp} \cdot \mathrm{Im}\left[-\frac{1}{\varepsilon(E_{vp})}\right], \tag{6.23}$$

where c_{vp} is a proportionality constant. Each of the possible energy losses E_{vp} are considered as competing Poisson point processes. However, as the excited electron cannot wind up in an occupied state, it cannot lose more energy than its energy relative to the Fermi level $\epsilon - \epsilon_F$. Writing the average travel interval of the excited electrons as t_e, the probability for an electron at energy ϵ of exciting a plasmon with energy $E_{vp} < \epsilon - \epsilon_F$ is (see Appendix B.1):

$$P_{vp}(\epsilon, E_{vp})dE_{vp} = \frac{c_{vp}L(E_{vp})dE_{vp}}{\int_0^{\epsilon-\epsilon_F} c_{vp}L(E)dE}\left[1 - e^{-\left(\int_0^{\epsilon-\epsilon_F} c_{vp}L(E)dE\right)t_e}\right] \tag{6.24}$$

$$= \frac{L(E_{vp})dE_{vp}}{\int_0^{\epsilon-\epsilon_F} L(E)dE}\left[1 - e^{-\left(\int_0^{\epsilon-\epsilon_F} L(E)dE\right)c_{vp}t_e}\right] \tag{6.25}$$

The product $\kappa_v = c_{vp} \cdot t_e$, the combination of the prefactor of the rate λ_{vp} and the average travel interval of the excited electrons, is a measure of the likelihood of plasmon excitations during the SEE process. We treat it as a parameter of the model.

For the surface plasmons, a plasmon excitation can occur instead of a single-electron excitation when the energy released in the neutralization of the incoming ion is close to that of the surface plasmon. In order to determine the possibility of a surface plasmon excitation, both processes are once again modeled as Poisson point processes, i.e. by considering the exponential distributions

$$f_{T_{sp}}(E_{sp}, t) = \lambda_{sp}(E_{sp})e^{-\lambda_{sp}(E_{sp})t} \tag{6.26}$$

$$f_{T_{aug}}(t) = \lambda_{aug}e^{-\lambda_{aug}t}. \tag{6.27}$$

The probability that the plasmon excitation will occur before the Auger Neutralization is then (see Appendix B.1):

$$P_{sp}(E_{sp}) = \Pr\{T_{sp} < T_{aug}\} = \frac{\lambda_{sp}(E_{sp})}{\lambda_{sp}(E_{sp}) + \lambda_{aug}} \tag{6.28}$$

The surface plasmon excitation rate is calculated from the dielectric response of the surface using the surface energy loss function [39]:

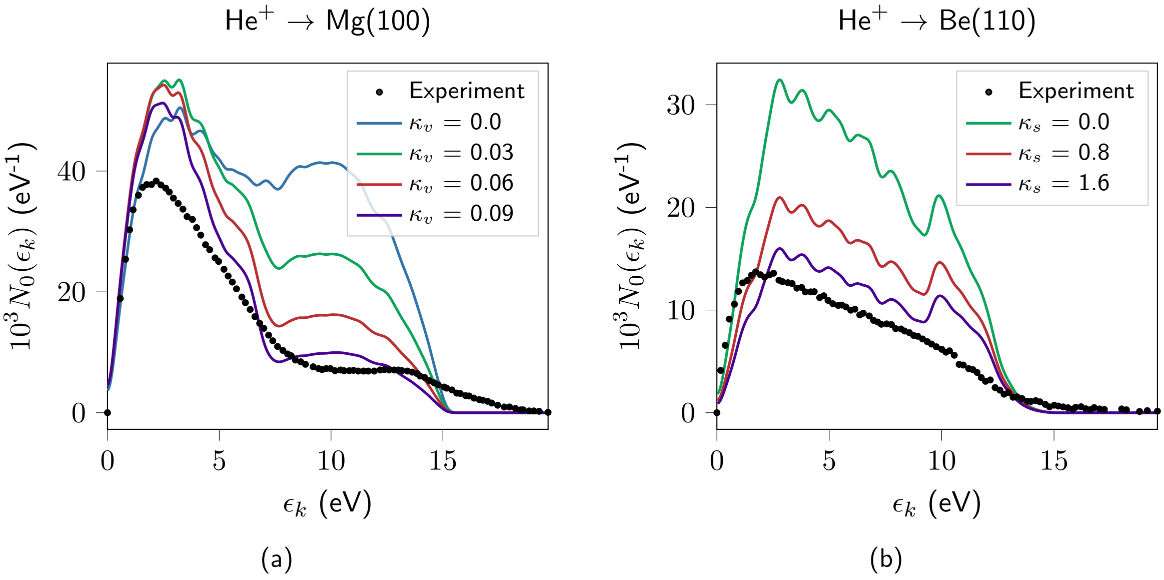
$$\lambda_{sp}(E_{sp}) = c_{sp} \cdot \mathrm{Im}\left[-\frac{1}{\varepsilon(E_{sp}) + 1}\right], \tag{6.29}$$

where c_{sp} is once again a proportionality constant. The expression for the probability of a plasmon excitation when an energy E_{sp} is transferred to the incoming ion becomes:

$$P_{sp}(E_{sp}) = \frac{c_{sp} \cdot \mathrm{Im}\left[-\frac{1}{\varepsilon(E_{sp})+1}\right]}{c_{sp} \cdot \mathrm{Im}\left[-\frac{1}{\varepsilon(E_{sp})+1}\right] + \lambda_{aug}} = \frac{\frac{c_{sp}}{\lambda_{aug}}\mathrm{Im}\left[-\frac{1}{\varepsilon(E_{sp})+1}\right]}{\frac{c_{sp}}{\lambda_{aug}}\mathrm{Im}\left[-\frac{1}{\varepsilon(E_{sp})+1}\right] + 1} \tag{6.30}$$

where the ratio $\kappa_s = c_{sp}/\lambda_{aug}$ is treated as a second parameter of the plasmon model[3]. In this description, it is assumed that the Auger neutralization rates are similar for the various ion material combinations and independent of the energy transferred to the secondary electron.

A more detailed derivation of Eqs. (6.24) and (6.30) can be found in Appendix B, along with a description of the exact implementation of the plasmonic excitation in the model. In short, the surface plasmon excitation probability (Eq. (6.30)) is used to remove a fraction of the energy distribution passed to the valence electrons in the Auger neutralization. The implementation of the volume plasmon is more complicated, as an electron traveling to the surface can excite volume plasmons with a range of energies E_{vp}. Moreover, volume plasmons decay into single particle excitations [40], and hence the spectrum of excited plasmons has to be considered for calculating a new distribution of excited electrons resulting from plasmon decay.



(a) (b)

Figure 6.7: Influence of the volume (a) and surface (b) plasmon parameter on the SEE spectra of Mg (100) and Be(001), respectively. The experimental data are taken from [32].

Figure 6.7 shows the influence of increasing the volume and surface plasmon parameters on the spectra of He$^+$ incident on Mg(100) and Be(001) surfaces, respectively. As previously discussed, surface plasmons are unlikely to be excited during the neutralization of He$^+$ on Mg, so this allows us to isolate the influence of volume plasmons. Similarly, the large plasmon frequency of Be (See Fig. 6.8) means that surface plasmons, having a lower frequency than volume plasmons, play a much larger role in the spectrum of Be. For Mg, volume plasmon excitations result in a large reduction of the yield spectrum at higher energies. By increasing the participation of volume plasmons through the parameter κ_v, we are able to match the experimental spectrum much more closely, largely reproducing the “plateau” feature a higher energies which was missing from our previous results. For the Be spectrum, increasing surface plasmon resonances through κ_v results in an overall reduction of the yield spectrum, most likely due to the broad plasmon peak in the surface energy loss function of Be (Fig. 6.8).

Although it would be possible to fit the plasmon parameters κ_v and κ_s for each ion/surface combination, similar to how Hagstrum fitted the escape function parameters α and β, this would

---

[3] We take this ratio to be the parameter - instead of its inverse - so that the amount of plasmonic activity is directly proportional to the parameter, which is more intuitive. This also allows us to turn off surface plasmons by setting the parameter to zero.

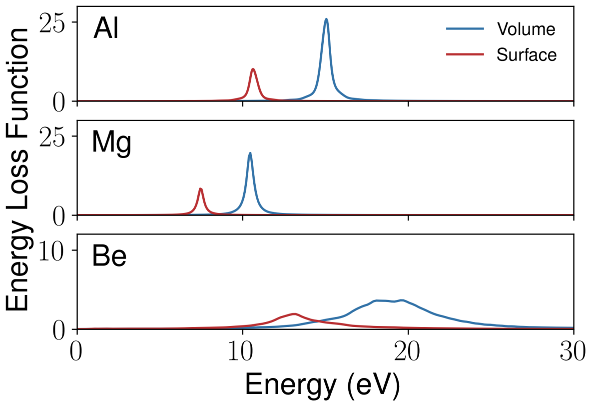


Figure 6.8: Volume and surface energy loss functions of Al, Mg an Be, calculated from the dielectric function of the bulk system of each element.

remove any predictive capability of the model. Hence, we have fitted a *single* set of parameters to the available quantitative experimental data[4] for He^+ and Ne^+ ions incident on Al and Mg from Baragiola et al. [32], as well as Takeishi and Hagstrums results for Ni(111) [41] and Cu(110) [42]. Unfortunately, as Baragiola et al. did not specify the surface orientation of each metal, we chose to compare the experimental result to an average of our calculated results. Based on an interactive Jupyter notebook, we first explored the influence of the parameters on the calculated parameters for each surface in order to ascertain a reasonable range for the fitting procedure. Next, we take the difference of the experimental and calculated values for each experimental data point and sum over their absolute values for each spectrum, normalizing each to the number of experimental data points in order to give each spectrum the same weight. Based on this analysis, the optimal values for the plasmon parameters are $\kappa_v = 0.105$ and $\kappa_s = 1.55$.

6.3.2 Comparison with experiment

Now that the model plasmon parameters κ_v and κ_s are fixed, we first compare the calculated secondary electron emission spectra for the surfaces of the metals for which experimental spectra are available. The results for the work function and calculated yields are compared with the experimental values [31, 32, 42] in Table 6.3. The calculated work functions are fairly close to the experimental ones for most surfaces. For Be, there is a significant difference in the calculated work function of the various surfaces. As was noted by Michaelson [29] and Green and Bauer [43], there has been some debate on the work function of Be, for which both values around 4 eV and 5 eV have been reported. This may be related to the observed large spread in calculated work function values between the surfaces.

There is also a good agreement between the calculated and experimental yield of He^+ an Ne^+ ions on Al, Mg and Ni, both for the total yield as well as the yield spectra (Fig. 6.9). For Al, our model tends to underestimate the total secondary electron emission. Comparing the full emission spectra, there is a feature of the experimental spectra that is missing in our computational results: the high energy tail of the spectrum. Hagstrum [6] already investigated the influence of the kinetic energy on the spectrum of He^+ on Ge(111), and found that increasing the kinetic energy of the ion leads to a broader tail of the spectrum at high energies. A similar result was found for Ne^+ on Al by Baragiola et al. [19], who also demonstrated an overall gain in electron

---

[4] Note that there are more experimental spectra available. I discuss the reason for their exclusion in the next section.

yield when the kinetic energy is increased. Hagstrum introduced a broadening in the Auger transform to simulate the kinetic effects, however, as the difference is rather minor, we simply stick to a general 0.2 eV Gaussian broadening of the yield spectra for visual purposes, similar to the spectra presented for Ge and Si.

In contrast with the results for Al and Mg, there is a much larger difference in the yield results for the surfaces of Be (Fig. 6.10), which can be expected considering the larger variation in the surface work function. Baragiola et al. [32] did not specify the Be surface for their SEE results, which complicates the quantitative comparison of our computational results with experiment, and hence we did not include them in our fitting procedure for the plasmon parameters. Qualitatively, however, the overall shape of the He^+ yield spectra are similar, and the yield of the (110) surface also matches well with the experimental results quantitatively. It is reassuring to see that a single set of plasmon parameters is able to reproduce a lot of varied experimental data adequately, which to some extent validates our approach in treating them as model parameters.

Table 6.3: Calculated work functions ϕ and SEE yields γ for each of the surfaces, compared with the available experimental data. For the work function of Mg and Be we do not have specific data for each surface. Similarly, we only know the surface for the yield results of Cu and Ni.

Element	Surface	ϕ (eV)	ϕ_{exp} (eV)	γ^{He}	γ^{He}_{exp}	γ^{Ne}	γ^{Ne}_{exp}
Al	(111)	4.05	4.28	0.193	0.231	0.157	0.202
	(100)	4.26	4.36	0.175	"	0.141	"
	(110)	4.04	4.21	0.195	"	0.159	"
Mg	(100)	3.65	3.66	0.286	0.257	0.222	0.202
	(001)	3.7	"	0.242	"	0.195	"
	(110)	3.49	"	0.316	"	0.245	"
Ni	(111)	5.05	5.28	0.173	0.175	0.142	0.124
	(100)	4.91	5.23	0.181	-	0.15	-
	(110)	4.67	4.64	0.203	-	0.171	-
Cu	(110)	4.4	4.48	0.145	0.155	0.12	0.12
	(111)	4.77	4.91	0.125	-	0.101	-
	(100)	4.52	4.57	0.139	-	0.114	-
Be	(100)	4.5	4.98	0.083	0.117	0.054	0.095
	(001)	5.31	"	0.049	"	0.029	"
	(110)	3.82	"	0.119	"	0.081	"
W	(111)	4.15	4.38	0.161	0.292	0.124	0.224
	(100)	4.1	4.57	0.158	"	0.121	"
	(110)	4.79	5.31	0.117	"	0.082	"

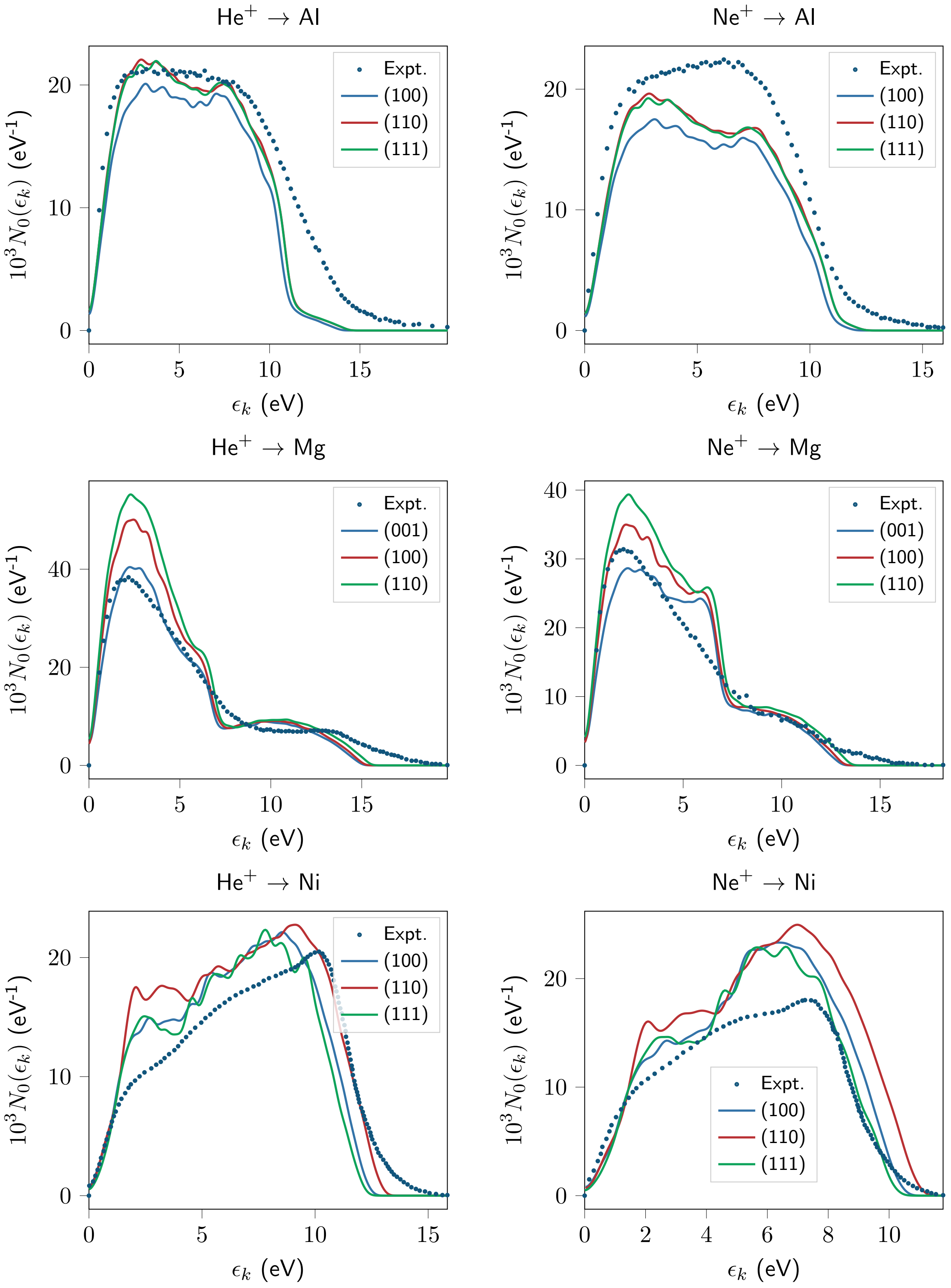


Figure 6.9: Experimental and calculated yield spectra for incoming He$^+$ and Ne$^+$ ions on the chosen surfaces of Al, Mg and Ni.

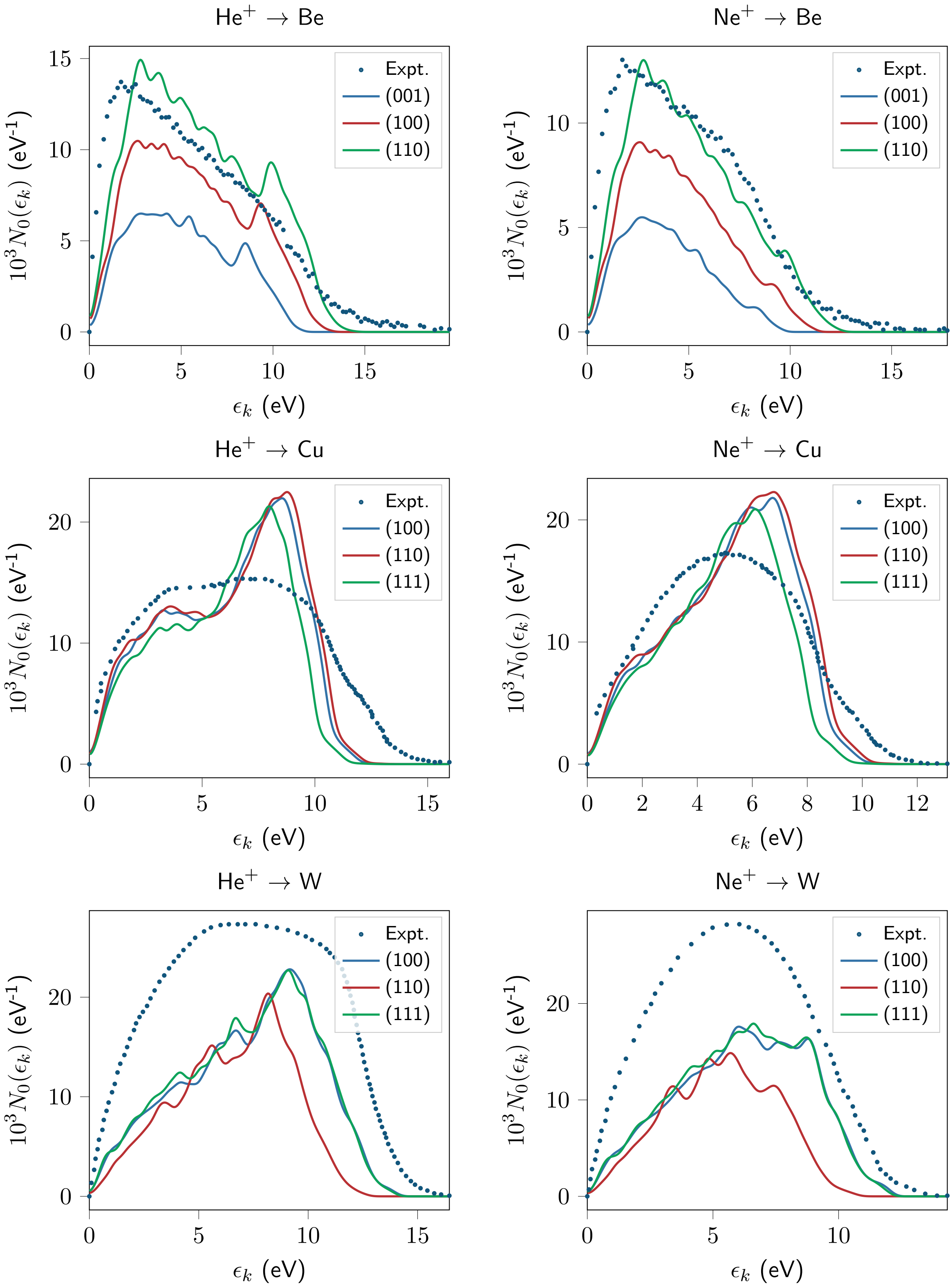


Figure 6.10: Experimental and calculated yield spectra for incoming He$^+$ and Ne$^+$ ions on the chosen surfaces of Be, Cu and W.

For Cu(110), the total yield is quite close to the experimental value, but the spectrum shape differs quite significantly from the experimental one, independently of the choice of plasmon parameters. As discussed in Section 6.2.1, the Auger neutralization rate λ_{aug} can depend significantly on the electronic state (s, p, d) [21]. In fact, Goebl et al. had specifically investigated the influence of the electron orbital on the neutralization rate for noble metals (Cu, Ag and Au, and found that the rate due to d electrons can be an order of magnitude higher compared to s or p). In order to analyze whether this has a large effect on the yield, we implemented an input argument in the yield calculation that allows the user to give more weight to the d orbitals in the neutralization of the incoming ion. However, as we can see in Fig. 6.11, even when we increase the weight of the d-orbitals by an order of magnitude, the influence on the final yield spectrum is negligible. Looking at the orbital projected density of states, this isn't surprising, as the electronic states close to the Fermi level already largely correspond to d orbitals. Currently, it is unclear as to why there is such a discrepancy between the calculated and experimental yield distributions of Cu(110).

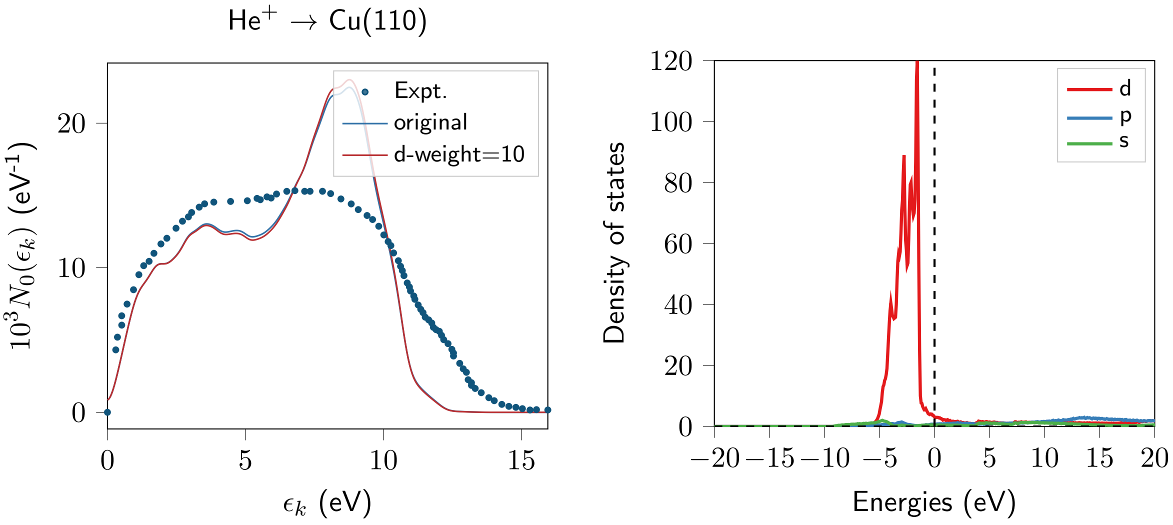


Figure 6.11: Influence of increasing the likelihood of d-electrons neutralizing the incoming ion.

Note that for W, the calculated spectrum is consistently lower than the experimental one from Hagstrum. At first glance, this seems to be related to the work function, as the position of the high energy tail is lower for the computational result. However, according to the results in Table 6.3, the computational work functions are lower for each surface compared to experiment. A possible explanation is that the W surface was contaminated during Hagstrum's measurement, lowering the effective work function. Because of this discrepancy, we did not consider W when fitting the plasmon parameters.

I end this section with a couple of notes. So far the discussion has been limited to He$^+$ and Ne$^+$ ions. For Ar$^+$, the calculated yield spectra are significantly lower than the experimental ones for most metals. This can be related to the fact that we do not consider resonance neutralization (RN) (See Fig. 6.1) in our calculations. Although Auger neutralization (AN) is believed to be the dominant process, neglecting RN can have a larger effect for ions with a lower ionization energy. To understand this, consider what happens when an electron first neutralizes the incoming ion via a resonant tunneling process, and subsequently is excited to an energy above the vacuum level via Auger de-excitation (AD). Compared to an electron that is excited inside the surface via AN, this electron is excited in the (approximately) spherical potential well of the ion. If the energy of the electron is higher than the vacuum level, the escape probability of this electron is much higher than that of an electron excited inside the surface. This difference in escape probabilities is especially large when the excited energy level is only slightly above the vacuum level, as in this case the excited electron in the surface can only escape in case its wave vector is directed almost straight at the surface.

For ions with a low ionization energy, a much larger fraction of the electrons are excited to an energy close to the vacuum level. If the electrons are excited via AN, their escape probability is low, especially when compared to RN+AD. Hence, even though only a small fraction of ions are neutralized via the latter process, the resulting electrons can make up a significant portion of the final electron yield. Although we could introduce the RN+AD process via another fitting parameter, it would be better to try and determine the likelihood of either process by calculating their rates for a model system as done by Goebl et al. [21]. This is beyond the scope of this work.

6.3.3 High-throughput results

Although it is gratifying to see that our model is able to reproduce both the experimental SEE yield values and spectra of many elemental surfaces, the purpose of developing the model is to apply it to predict the yield of surfaces for which no experimental results are available. This section presents a high throughput screening of a list of elemental surfaces spanning the periodic table, based on the workflow discussed in Section 3.3.4. As is described in that section, the calculation of the surface properties for our model input requires a choice of sufficient atomic layers and vacuum thickness. Fortunately, here we can rely on the extensive testing of De Waele et al. [31], instead of performing the necessary convergence tests ourselves. Based on the details provided in the supplemental material of their paper, we have calculated the required input[5] for our version of Hagstrum's model for all of the tabulated surfaces. Using the model plasmon parameters from our fitting procedure, we have then calculated the total yield for He^+ and Ne^+ for each surface. Figure 6.12 shows a map of the averaged total yield for He^+ on the periodic table.

The average yield of the group IA elements is excessively high. This is a result of the combination of the low work functions of the surfaces of these elements as well as their small width of the density of occupied states $D_v(\epsilon)$. The latter increases the yield because the average energy of an electron is higher in case the electrons are all close to the Fermi level. Hence, due to these two properties, the electrons only have to overcome a relatively small barrier when trying to escape, and there are always relatively high energy electrons that participate in both the Auger neutralization and the electron scattering processes. Most likely, our model is rather optimistic in its treatment of the electron scattering, as it does not consider the depth of the scattering electrons, allowing them a chance to escape at every iteration. This results in an overestimation of the yield for these elements, as well as other elements with similar surface properties (Ca, Y, Sc).

On the other side of the periodic table, elements with full d orbitals and only a couple of electrons in the s and/or p orbitals (group 11-13, excluding Al) have a noticeably lower yield. This is connected to the electronic structure of these elements near the Fermi level. Figure 6.14 shows the projected DOS of Ni (100), Cu (100) and Zn (100). For all of these surfaces, most of the occupied states near the Fermi level correspond to d states. However, for Zn (100) these lie significantly below the Fermi level, which means that the average energy of electrons that

---

[5] Note that this input cannot simply be extracted from the output files of De Waele et al., as our model needs the density of states for a large number of unoccupied bands.

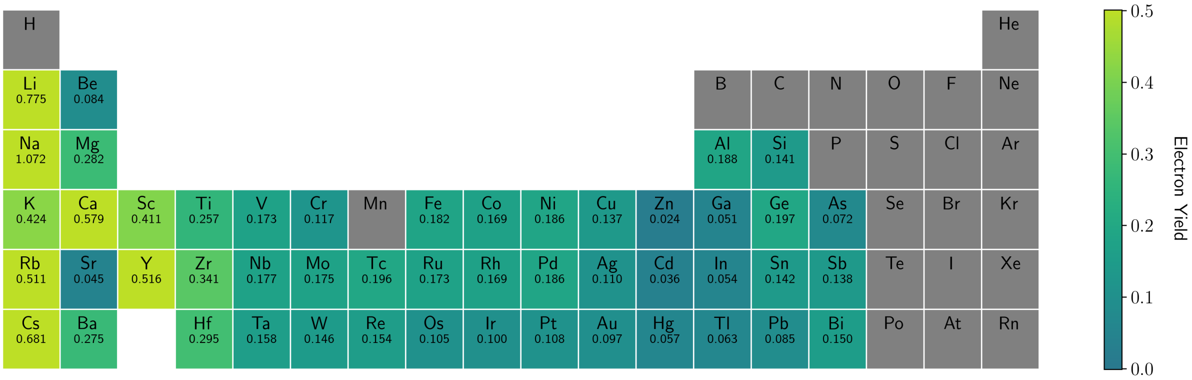


Figure 6.12: Average yield results for He$^+$ ions on the surfaces of the studied elements.

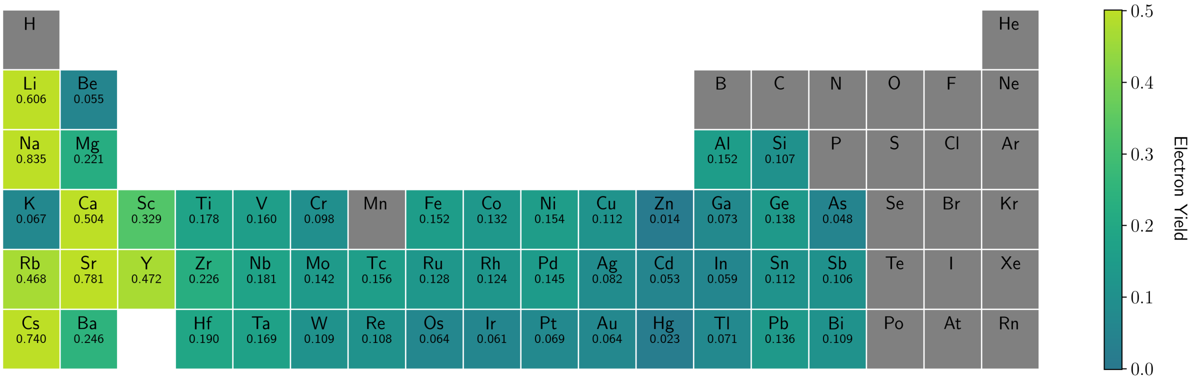


Figure 6.13: Average yield results for Ne$^+$ ions on the surfaces of the studied elements.

participate in the Auger neutralization and electron scattering processes is relatively low. This reduces the average energy of the excited electron, which results in a lower chance of escape and hence a lower yield. The electronic structures of other elements with full *d* orbitals is similar, resulting in a lower yield for groups 11-13. Cu, and by extension Ag and Au, suffer from a similar effect, but to a lesser extent because the *d* states are closer to the Fermi level. Finally, this also explains the relatively low results for He$^+$ ions on K, Sr and Ba.

The yield for Ne$^+$ results in a reduction for the yield for most elements compared to He$^+$, which is to be expected considering the lower ionization energy of Ne$^+$. There are some notable exceptions, however, such as Sr, Sn, Pb and Cs. The first three can be explained by the fact that due to the lower ionization energy of Ne$^+$, the deep *d* orbitals can no longer neutralize the incoming ion, which means that only the higher energy *s* and *p* electrons take part, resulting in a higher average energy of the excited Auger electrons. For He$^+$ on Cs, there is an increased resonance between the energy released during the neutralization and the surface plasmon excitation. Interesting here is also the low average yield value for Ne$^+$ on K. This is due to the fact that the $3p$ orbitals of K are just barely able to neutralize the incoming ion, resulting in a large fraction of electrons in the excited density that are just above the vacuum level, and hence a severely reduced yield.

So far, I have found no results in the literature where the yield of Ne$^+$ is higher than He$^+$ for an element. It is conceivable that these lower shell electrons have very little overlap between their wave function and that of the lowest occupied state of the incoming ion, resulting in a low likelihood of them neutralizing the ion. In this case, the participation of the higher energy *s* and *p* states increases the yield for He$^+$ and can restore the expected trend. However, the same can be said about the $3d$ orbitals in Cu compared to its $4s$ orbitals, and Goebl et al. [21] found that the *d* orbitals actually have a higher rate of neutralizing the ion. In future work, it might be

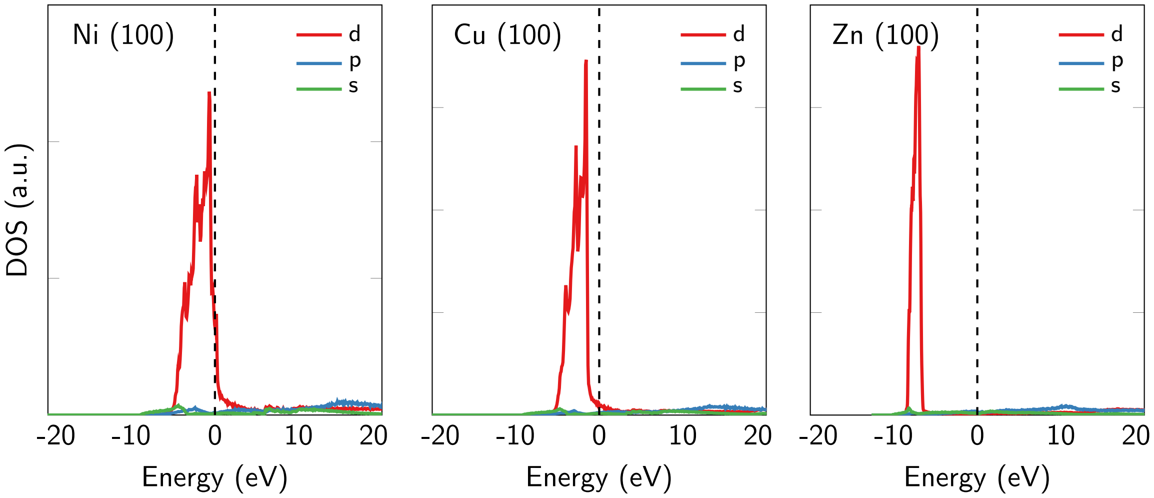


Figure 6.14: Projected density of states of Ni (100), Cu (100) and Zn (100).

interesting to supplement our model with rate calculations similar to those performed by Goebl et al. for other elements, so we can weigh the contribution of each orbital accordingly.

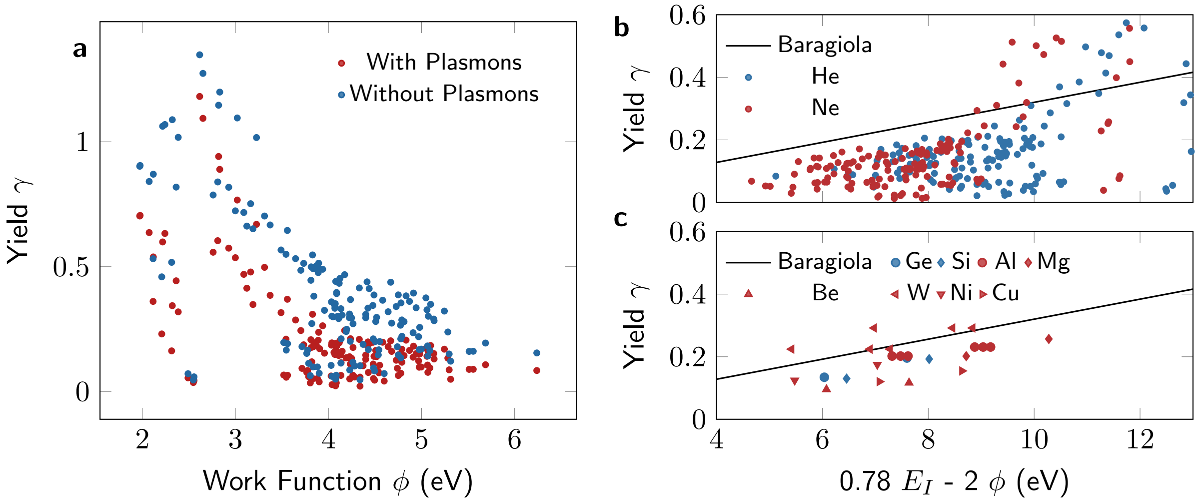


Figure 6.15: (**a**) Calculated yields of all surfaces versus their work function. The regular data, including our model for the plasmonic excitations, is shown in red. For comparison, we have also added the results without plasmonic effects ($\kappa_v = \kappa_s = 0$) in blue. Calculated (**b**) and experimental (**c**) secondary electron yields compared to the fit of Baragiola et al. [44]. The experimental values are tabulated in Table 6.3. For materials where we do not know the surface of the corresponding yield result, we have plotted one data point for each experimental work function value.

Figure 6.15a plots the calculated yield for each surface versus its work function, for both the model with and without plasmons. It is clear that the plasmons have a significant influence on the yield for most materials. The work function also has an important influence on the yield, which is to be expected, however we also find that for a single value of the work function there can be a wide range of yields. The fact that this is true for both the results with and without plasmons indicates that this is not simply a consequence of the electronic response of the material, but most likely due to differences in the density of states of the various surfaces.

Finally, we compare both our calculated yield spectra and work functions, as well as the experimental results tabulated in Table 6.3, with an empirical fit from Baragiola et al. [44] in Fig. 6.15b-c. It is clear that for surfaces with higher work functions (left side of the figure), our calculated yields lie consistently below Baragiola's fit. This is also the case for most experimental data, however, with the sole exception of tungsten, for which we observed a significant discrepancy between our calculated spectra and the experimental ones in Section 6.3.2. For surfaces with lower work functions (right side of Figs. 6.15b-c), the calculated yield can be significantly higher than the value corresponding to Baragiola's fit. This is another indication that our model overestimates the yield of these surfaces. Besides the results of Hagstrum, Baragiola et al. relied on the results of Oechsner [45] and Arifov [46]. The former, however, used Ar^+ ions with a kinetic energy of 1 keV, which can significantly increase the electron yield compared to relatively slow ions (4-100 eV). As our model does not consider kinetic mechanisms in the calculation of the yield, this could be one explanation as to why our results lie below Baragiola's fit. Moreover, considering the importance of plasmon excitations, and the highly material dependent energy loss spectra in Fig. 6.8, it seems unlikely that an accurate electron yield can be obtained from a linear fit relying solely on the ionization energy E_I and work function ϕ.

6.4 Conclusions and outlook

Starting from Hagstrum's model, I have presented a method for calculating the SEE yield of incoming ions on semiconductor surfaces, and have applied it to incident He^+ and Ne^+ ions on the (111) surface of Ge and Si. Determining the required input using *ab initio* DFT calculations, as well as using a different expression for the escape function, removes the parameters Hagstrum used in his model. Furthermore, introducing electron cascades recovers the low energy electrons missing in the initial spectra, without the introduction of new parameters. Using this approach, our model is able to produce spectra that match closely to experiment for Ge (111) and Si (111), both in shape as well as the total calculated yield.

For metals, we have extended the model with an implementation of plasmonic excitations. Here, the model makes a distinction between surface and volume plasmons, both of which play an important role in the secondary electron emission process. The calculated spectra match reasonably well with experiment, but there are some discrepancies, e.g. for Cu and W. Moreover, when applying the model to group I-II elemental surfaces, the calculated yield is very high, and most likely an overestimation. Finally, due to the electronic structure of the surface, it is possible that for some elements the calculated yield is lower for He^+ than Ne^+, despite the higher ionization energy of He^+. Such a trend has so far not been observed in the literature.

Although the electron cascades are an important process that allows our model to retrieve the low energy electrons missing in Hagstrum's original model, it is possible that our idealistic representation of the process overestimates the number of cascading electrons that contribute to the total yield. For many elements, this overestimation can be balanced by the fact that we use an isotropic distribution for the wave vector of the excited electron. Hagstrum was right to consider the distribution to be more skewed along the normal of the surface, but most likely overemphasized this effect in order to obtain sufficient electrons at low energy. In our case, the fully isotropic distribution means that every excited electron has a lower chance to escape than in Hagstrum's case. This is also true for the I-II group elements, but here the idealistic electron cascades implementation combined with the low work function of their surfaces results in a far greater overestimation.

Calculating accurate SEE yield coefficients from first principles is a challenging undertaking, both due to the complexity of the processes involved, as well as their sensitivity to the vacuum level. The model in its current state offers a pragmatic and effective approach for calculating yield spectra, but can still benefit from several improvements. First, using an approach similar to Goebl et al. [21] to calculate transition rates for atomic orbitals can be used to weigh the participation of s, p and d electrons. Moreover, if it is possible to extend such calculations to the comparison of the Auger- and resonant neutralization process, it can allow us to include the resonant mechanism in our calculation without introducing more parameters. Second, both the implementation of the electron cascades and the plasmonic excitations can be improved by including the wave vector of the electron in the calculation. Finally, the ionization energy can be better described by considering a range of energies that depend on the distance of the ion to the surface when it is neutralized. However, this requires an accurate function for both the ionization energy as well as the probability of neutralization versus the distance.

Bibliography

[1] D. Depla et al. “Discharge voltage measurements during magnetron sputtering”. *Surf. Coatings Technol.* 200 (2006), pp. 4329–4338. DOI: `10.1016/j.surfcoat.2005.02.166`.

[2] D. Depla, S. Mahieu, and R. De Gryse. “Magnetron sputter deposition: Linking discharge voltage with target properties”. *Thin Solid Films* 517 (2009), pp. 2825–2839. DOI: `10.1016/j.tsf.2008.11.108`.

[3] Y. Motoyama and F. Sato. “Calculation of secondary electron emission yield γ from MgO surface”. *IEEE Trans. Plasma Sci.* 34 (2006), pp. 336–342. DOI: `10.1109/TPS.2006.872443`.

[4] Y. Cho et al. “First-principles study on secondary electron emission of MgO surface”. *J. Appl. Phys.* 101 (2007), p. 083710. DOI: `10.1063/1.2721857`.

[5] H. D. Hagstrum. “Electron ejection from Mo by He^+, He^{++}, and He_2^+”. *Phys. Rev.* 89 (1953), pp. 244–255. DOI: `10.1103/PhysRev.89.244`.

[6] H. D. Hagstrum. “Auger electron ejection from germanium and silicon by noble gas ions”. *Phys. Rev.* 119 (1960), pp. 940–952. DOI: `10.1103/PhysRev.119.940`.

[7] P. S. Kothnur, X. Yuan, and L. L. Raja. “Structure of direct-current microdischarge plasmas in helium”. *Appl. Phys. Lett.* 82 (2003), pp. 529–531. DOI: `10.1063/1.1540246`.

[8] P. S. Kothnur and L. L. Raja. “Two-dimensional simulation of a direct-current microhollow cathode discharge”. *J. Appl. Phys.* 97 (2005). DOI: `10.1063/1.1849816`.

[9] H. D. Hagstrum. “Theory of auger ejection of electrons from metals by ions”. *Phys. Rev.* 96 (1954), pp. 336–365. DOI: `10.1103/PhysRev.96.336`.

[10] H. D. Hagstrum. “Theory of auger neutralization of ions at the surface of a diamond-type semiconductor”. *Phys. Rev.* 122 (1961), pp. 83–113. DOI: `10.1103/PhysRev.122.83`.

[11] F. M. Propst. “Energy distribution of electrons ejected from tungsten by He^+”. *Phys. Rev.* 129 (1963), pp. 7–11. DOI: `10.1103/PhysRev.129.7`.

[12] A. A. Almulhem and M. D. Girardeau. “Theory of ion neutralization at metal surfaces by surface plasmon excitation”. *Surf. Sci.* 210 (1989), pp. 138–162. DOI: `10.1016/0039-6028(89)90108-8`.

[13] N. Lorente and R. Monreal. “Neutralization of slow He^{2+} on metal surfaces: theory for Auger and cascade electron emission”. *Surf. Sci.* 303 (1994), pp. 253–265. DOI: `10.1016/0039-6028(94)90637-8`.

[14] R. Monreal and N. Lorente. “Dynamical screening in Auger processes near metal surfaces”. *Phys. Rev. B* 52 (1995), pp. 4760–4763. DOI: `10.1103/PhysRevB.52.4760`.

[15] C. R. Monreal. “Auger neutralization and ionization processes for charge exchange between slow noble gas atoms and solid surfaces”. *Prog. Surf. Sci.* 89 (2014), pp. 80–125. DOI: `10.1016/j.progsurf.2014.01.001`.

[16] J. Burgdörfer and C. Lemell. “Theoretical concepts and methods for electron emission from solid surfaces”. In: *Slow Heavy-Particle Induced Electron Emission from Solid Surfaces.* Vol. 225. Springer Tracts in Modern Physics. Berlin, Heidelberg: Springer, Berlin, Heidelberg, 2007, pp. 1–38. ISBN: 978-3-540-70788-2. DOI: `10.1007/3-540-70789-1_1`.

[17] F. Aumayr and H. Winter. “Potential Electron Emission from Metal and Insulator Surfaces”. In: *Slow Heavy-Particle Induced Electron Emission from Solid Surfaces.* Vol. 225. Springer Tracts in Modern Physics. Berlin, Heidelberg: Springer, Berlin, Heidelberg, 2007, pp. 79–112. ISBN: 978-3-540-70788-2. DOI: `10.1007/3-540-70789-1_3`.

[18] A Cobas and W. E. Lamb. “On the extraction of electrons from a metal surface by ions and metastable atoms”. *Phys. Rev.* 65 (1944), pp. 327–337. DOI: `10.1103/PhysRev.65.327`.

[19] R. A. Baragiola and C. A. Dukes. “Plasmon-Assisted Electron Emission from Al and Mg Surfaces by Slow Ions”. *Phys. Rev. Lett.* 76 (1996), pp. 2547–2550. DOI: `10.1103/PhysRevLett.76.2547`.

[20] P. Riccardi et al. “Bulk and surface plasmon excitation in the interaction of He^+ with Mg surfaces”. In: *Nucl. Instruments Methods Phys. Res. Sect. B Beam Interact. with Mater. Atoms.* Vol. 212. 1-4. North-Holland, 2003, pp. 339–345. DOI: `10.1016/S0168-583X(03)01424-1`.

[21] D. Goebl et al. “Band structure effects in Auger neutralization of He ions at metal surfaces”. *Phys. Rev. B* 84 (2011), p. 165428. DOI: `10.1103/PhysRevB.84.165428`.

[22] D Goebl et al. “Quasi-resonant neutralization of He^+ ions at a germanium surface”. *J. Phys. Condens. Matter* 25 (2013), p. 485006. DOI: `10.1088/0953-8984/25/48/485006`.

[23] D. Valdés et al. “Role of d Electrons in Auger Neutralization at Metal Surfaces”. *Phys. Rev. Lett.* 97 (2006), p. 047601. DOI: `10.1103/PhysRevLett.97.047601`.

[24] R. A. Baragiola. “Principles and mechanisms of ion induced electron emission”. *Nucl. Inst. Methods Phys. Res. B* 78 (1993), pp. 223–238. DOI: `10.1016/0168-583X(93)95803-D`.

[25] M. O. Aboelfotoh and J. A. Lorenzen. “Influence of secondary-electron emission from MgO surfaces on voltage-breakdown curves in Penning mixtures for insulated-electrode discharges”. *J. Appl. Phys.* 48 (1977), pp. 4754–4759. DOI: `10.1063/1.323490`.

[26] S. J. Yoon et al. “Intrinsic and effective secondary electron emission coefficients in AC plasma display panel”. *Jpn. J. Appl. Phys* 40 (2001), pp. 809–812. DOI: `10.1143/jjap.40.809`.

[27] N. Lorente, R. Monreal, and M. Maravall. “Electron emission in the neutralization of multiply-charged ions at low velocities on metal surfaces: the effect of secondary-electron cascades”. *Nucl. Inst. Methods Phys. Res. B* 100 (1995), pp. 290–295. DOI: `10.1016/0168-583X(94)00833-7`.

[28] S. De Waele. “Private Communication”. 2018.

[29] H. B. Michaelson. “The work function of the elements and its periodicity”. *J. Appl. Phys.* 48 (1977), pp. 4729–4733. DOI: `10.1063/1.323539`.

[30] H. Kawano. “Effective work functions for ionic and electronic emissions from mono- and polycrystalline surfaces”. *Prog. Surf. Sci.* 83 (2008), pp. 1–165. DOI: `10.1016/j.progsurf.2007.11.001`.

[31] S. De Waele et al. “Error estimates for density-functional theory predictions of surface energy and work function”. *Phys. Rev. B* 94 (2016), p. 235418. DOI: `10.1103/PhysRevB.94.235418`.

[32] R. A. Baragiola, C. A. Dukes, and P. Riccardi. “Plasmon excitation in ion-solid interactions”. In: *Nucl. Instruments Methods Phys. Res. Sect. B Beam Interact. with Mater. Atoms.* Vol. 182. 1-4. North-Holland, 2001, pp. 73–83. DOI: `10.1016/S0168-583X(01)00723-6`.

[33] R. A. Baragiola and R. C. Monreal. “Electron emission from surfaces mediated by ion-induced plasmon excitation”. In: *Slow Heavy-Particle Induced Electron Emission from Solid Surfaces.* Vol. 225. Springer Tracts in Modern Physics. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 185–211. DOI: `10.1007/3-540-70789-1_6`.

[34] P Apell. “Surface plasmon de-excitation of multiply charged ions”. *J. Phys. B At. Mol. Opt. Phys.* 21 (1988), pp. 2665–2673. DOI: `10.1088/0953-4075/21/14/019`.

[35] M. Vicente Alvarez, V. Ponce, and E. Goldberg. “Auger, resonant, and plasmon-assisted charge-transfer processes in atom-surface collisions”. *Phys. Rev. B* 57 (1998), pp. 14919–14929. DOI: `10.1103/PhysRevB.57.14919`.

[36] R. H. Ritchie. “Plasma Losses by Fast Electrons in Thin Films”. *Phys. Rev.* 106 (1957), pp. 874–881. DOI: `10.1103/PhysRev.106.874`.

[37] R. F. Egerton. “Electron energy-loss spectroscopy in the TEM”. *Reports Prog. Phys.* 72 (2009), p. 016502. DOI: `10.1088/0034-4885/72/1/016502`.

[38] R. Gallager. *Poisson Processes.* MIT open courseware. 2011.

[39] H. Raether. *Excitation of Plasmons and Interband Transitions by Electrons.* Springer-Verlag, 1980. DOI: `10.1007/bfb0045951`.

[40] S. A. Maier. *Plasmonics: Fundamentals and Applications.* Springer US, 2007. DOI: `10.1007/0-387-37825-1`.

[41] Y. Takeishi and H. D. Hagstrum. “Auger-type electron ejection from the (111) face of nickel by slow He+, Ne+, and Ar+ ions”. *Phys. Rev.* 137 (1965). DOI: `10.1103/PhysRev.137.A641`.

[42] H. D. Hagstrum. “Ion-neutralization spectroscopy of solids and solid surfaces”. *Phys. Rev.* 150 (1966), pp. 495–515. DOI: `10.1103/PhysRev.150.495`.

[43] A. K. Green and E. Bauer. “Work function and purity of the Beryllium (0001) surface”. *Surf. Sci.* 74 (1978), pp. 676–681. DOI: `10.1016/0039-6028(78)90023-7`.

[44] R. Baragiola et al. “Ion-induced electron emission from clean metals”. *Surf. Sci.* 90 (1979), pp. 240–255. DOI: `10.1016/0039-6028(79)90341-8`.

[45] H. Oechsner. “Electron yields from clean polycrystalline metal surfaces by noble-gas-ion bombardment at energies around 1 keV”. *Phys. Rev. B* 17 (1978), pp. 1052–1056. DOI: `10.1103/physrevb.17.1052`.

[46] U. A. Arifov. *Interaction of Atomic Particles with a Solid Surface.* Springer US, 1969. DOI: `10.1007/978-1-4899-4809-0`.

Appendix A

Computational Details

All calculations in this thesis were performed in the Density Functional Theory [1, 2] (DFT) framework, as implemented in the Vienna Ab initio Simulation Package [3–5] (VASP), save for the calculation of the energy landscapes presented in Section 5.4.2. This appendix starts with a brief description of the VASP software package, as well as its input and output files (Sec. A.1.1) and most important input parameters (Sec. A.1.2). Next, the chapter details the settings used to obtain all of the results presented in this thesis, organized per chapter and section in the order they are presented. Finally, there are still two brief sections, one on parallelization tests (Sec. A.3) and one on an issue with the way VASP calculates the real part of the dielectric tensor using the Kramers-Kronig relation.

A.1 Vienna Ab initio Simulation Package

In order to solve the many-body problem using DFT, we need a software package that is able to implement the theory numerically on a computer cluster. Currently, there is a wide selection of such packages available to computational scientists, each with their respective advantages and disadvantages.

VASP is particularly suited for materials science, relying on an unbiased plane wave basis set and offering several projector augmented wave [6, 7] (PAW) data sets for most atomic species of the periodic table. It can calculate an approximate solution to the many-body Schrödinger equation within the DFT formalism or the HF approximation, including the possibility of mixing to utilize hybrid functionals. VASP also employs a set of efficient iterative procedures to find the ground state of a system, and allows parallelization of the calculations on multi-core machines.

This section presents a concise overview of the different files used by VASP, discussing their purpose, and takes a closer look at the input parameters and their relation to the theory.

A.1.1 Files

VASP uses four basic input files for its calculations, which must always be in the directory where it is executed:

- `INCAR`: Contains the input parameters for the calculation. Various settings can be adjusted according to the needs of the user through a large number of *tags*, which are described in Section A.1.2. This can be considered the most important input file, in the sense that it has the most diverse content and therefore has a lot of control over the calculation. Because of this, it is also more prone to be the cause of errors.

- `POSCAR`: Contains the lattice vectors of the unit cell, as well as the atomic positions of the structure. The user is free to specify the atomic positions in Cartesian or direct coordinates. In case selective dynamics is used to fix certain atom coordinates in the unit cell, this is also indicated in this file. In order to make sure there are no mistakes in the POSCAR file, it is important to first visualize the structure.

- `KPOINTS`: Defines the $\mathbf{k}$ point mesh (Sec. 2.2.5), either by explicitly entering all the points or using an automatically generated Monkhorst-Pack [8] grid. For band structure calculations, there is the useful *line mode*. Here the user can specify certain (symmetry) lines along which to calculate the band structure of the crystal. These lines are described pairwise via the coordinates of their end points.

- `POTCAR`: Concatenation of the projector augmented wave [6, 7] (PAW) data sets for the different elements present in the crystal structure. VASP supplies a set of POTCAR files for each element and supported functional, corresponding to different choices for the number of valence electrons and radii of the PAW sphere. It is important to make sure the order of the atoms is the same for the POTCAR and POSCAR file.

Besides these four essential files, a few other files can serve as input files for the VASP calculation:

- `STOPCAR`: This file can be used to stop the calculation without killing the VASP process. By writing either `LABORT = True` or `LSTOP = True` to the `STOPCAR` file, the user can stop the calculation after the next electronic or ionic step, respectively.

- `CHGCAR`: The electronic charge density of the unit cell is written to this file. Although the file is an output file, it can be used as an input file to start a calculation with a desired charge density.

- `WAVECAR`: The plane wave coefficients of the wave functions are written to this file. Although the file is an output file, it can be used as an input file to start a calculation with a desired set of wave functions. Note that this is only possible if neither the number of bands nor the set of plane waves has changed.

During the calculation, VASP produces a set of output files from which the user can extract the data necessary for his or her research. Here I present a (non-exhaustive) list of the most important VASP output files, besides the `CHGCAR` and `WAVECAR` already presented previously.

- `OUTCAR`: The general output file of VASP, a lot of information is printed in the this file during the calculation. The verbosity of the output is determined by the `NWRITE` tag in the `INCAR` file.

- `OSZICAR`: Presents an overview of the total energy at each SCF iteration, as well as some other properties interesting for monitoring the convergence of the calculation, for each ionic step in the case a geometry optimization is performed.

- `CONTCAR`: The final lattice vectors and atom positions of the unit cells are written to this file. Obviously this is mostly important when performing a geometric optimization.

- `vasprun.xml`: XML formatted output file, written at the end of the calculation. A lot of output from other files is gathered into this one file, which makes it the most useful for post processing.

- `DOSCAR`: Contains the density of states (DOS) of the system, as well as the integrated DOS and the projected DOS, in case `LORBIT` is set correctly.

- `IBZKPT`: The set of irreducible $\mathbf{k}$-points, along with their respective weights, can be found in this file.

- `EIGENVAL`: Details the Kohn-Sham eigenvalues for all **k**-points, which can for example be used to plot the band structure.

Note that some output is present in several files, but unfortunately VASP is not always entirely sensible in which output is printed where. An example here is the fact that although the `vasprun.xml` file contains the dielectric tensor, it does not contain the plasma frequencies.

A.1.2 Input Parameters

This section presents some of the input tags which are set by the user to determine the specifics of the calculation. The list below is by no means exhaustive; we simply focus on a selection of tags that were especially relevant for the results presented in this thesis. For the complete list, we refer the reader to the VASP manual.

- `ENCUT`: Energy cutoff (in eV) used for determining the size of the plane wave basis set, as per Eq. (2.50).
- `PREC`: Determines several settings that influence the precision of the calculations. First, the default value of the energy cutoff is increased for higher precision settings. Second, `PREC` sets the density of the grid used for the Fourier transformation. Finally, the precision of the representation of the PAW projectors is set by PREC.
- `ALGO`: Sets the algorithm used for the diagonalization of the Hamiltonian matrix when solving the Kohn-Sham equations.
- `EDIFF`: Convergence criterion on the self-consistency cycle during the electronic optimization, i.e. when determining the electron charge density.
- `ISMEAR`: Specifies the *smearing* method. Smearing is a technique that is designed to improve the convergence of calculation with respect to the sampling of the first Brillouin zone, specifically for metals. An excellent discussion by Prof. Marzari on the topic can be found here. The basic idea is to replace the step function in the calculation of the total energy:

  $$\sum_n \frac{1}{\Omega_{BZ}} \int_{BZ} \epsilon_{n\mathbf{k}} \Theta(\epsilon_{n\mathbf{k}} - E_F) d\mathbf{k}, \tag{A.1}$$

  by a smooth function $f(\{\epsilon_{n\mathbf{k}}\})$. The main advantage of using smearing methods is that the integral in Eq. (A.1) can be calculated accurately using a relatively sparse **k**-mesh (Sec. 2.2.5). Another method used to solve the integral in Eq. (A.1) is the *tetrahedron* method [9] (ISMEAR = -5), which linearly interpolates $\epsilon_{n\mathbf{k}}$ between each p **k**-points. Which method is preferable depends on the calculation being performed, and is specified by the input set loaded by the WriteVaspFromIOSet task.
- `SIGMA`: Smearing width used for the smearing of the occupies, as per the method specified with `ISMEAR`. Note that specifying `SIGMA` for the tetrahedron method makes little sense, as this method does not apply any smearing.
- `IBRION`: Determines the algorithm for the geometry optimization. A common and stable choice here is the conjugate gradient algorithm [10] (`IBRION` = 2).
- `ISIF`: Specifies the degrees of freedom for the geometry optimization, e.g. whether to optimize only the atomic positions (`ISIF = 2`), or perform a full optimization of the structure (`ISIF = 3`).

- `EDIFFG`: Sets the value for the convergence condition of the geometry optimization. The condition can be either applied to the total energy by setting a positive value, or the forces, when a negative value is provided.
- `LOPTICS`: Boolean setting that indicates that the frequency dependent dielectric tensor should be calculated.
- `CSHIFT`: Complex shift used in the Kramers-Kronig relation used to calculate the real part of the dielectric tensor. We refer the reader to Appendix A.4 for more details.
- `NBANDS`: Number of bands to include in the calculation. VASP includes a limited number of empty bands by default, but in order to calculate the optical properties or density of states over a larger energy range, this number should be increased to at least 2-3 times the default value.
- `NEDOS`: Number of points in the energy mesh for the calculation of the density of states and dielectric tensor.
- `ISPIN`: Specifies the spin-polarization setting, i.e. 1 or 2. The default is to perform a non-spin-polarized calculation (`ISPIN` = 1). Using `ISPIN` = 2 starts a spin-polarized calculation, i.e. for a system with collinear spins.
- `MAGMOM`: Allows the user to set the magnetic moments of all the atoms in the unit cell, when performing a spin-polarized calculation with collinear spins (`ISPIN` = 1). For non-collinear calculations, the magnetization density is a vectorial quantity, and the components of the magnetization density should be provided for each atom, with respect to the spin quantization axis (see SAXIS).
- `LHFCALC`: Boolean tag that indicates that the Hartree-Fock/DFT hybrid calculations should be performed.
- `AEXX`: Determines the mixing parameter for the hybrid calculations, i.e. the fraction a of Hartree-Fock exact exchange energy that should be included for the short range interaction.
- `HFSCREEN`: Specifies the separation parameter ω for hybrid calculations, i.e. beyond what distance the interaction energy is considered to be long range instead of short range.

A.2 Results

This section provides a more conventional description of the computational settings used for the calculations. Most of these have been directly copied from the corresponding papers, and updated slightly where necessary. As these sections are supposed to serve as a reference which each header of sections that discuss results links to, there is quite a bit of repetition if the reader goes through this section in one go.

A.2.1 Solar Cells

Structure and formation energy

We make a selection of ten compounds for which we can compare the calculated efficiency of the CuAu-like (CA) phase with the chalcopyrite (CH) results of Yu and Zunger [11]. The CA and CH structure are studied using a first-principles approach within the Density Functional Theory [1, 2] (DFT) formalism, as implemented in the Vienna Ab initio Simulation Package [3–5] (VASP). The projector augmented wave [6, 7] (PAW) method is applied, and the electrons that are treated as valence electrons are underlined in Table A.1. The exchange-correlation functional is calculated using the generalized gradient approximation (GGA) of Perdew-Burke-Ernzerhof [12] (PBE). The energy cutoff for the plane wave basis is set to 350 eV, and a 4×4×4 Monkhorst-Pack [8] mesh is used for sampling the first Brillouin zone. The charge density is considered converged when the energy difference between two electronic steps is smaller than 10^{-4} eV. The geometry is considered optimized when the forces on the atoms are all below 10^{-2} eV/Å.

Table A.1: Electron configuration of the atoms.

Element	Configuration
Cu	$[\mathrm{Ar}]\ \underline{3d^{10}4s^{1}}$
Ag	$[\mathrm{Kr}]\ \underline{4d^{10}5s^{1}}$
Ga	$[\mathrm{Ar}]\ 3d^{10}\underline{4s^{2}4p^{1}}$
In	$[\mathrm{Kr}]\ 4d^{10}\underline{5s^{2}4p^{1}}$
S	$[\mathrm{Ne}]\ \underline{3s^{2}3p^{4}}$
Se	$[\mathrm{Ar}]\ 3d^{10}\underline{4s^{2}4p^{4}}$
Te	$[\mathrm{Kr}]\ 4d^{10}\underline{5s^{2}5p^{4}}$

Absorber layer efficiency

Because an accurate band gap is important for the correct evaluation of the efficiency, we perform single shot G_0W_0 [13] calculations on top of hybrid HSE06 [14]. However, in order to update the quasiparticle energies within the G_0W_0 approximation with sufficient precision, it is necessary to consider the semi-core electrons as valence electrons within the projector augmented wave [6, 7] (PAW) framework. Hence, we treat the $3s$, $3p$ and $3d$ ($4s$, $4p$ and $4d$) orbitals as valence states for the Ga (In) atoms for the G_0W_0@HSE06 calculations of the band gap, on top of those underlined in Table A.1. In addition, we use a well converged 8×8×8 Monkhorst-Pack [8], an energy cutoff of 400 eV and a large amount of unoccupied bands (600 in total).

The optical properties are calculated within the Random Phase Approximation (RPA), using the long wavelength expression for the imaginary part of the dielectric tensor [15, 16]. The real part of the dielectric tensor is determined using the Kramers-Kronig relation[1]. In order to get an accurate description of the energy levels, the exchange-correlation energy is calculated with the HSE06 [14] functional, which has been reported [17] to produce optical properties close to

---

[1] The Kramers-Kronig relation is calculated by VASP using a complex shift (“`CSHIFT`”). After calculating the real part, however, VASP also recalculates the corresponding imaginary part. Since the complex shift introduces a broadening, this causes an earlier onset of the imaginary part, and consequently in the absorption coefficient. In order to prevent this, we commented out the line in the VASP code that recalculates the imaginary part. Note that in a more recent version of VASP, this smearing of the imaginary part can also be avoided by choosing a smaller CSHIFT settings. See Appendix A.4.

those obtained from experiment for $\mathrm{CuIn(S}_x\mathrm{Se}_{x-1})_2$. We find that it is sufficient to sample the Brillouin zone using a 12×12×12 Monkhorst-Pack [8] mesh to obtain a reasonably converged dielectric tensor. The number of unoccupied bands is increased to at least three times the number of occupied bands. Because of the tetragonal symmetry of the CA structure, the resulting dielectric tensor is diagonal and has two independent components ε_{xx} $(=\varepsilon_{yy})$ and ε_{zz}. Since we make no assumptions about the direction from which the photons enter the absorber layer, we average the diagonal components to derive the dielectric function $\varepsilon(E) = \varepsilon^{(1)}(E) + i\varepsilon^{(2)}(E)$ at energy E. Finally, in order to obtain a more accurate onset of the absorption spectrum, we shift the imaginary part of the dielectric function to the $\mathrm{G_0W_0}$@HSE06 band gap, and recalculate the real part using the Kramers-Kronig relations.

A.2.2 Li-ion Batteries

All calculations are performed based on the Density Functional Theory [1, 2] (DFT) formalism, as implemented in the Vienna Ab initio Simulation Package [3–5] (VASP). The projector augmented wave [6, 7] (PAW) is used to make a distinction between the core and valence electrons, with the standard VASP recommended choice for the number of valence electrons. The wave functions of the valence electrons are expanded in a plane wave basis set, using a high energy cutoff equal to 500 eV, which is advisable for structures containing oxygen.

Structure and Li-configurations

The configurations are optimized with PBE+U [18, 19], where a range of choices for the U parameter were tested to closely match the magnetic moments and lattice constants of a HSE06 [14] calculation for bulk O3-$\mathrm{Li_2MnO_3}$ using the same settings as described below. We settled on a U correction of 3.9 eV, applied to the $3d$ orbitals of Mn. A Monkhorst-Pack [8] mesh with a reciprocal density of 100 $\mathrm{\mathring{A}}^{-3}$ is chosen for the k-point sampling of the Brillouin zone. Geometry optimizations were performed with a Gaussian smearing of 0.05 eV, followed by a static calculation using the tetrahedron method [9] for a more precise calculation of the total energy, where we have doubled the density of the k-point mesh. The energy convergence criterion on the electronic optimization is set at 10^{-5} eV, and 10^{-3} eV for the geometric optimization, i.e. the difference in energy between ionic steps. For the static calculation, the electronic energy convergence criterion is tightened slightly to 10^{-6} eV. All lowest energy bulk structures are further optimized with the HSE06 [14] functional to obtain the final geometries discussed in the text.

Oxidation

All calculations of the magnetic moments and density of states were performed with the hybrid HSE06 [14] functional, based on the geometries obtained from the calculations presented in the previous section. The density of the Monkhorst-Pack [8] mesh was doubled to 200 $\mathrm{\mathring{A}}^{-3}$, and the tetrahedron method [9] was used for the integration of the Brillouin zone. As Ir is known to exhibit a strong spin-orbit interaction, non-collinear calculations including spin-orbit coupling were performed for calculating the magnetic moment and density of states of $\mathrm{Li}_x\mathrm{IrO_3}$. The charge density is considered converged when the difference in energy between electronic steps is smaller than 10^{-5} eV.

Dimer

The dimer reaction energy and kinetics of both compounds were calculated in a 2×2×2 supercell, where we once again switched to the PBE+U [18, 19] functional in order to make the dimer screening computationally feasible. We applied a U correction of 3.9 eV to the 3*d* orbitals of Mn. For Ir, we used the same value as McCalla et al. [20] (4.0 eV). Activation energies were calculated using the nudged elastic band [21, 22] (NEB) method.

Thermodynamic Stability of Sn substitution

The formation energies of Sn-substituted structures for a range of x-values have been calculated within the Density Functional Theory [1, 2] (DFT) framework, as implemented in the Vienna Ab initio Simulation Package [3–5] (VASP). The projector augmented wave [6, 7] (PAW) method was used to make a distinction between the core and valence electrons, with the standard VASP recommended choice for the number of valence electrons. The exchange-correlation energy was calculated using the SCAN+rVV10 [23, 24] functional to include the van der Waals interaction, which is especially important for a layered structure such as SnO [25]. The wave functions of the valence electrons are expanded in a plane wave basis set, using a high energy cutoff equal to 500 eV, which is advisable for structures containing oxygen. For all 2×2×2 supercell calculations, a 3×3×3 Monkhorst-Pack [8] mesh was used for sampling the Brillouin zone, whereas a 6×6×3 and 9×9×7 mesh were used for Li_2SnO_3 and SnO, respectively. Geometry optimizations were performed with a Gaussian smearing of 0.05 eV, followed by a static calculation using the tetrahedron method [9], for a precise calculation of the total energies. The convergence criterion on the electronic optimization is set at 10^{-4} eV, and 10^{-3} eV for the geometric optimization.

Influence of Mn^{4+} substitution on oxygen stability

For the calculation of the density of states and kinetic barrier of the dimer formation, both the PBE+U [18, 19] and SCAN [23] functionals have been used. Similar to our previous calculations, we have applied a U correction of 3.9 eV to the 3*d* orbitals of Mn. For V and Mo, we have applied a U correction of 3.1 eV [26] and 4.38 eV [27], respectively. A 3×3×3 Monkhorst-Pack [8] mesh was used for sampling the Brillouin zone for all geometry optimizations, including those of the nudged elastic band [21, 22] (NEB) calculations with the climbing image modification [28] used to calculate the kinetic barrier. A more dense 7×7×9 mesh was used for the calculation of the density of states. For the calculation of the kinetic barrier, the FIRE [29] force-based optimizer was used, and the convergence condition on the forces was set at 10^{-2} eV Å^{-1}.

Energy landscape of $[CB_{11}H_{12}]^-$

All calculations for the landscapes were performed within the Density Functional Theory [1, 2] (DFT) framework, as implemented in the NWChem software package [30]. All atoms adopted correlation consistent local basis sets at the double zeta level with diffuse augmentation (aug-cc-pVDZ) [31–33], as provided by the EMSL basis set exchange [34, 35]. The exchange-correlation energy was calculated using the generalized gradient approximation (GGA) of Perdew-Burke-Ernzerhof [12] (PBE). Each point in the energy landscapes $LiCB_{11}H_{12}$ and $NaCB_{11}H_{12}$ molecules were computed using a static calculation of the energy of the isolated cation-anion system.

A.2.3 Ion-Induced Secondary Electron Emission

Semiconductors

Hagstrum's model requires the density of states of the valence $D_v(\epsilon)$ and conduction $D_c(\epsilon)$ band as input, as well as the vacuum level. We calculate the density of states and vacuum level of

Ge(111) and Si(111) using a Density Functional Theory [1, 2] (DFT) approach, as implemented in the Vienna Ab initio Simulation Package [3–5] (VASP). Within the projector augmented wave [6, 7] (PAW) formalism, the recommended number of valence electrons is included for both Ge and Si. The energy cutoff is set at 500 eV in order to obtain a well converged plane wave basis set, and the exchange correlation energy is calculated using the generalized gradient approximation (GGA) approximation of Perdew-Burke-Ernzerhof [12] (PBE). A well converged 18×36×1 Monkhorst-Pack [8] k-point mesh is used for sampling the Brillouin zone.

To simulate a surface within the periodic boundary framework of VASP, it is conventional to take a slab approach, where a certain number of atomic layers are separated by a suitably large vacuum layer (See 3.3.4). For Si and Ge, it is well known that the (111) surfaces reconstruct, forming dimers at the surface with a 2×1 periodicity. We take the reconstructed structures from the supplementary material of De Waele et al. [36] and subsequently optimize the geometry using the computational parameters described in the previous paragraph. The slab consists of 14 atomic layers and at least 20 Å of vacuum spacing is present. The vacuum level is obtained by averaging the one-electron electrostatic potential over planes parallel to the surface and determining the potential in the vacuum, which should be constant in case the vacuum layer is sufficiently thick. The work function ϕ of the surface is then calculated by comparing the vacuum level with the top of the valence band $\phi = \epsilon_0 - \epsilon_v$.

Metals

Hagstrum's model requires the density of states of the occupied and unoccupied ($D_v(\epsilon)$ and $D_c(\epsilon)$) states as input, as well as the vacuum level. We calculate the density of states and vacuum level of all metal surfaces using a Density Functional Theory [1, 2] (DFT) approach, as implemented in the Vienna Ab initio Simulation Package [3–5] (VASP). Within the projector augmented wave [6, 7] (PAW) formalism, the recommended number of valence electrons is included for all metals. The energy cutoff is set at 500 eV in order to obtain a well converged plane wave basis set, and the exchange correlation energy is calculated using the generalized gradient approximation (GGA) of Perdew-Burke-Ernzerhof [12] (PBE). For sampling the Brillouin zone, Monkhorst-Pack [8] k-point mesh is used for which the spacing in each direction is smaller than 0.05 Å^{-1}.

To simulate a surface within the periodic boundary framework of VASP, it is conventional to take a slab approach, where a certain number of atomic layers are separated by a suitably large vacuum layer (See 3.3.4). We take the structures of all surfaces from the supplementary material of De Waele et al. [36] and subsequently optimize the geometry using the computational parameters described in the previous paragraph. The slab consists of 14 atomic layers and at least 20 Å of vacuum spacing is present. The vacuum level is obtained by averaging the one-electron electrostatic potential over planes parallel to the surface and determining the potential in the vacuum, which should be constant in case the vacuum layer is sufficiently thick. The work function ϕ of the surface is then calculated by comparing the vacuum level with the top of the Fermi level $\phi = \epsilon_0 - \epsilon_v$.

The optical properties of the bulk are calculated within the Random Phase Approximation (RPA), using the long wavelength expression for the imaginary part of the dielectric tensor [15, 16]. The real part of the dielectric tensor is determined using the Kramers-Kronig relations. For the damping parameter in the Drude expression of the intraband part of the dielectric tensor a value of 50 meV is used.

A.3 Parallelization

In order to speed up the calculations of the workflows, it is important to set reasonably good parallelization parameters for VASP (`NPAR`/`NCORE`, `KPAR`). Knowing beforehand what settings are optimal is tricky, as this depends on the system being studied, the computational settings and the machine the calculations are being run on. To come up with the somewhat rudimentary algorithm behind `VaspParallelizationTask`, I have performed a whole series of tests on a selected set of systems. An illustration of the (partial) results of such a test can be seen in Figure A.1. The square with a thick red edge indicates the setting which was found by `VaspParallelizationTask`. The algorithm doesn't always find the optimal setting, but does a fairly good job of finding one that does not waste significant amount of resources compared to the optimal choice. Clearly, it does a much better job compared to the default (`KPAR` = `NPAR` = 1). All other parallelization tests can be explored interactively via the Jupyter notebook corresponding to this section.

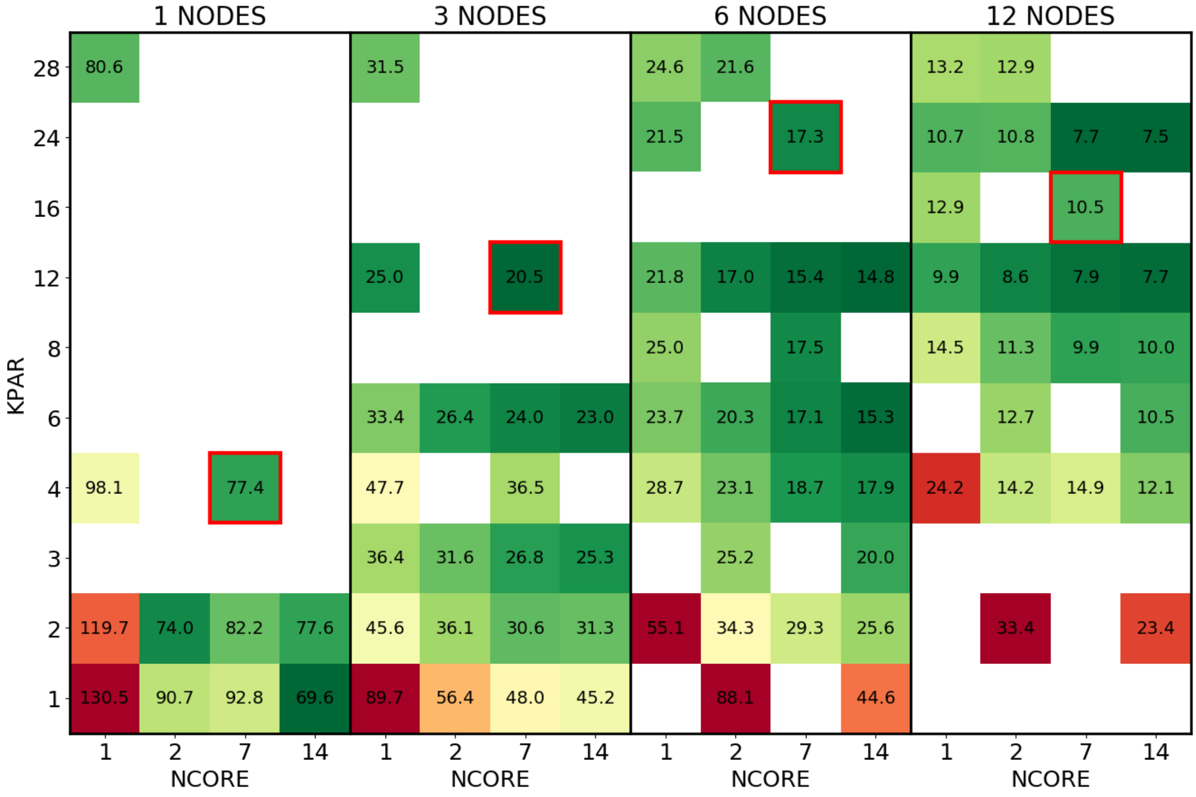


Figure A.1: Example of a parallelization test for U_3O_8. The number in each square represents the average electronic time step.

A.4 CSHIFT

In order to calculate the real part of the dielectric tensor, VASP uses the Kramers-Kronig transformation (Eq. (2.86)):

$$\varepsilon^{(1)}_{\alpha\beta}(\omega) = 1 + \frac{2}{\pi}\mathcal{P}\int_0^\infty \frac{\varepsilon^{(2)}_{\alpha\beta}(\omega')\omega'}{(\omega')^2 - \omega^2} d\omega'. \tag{A.2}$$

However, this complex integral has a pole at the origin. This is solved by relying on a small complex shift, i.e. the whole integration path is shifted slightly into the complex plane:

$$\varepsilon^{(1)}_{\alpha\beta}(\omega) = 1 + \frac{2}{\pi}\int_0^\infty \frac{\varepsilon^{(2)}_{\alpha\beta}(\omega')\omega'}{(\omega')^2 - \omega^2 + i\eta} d\omega', \tag{A.3}$$

where η is the complex shift, which can be set in the `INCAR` file using the `CSHIFT` tag. Relying on a complex shift introduces a slight broadening to the real part of the dielectric tensor. So far, this does not affect the optical properties such as the absorption coefficient significantly. However, VASP also calculates the corresponding broadened imaginary part. This leads to an earlier onset of the imaginary part of the dielectric tensor, and subsequently an earlier onset of the absorption coefficient and absorptivity. Effectively, the band gap of the material is fictitiously reduced, which has a large effect on the efficiency calculated in Chapter 4.

One way of mitigate this issue is by removing all absorption below the band gap, i.e. setting the absorptivity $\alpha(E)$ to zero for $E < E_g$. This is definitely an improvement, but still can have quite a significant influence on the efficiency. Fortunately, VASP overwrites the broadened imaginary part of the dielectric tensor by the original calculated using Eq. (2.85) under certain conditions (which are unfortunately not usually obtained using the default settings). More specifically, in case the complex shift (`CSHIFT`) is smaller than the distance between two points in the energy mesh, VASP will output the original - unbroadened - imaginary part of the dielectric tensor.

To understand the influence of these various effects, I have set up a notebook studying the onset of the absorptivity at 500 nm for low and high `CSHIFT` value, based on the dielectric function of CA-CuInS$_2$. Figure A.2 summarizes the results for the absorptivity. The broadening of the absorptivity for the (default) high `CSHIFT` value is clear. Due to the severely reduced effective band gap, the calculated efficiency for the broadened absorptivity is 10 times smaller than the unbroadened value. We can also that that when the absorptivity is cut below the band gap, there is still a large difference in the onset when compared to the unbroadened (low `CSHIFT`) case. The difference in efficiency is about 2 %.

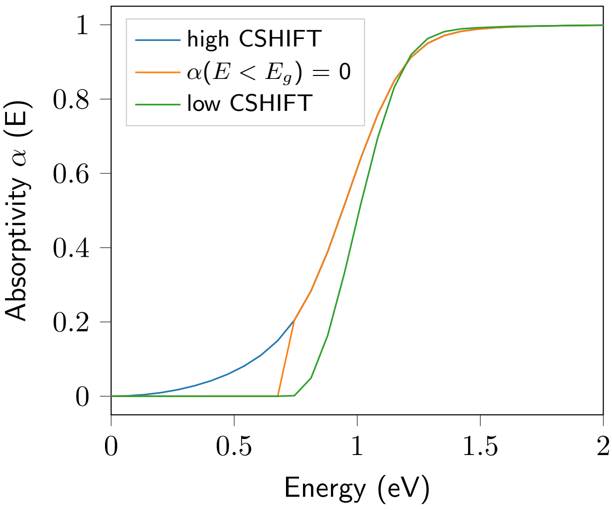


Figure A.2: Difference in onset of the absorptivity for the settings described in the text.

Appendix B

Plasmonic Excitations

B.1 Poisson Processes

A Poisson process is defined as (taken from [37]):

> *A Poisson process is a renewal process in which the inter arrival intervals* T_p *follow an exponential distribution function; i.e., for some real* $\lambda_p > 0$, T_p *has the probability density function*

$$f_{T_p}(t) = \begin{cases} \lambda_p e^{-\lambda_p t} & \text{for } t \geq 0 \\ 0 & \text{for } t < 0 \end{cases} \tag{B.1}$$

Let's start with some basic expressions that facilitate the derivations of the plasmon probabilities. First, note that the probability density function is properly normalized to 1:

$$\begin{aligned} \int_{-\infty}^{\infty} \lambda_p e^{-\lambda_p t} dt &= \int_{0}^{\infty} \lambda_p e^{-\lambda_p t} dt \\ &= \lambda_p \left[\frac{1}{-\lambda_p} e^{-\lambda_p t} \right]_0^{\infty} \\ &= -(0-1) = 1. \end{aligned}$$

The probability of the interval of process p being smaller than some specified interval T is:

$$\begin{aligned} \Pr\{T_p \leq T\} &= \int_0^T \lambda_p e^{-\lambda_p t} dt \\ &= \lambda_p \left[\frac{1}{-\lambda_p} e^{-\lambda_p t} \right]_0^{T} \\ &= -(e^{-\lambda_p T} - 1) = 1 - e^{-\lambda_p T}, \end{aligned} \tag{B.2}$$

which is simply the cumulative distribution function of the Poisson process. Similarly, the probability of the interval being larger than some specified interval T is:

$$\begin{aligned} \Pr\{T_p > T\} &= 1 - \Pr\{T_p \leq T\} \\ &= 1 - (1 - e^{-\lambda_p T}) = e^{-\lambda_p T} \end{aligned} \tag{B.3}$$

For two competing Poisson processes with rates λ_1 and λ_2, the probability of the interval T_1 being smaller than T_2 is:

$$\begin{aligned}
\Pr\{T_1 < T_2\} &= \int_0^\infty \Pr\{T_1 < T_2 | T_1 = t\} \lambda_1 e^{-\lambda_1 t} dt \\
&= \int_0^\infty \Pr\{T_2 > t\} \lambda_1 e^{-\lambda_1 t} dt \\
&= \int_0^\infty e^{-\lambda_2 t} \lambda_1 e^{-\lambda_1 t} dt \\
&= \lambda_1 \int_0^\infty e^{-(\lambda_1+\lambda_2)t} dt \\
&= \frac{\lambda_1}{-(\lambda_1+\lambda_2)}(0-1) \\
&= \frac{\lambda_1}{\lambda_1+\lambda_2} \qquad \text{(B.4)}
\end{aligned}$$

Equation (6.28) is now fairly easy to derive. Consider two Poisson processes, one for the surface plasmon excitation (*sp*) and one for the Auger neutralization (*aug*):

$$\begin{aligned}
f_{T_{sp}}(E_{sp}, t) &= \lambda_{sp}(E_{sp}) e^{-\lambda_{sp}(E_{sp})t} \\
f_{T_{aug}}(t) &= \lambda_{aug} e^{-\lambda_{aug} t}.
\end{aligned}$$

The probability of a surface plasmon excitation is equal to that of the interval T_{sp} being smaller than T_{aug}:

$$\begin{aligned}
P_{sp}(E_{sp}) &= \Pr\{T_{sp} < T_{aug}\} \\
&\overset{(B.4)}{=} \frac{\lambda_{sp}(E_{sp})}{\lambda_{sp}(E_{sp}) + \lambda_{aug}} \qquad \text{(B.5)} \\
& \qquad \text{(B.6)}
\end{aligned}$$

For the volume plasmons, we have to consider multiple competing Poisson processes. First, note that two independent Poisson processes with rates λ_1 and λ_2 can be merged into a new Poisson process with rate $\lambda' = \lambda_1 + \lambda_2$. So, if we consider a third process with rate λ_3, the probability of the interval of this process being smaller than that of process 1 and 2 is:

$$\begin{aligned}
\Pr\{T_3 < T'\} &= \frac{\lambda_3}{\lambda' + \lambda_3} \\
&= \frac{\lambda_3}{\lambda_1 + \lambda_2 + \lambda_3}. \qquad \text{(B.7)}
\end{aligned}$$

By extension, the probability of a Poisson process with rate λ_n having the shortest interval among a set of independent processes $\{\lambda_i\}$, with $i = 1, ..., N$, becomes:

$$\Pr\{T_n < \min(T_1, T_2, ..., T_{n-1}, T_{n+1}, ..., T_N)\} = \frac{\lambda_n}{\lambda_1 + ... + \lambda_N}. \qquad \text{(B.8)}$$

Within our model, an electron with energy ϵ can excite plasmons up to energy $\epsilon - \epsilon_F$, and the rate of exciting a volume plasmon with energy E_{vp} is proportional to the energy loss function

$L(E)$:

$$\lambda_{vp}(E_{vp}) = \begin{cases} c_{vp} \cdot L(E_{vp}) & \text{for } 0 < E_{vp} \leq \epsilon - \epsilon_F \\ 0 & \text{for } E_{vp} \leq 0 \text{ or } E_{vp} > \epsilon - \epsilon_F. \end{cases} \tag{B.9}$$

In this case, we have a continuous set of independent Poisson processes, each with rate equal to $\lambda_{vp}(E_{vp})$, so the probability distribution of first exciting a volume plasmon with energy E_{vp} becomes:

$$P_{vp}(\epsilon, E_{vp}) = \Pr\{T_{vp}(E_{vp}) = \min(T_{vp}(E)\} = \frac{\lambda_{vp}(E_{vp})}{\int_0^{\epsilon-\epsilon_F} \lambda_{vp}(E)dE}. \tag{B.10}$$

Note that this distribution is properly normalized to 1:

$$\begin{aligned} \int_0^\infty P_{vp}(\epsilon, E_{vp})dE_{vp} &= \int_0^\infty P_{vp}(\epsilon, E_{vp})dE_{vp} \\ &= \frac{\int_0^\infty \lambda_{vp}(E_{vp})dE_{vp}}{\int_0^{\epsilon-\epsilon_F} \lambda_{vp}(E)dE} \\ &= \frac{c_{vp}\int_0^{\epsilon-\epsilon_F} L(E_{vp})dE_{vp}}{c_{vp}\int_0^{\epsilon-\epsilon_F} L(E)dE} = 1 \end{aligned}$$

However, we also have to consider the average travel interval of the electrons t_e. As the total rate of the volume plasmon excitation process is $\int_0^{\epsilon-\epsilon_F} \lambda_{vp}(E)dE$, the probability of any volume plasmon being excited in the interval t_e is:

$$\Pr\{\min(T_{vp}(E)) < t_e\} = 1 - e^{-(\int_0^{\epsilon-\epsilon_F} \lambda_{vp}(E)dE)t_e}. \tag{B.11}$$

Combining both leads to the probability distribution of a volume plasmon excitation for an electron with average travel interval t_e:

$$\begin{aligned} P_{vp}(\epsilon, E_{vp})dE_{vp} &= \frac{\lambda_{vp}(E_{vp})dE_{vp}}{\int_0^{\epsilon-\epsilon_F} \lambda_{vp}(E)dE}\left[1 - e^{-\left(\int_0^{\epsilon-\epsilon_F} \lambda_{vp}(E_{vp})dE\right)t_e}\right] \\ &= \frac{c_{vp}L(E_{vp})dE_{vp}}{\int_0^{\epsilon-\epsilon_F} c_{vp}L(E)dE}\left[1 - e^{-\left(\int_0^{\epsilon-\epsilon_F} c_{vp}L(E)dE\right)t_e}\right] \\ &= \frac{L(E_{vp})dE_{vp}}{\int_0^{\epsilon-\epsilon_F} L(E)dE}\left[1 - e^{-\left(\int_0^{\epsilon-\epsilon_F} L(E)dE\right)c_{vp}t_e}\right] \end{aligned} \tag{B.12}$$

B.2 Volume Plasmons

Consider the distribution of excited electrons $N_i(\epsilon)$. The probability that an excited electron at energy ϵ induces a volume plasmon of energy $E_{vp} < \epsilon - \epsilon_F$ is given by Eq. (B.12). The fraction of electrons which excite a volume plasmon is then equal to:

$$N_{vp}^-(\epsilon) = \int_0^\infty N_i(\epsilon)P_{vp}(\epsilon, E_{vp})dE_{vp}. \tag{B.13}$$

These are subtracted from the excited electron density. However, these electrons are not lost, they have simply lost an energy equal to the volume plasmon energy E_{vp}. Hence, we have to add the following density to the excited density again:

$$N_{vp}^{+}(\epsilon) = \int_0^\infty N_i(\epsilon + E_{vp}) P_{vp}(\epsilon, E_{vp}) dE_{vp}. \tag{B.14}$$

The story doesn't end there, however. Volume plasmons do not couple with transversal waves [38], and hence cannot decay radiatively. Most likely, the plasmon will decay as a single electron excitation. In order to model this, we first calculate the distribution of plasmons with an energy E_{vp}:

$$D_{vp}(E_{vp}) = \int_0^\infty N_i(\epsilon) P_{vp}(\epsilon, E_{vp}) d\epsilon \tag{B.15}$$

When these plasmons decay, they release an energy equal to E_{vp}. In other words, this distribution should be convoluted with the density of the occupied states $D_v(\epsilon)$ in order to obtain the density of excited electrons due to volume plasmon decay:

$$\begin{aligned}
N_{vp}^{d}(\epsilon) &= \frac{1}{n_{vp}} (D_v * D_{vp})(\epsilon) && \text{(B.16)}\\
&= \frac{1}{n_{vp}} \int_{\Delta\epsilon_v} D_v(\epsilon_1) D_{vp}(\epsilon - \epsilon_1) d\epsilon_1 && \text{(B.17)}\\
&= \frac{1}{n_{vp}} \int_{\Delta\epsilon_v} D_v(\epsilon_1) \int_0^\infty N_i(\epsilon_2) P_{vp}(\epsilon_2, \epsilon - \epsilon_1) d\epsilon_2 d\epsilon_1 && \text{(B.18)}
\end{aligned}$$

where n_{vp} is a normalization factor. Of course, this distribution should be normalized to the number of excited plasmons, i.e.:

$$n_{vp} = \int_0^\infty N_{vp}^{-}(\epsilon) d\epsilon \tag{B.19}$$

B.3 Surface Plasmons

Hagstrum's Auger transform can be written as a convolution:

$$\begin{aligned}
T\left[\frac{\epsilon + \epsilon_0 - E_I}{2}\right] &= \int_{\Delta\epsilon_v} \int_{\Delta\epsilon_v} D_v(\epsilon_1) D_v(\epsilon_2) \delta(\epsilon - \epsilon_1 - \epsilon_2 + \epsilon_0 - E_I) d\epsilon_1 d\epsilon_2. \\
&= \int_{\Delta\epsilon_v} D_v(\epsilon_1) D_v(\epsilon + \epsilon_0 - E_I - \epsilon_1) d\epsilon_1 \\
&= (D_v * D_v)(\epsilon + \epsilon_0 - E_I)
\end{aligned} \tag{B.20}$$

which can also be written as:

$$T\left[\frac{\epsilon + \epsilon_0 - E_I}{2}\right] = D_v(\epsilon) * D_v(\epsilon + \epsilon_0 - E_I) \tag{B.21}$$

Here, $D_E(\epsilon) = D_v(\epsilon + \epsilon_0 - E_I)$ corresponds to the distribution of energies released by the neutralization of the incoming ion[1]. In order to include surface plasmon excitations, we consider the probability that a plasmon with energy E_{sp} is excited instead of an Auger process:

$$P_{sp}(E_{sp}) = \frac{\frac{c_{sp}}{\lambda_{aug}}\mathrm{Im}\left[-\frac{1}{\varepsilon(E_{sp})+1}\right]}{\frac{c_{sp}}{\lambda_{aug}}\mathrm{Im}\left[-\frac{1}{\varepsilon(E_{sp})+1}\right]+1} \tag{B.22}$$

As the surface plasmon excitation process is considered to be resonant, E_{sp} must correspond to the energy released in the neutralization of the incoming ion, i.e. ϵ in $D_E(\epsilon)$. The leftover distribution of energies passed to other valence electrons through an Auger process is:

$$D_{aug}(\epsilon) = D_E(\epsilon)(1 - P_{sp}(\epsilon)) \tag{B.23}$$

We substitute this distribution in the expression for the Auger transform:

$$\begin{aligned} T'\left[\frac{\epsilon + \epsilon_0 - E_I}{2}\right] &= D_v(\epsilon) * D_{aug}(\epsilon) \\ &= D_v(\epsilon) * (D_v(\epsilon + \epsilon_0 - E_I)(1 - P_{sp}(\epsilon))) \end{aligned} \tag{B.24}$$

And use this Auger transform instead of the original one in the nominator of Eq. (6.4). Note that the normalization in the denominator of Eq. (6.4) should still use the original Auger transform, else the number of excited electrons due to the neutralization are still normalized to one per incoming ion.

---

[1] This should be more clear when we substitute e.g. $\epsilon = 0$ in $D_E(\epsilon)$. We then obtain $D_v(\epsilon_0 - E_I)$, i.e. the number of electrons in the valence band that are at an equal depth (in relation to the vacuum level) as the lowest unoccupied state of the incoming ion. Clearly, electrons at this level do not pass energy through any Auger process, as this is a resonant neutralization. Another way of understanding this is that an electron at energy level $\epsilon_0 - E_I + \epsilon$ releases an energy of ϵ when it neutralizes the incoming ion.

Bibliography

[1] P. Hohenberg and W. Kohn. "Inhomogeneous electron gas". *Phys. Rev.* 136 (1964), B864–B871. DOI: `10.1103/PhysRev.136.B864`.

[2] W. Kohn and L. J. Sham. "Self-consistent equations including exchange and correlation effects". *Phys. Rev.* 140 (1965), A1133–A1138. DOI: `10.1103/PhysRev.140.A1133`.

[3] G. Kresse and J. Hafner. "Ab initio molecular dynamics for liquid metals". *Phys. Rev. B* 47 (1993), pp. 558–561. DOI: `10.1103/PhysRevB.47.558`.

[4] G. Kresse and J. Furthmüller. "Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set". *Comput. Mater. Sci.* 6 (1996), pp. 15–50. DOI: `10.1016/0927-0256(96)00008-0`.

[5] G. Kresse and J. Furthmüller. "Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set". *Phys. Rev. B - Condens. Matter Mater. Phys.* 54 (1996), pp. 11169–11186. DOI: `10.1103/PhysRevB.54.11169`.

[6] P. E. Blöchl. "Projector augmented-wave method". *Phys. Rev. B* 50 (1994), pp. 17953–17979. DOI: `10.1103/PhysRevB.50.17953`.

[7] D. Joubert. "From ultrasoft pseudopotentials to the projector augmented-wave method". *Phys. Rev. B - Condens. Matter Mater. Phys.* 59 (1999), pp. 1758–1775. DOI: `10.1103/PhysRevB.59.1758`.

[8] H. J. Monkhorst and J. D. Pack. "Special points for Brillouin-zone integrations". *Phys. Rev. B* 13 (1976), pp. 5188–5192. DOI: `10.1103/PhysRevB.13.5188`.

[9] P. E. Blöchl, O. Jepsen, and O. K. Andersen. "Improved tetrahedron method for Brillouin-zone integrations". *Phys. Rev. B* 49 (1994), pp. 16223–16233. DOI: `10.1103/PhysRevB.49.16223`.

[10] W. H. Press et al. *Numerical Recipes.* Cambridge University Pr., Sept. 6, 2007. 1248 pp. ISBN: 0521880688.

[11] L. Yu and A. Zunger. "Identification of potential photovoltaic absorbers based on first-principles spectroscopic screening of materials". *Phys. Rev. Lett.* 108 (2012), p. 068701. DOI: `10.1103/PhysRevLett.108.068701`.

[12] J. P. Perdew, K. Burke, and M. Ernzerhof. "Generalized gradient approximation made simple". *Phys. Rev. Lett.* 77 (1996), pp. 3865–3868. DOI: `10.1103/PhysRevLett.77.3865`.

[13] L. Hedin. "New Method for Calculating the One-Particle Green's Function with Application to the Electron-Gas Problem". *Physical Review* 139 (1965), A796–A823. DOI: `10.1103/physrev.139.a796`.

[14] J. Heyd, G. E. Scuseria, and M. Ernzerhof. "Hybrid functionals based on a screened Coulomb potential". *J. Chem. Phys.* 118 (2003), pp. 8207–8215. DOI: `10.1063/1.1564060`.

[15] M. Gajdoš et al. "Linear optical properties in the projector-augmented wave methodology". *Phys. Rev. B - Condens. Matter Mater. Phys.* 73 (2006), p. 045112. DOI: `10.1103/PhysRevB.73.045112`.

[16] J Harl et al. "Ab initio reflectance difference spectra of the bare and adsorbate covered Cu(110) surfaces". *Phys. Rev. B* 76 (2007), p. 35436. DOI: `10.1103/physrevb.76.035436`.

[17] F. C. Wan et al. "First-principles investigation of the optical properties of {CuIn}({SxSe}1{\textendash}x)2". *Mater. Sci. Semicond. Process.* 16 (2013), pp. 1422–1427. DOI: `10.1016/j.mssp.2013.05.009`.

[18] S. L. Dudarev et al. "Electron-energy-loss spectra and the structural stability of nickel oxide: An LSDA+U study". *Phys. Rev. B* 57 (1998), pp. 1505–1509. DOI: `10.1103/PhysRevB.57.1505`.

[19] O. Bengone et al. "Implementation of the projector augmented-wave LDA+U method: Application to the electronic structure of NiO". *Phys. Rev. B* 62 (2000), pp. 16392–16401. DOI: `10.1103/PhysRevB.62.16392`.

[20] E. McCalla et al. "Visualization of O-O peroxo-like dimers in high-capacity layered oxides for Li-ion batteries". *Science* 350 (2015), pp. 1516–1521. DOI: `10.1126/science.aac8260`.

[21] H. Jónsson, G. Mills, and K. W. Jacobsen. "Nudged elastic band method for finding minimum energy paths of transitions". In: *Classical and Quantum Dynamics in Condensed Phase Simulations*. World Scientific, 1998, pp. 385–404. ISBN: 978-981-02-3498-0. DOI: `10.1142/9789812839664_0016`.

[22] G. Henkelman and H. Jónsson. "Improved tangent estimate in the nudged elastic band method for finding minimum energy paths and saddle points". *J. Chem. Phys.* 113 (2000), pp. 9978–9985. DOI: `10.1063/1.1323224`.

[23] J. Sun, A. Ruzsinszky, and J. Perdew. "Strongly Constrained and Appropriately Normed Semilocal Density Functional". *Phys. Rev. Lett.* 115 (2015), p. 036402. DOI: `10.1103/PhysRevLett.115.036402`.

[24] H. Peng et al. "Versatile van der Waals Density Functional Based on a Meta-Generalized Gradient Approximation". *Physical Review X* 6 (2016). DOI: `10.1103/physrevx.6.041005`.

[25] K. Govaerts et al. "van der Waals bonding and the quasiparticle band structure of SnO from first principles". *Phys. Rev. B* 87 (2013). DOI: `10.1103/physrevb.87.235210`.

[26] W. D. Richards et al. "Fluorination of Lithium-Excess Transition Metal Oxide Cathode Materials". *Adv. Energy Mater.* 8 (2017), p. 1701533. DOI: `10.1002/aenm.201701533`.

[27] M. Kocher et al. *The Materials Project*. URL: `https://materialsproject.org/materials/mp-35975/` (visited on 01/13/2020).

[28] G. Henkelman, B. P. Uberuaga, and H. Jónsson. "A climbing image nudged elastic band method for finding saddle points and minimum energy paths". *J. Chem. Phys.* 113 (2000), p. 9901. DOI: `10.1063/1.1329672`.

[29] E. Bitzek et al. "Structural Relaxation Made Simple". *Phys. Rev. Lett.* 97 (2006). DOI: `10.1103/physrevlett.97.170201`.

[30] M. Valiev et al. "NWChem: A comprehensive and scalable open-source solution for large scale molecular simulations". *Comput. Phys. Commun.* 181 (2010), pp. 1477–1489. DOI: `10.1016/j.cpc.2010.04.018`.

[31] T. H. Dunning. "Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen". *J. Chem. Phys.* 90 (1989), pp. 1007–1023. DOI: `10.1063/1.456153`.

[32] R. A. Kendall, T. H. Dunning, and R. J. Harrison. "Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions". *J. Chem. Phys.* 96 (1992), pp. 6796–6806. DOI: `10.1063/1.462569`.

[33] B. P. Prascher et al. "Gaussian basis sets for use in correlated molecular calculations. VII. Valence, core-valence, and scalar relativistic basis sets for Li, Be, Na, and Mg". *Theor. Chem. Acc.* 128 (2010), pp. 69–82. DOI: `10.1007/s00214-010-0764-0`.

[34] D. Feller. "The role of databases in support of computational chemistry calculations". *J. Comput. Chem.* 17 (Oct. 1996), pp. 1571–1586. DOI: `10.1002/(sici)1096-987x(199610)17:13⟨1571::aid-jcc9⟩3.0.co;2-p`.

[35] K. L. Schuchardt et al. “Basis Set Exchange: A Community Database for Computational Sciences”. *J. Chem. Inf. Model.* 47 (2007), pp. 1045–1052. DOI: `10.1021/ci600510j`.

[36] S. De Waele et al. “Error estimates for density-functional theory predictions of surface energy and work function”. *Phys. Rev. B* 94 (2016), p. 235418. DOI: `10.1103/PhysRevB.94.235418`.

[37] R. Gallager. *Poisson Processes.* MIT open courseware. 2011.

[38] S. A. Maier. *Plasmonics: Fundamentals and Applications.* Springer US, 2007. DOI: `10.1007/0-387-37825-1`.

Publication list

A1 - Journals

- **M. Bercx**, N. Sarmadian, R. Saniz, B. Partoens and D. Lamoen, “First-principles analysis of the spectroscopic limited maximum efficiency of photovoltaic absorber layers for CuAu-like chalcogenides and silicon”, *Physical Chemistry Chemical Physics* **18**, pp. 20542-20549 (2016).
- T. Willhammar, K. Sentosun, S. Mourdikoudis, B. Goris, M. Kurttepeli, **M. Bercx**, D. Lamoen, B. Partoens, I. Pastoriza-Santos, J. Pérez-Juste, L. M. Liz-Marzán, S. Bals, Gustaaf van Tendeloo, “Structure and Vacancy Distribution in Copper Telluride Nanoparticles Determine Plasmon Activity in the Near Infrared”, *Nature Communications* **8**, 14925 (2017).
- M. Dimitrievska, P. Shea, K. E. Kweon, **M. Bercx**, J. B. Varley, W. S. Tang, A. V. Skripov, V. Stavila, T. J. Udovic, and B. C. Wood, “Carbon Incorporation and Anion Dynamics as Synergistic Drivers for Ultrafast Diffusion in Superionic $LiCB_{11}H_{12}$ and $NaCB_{11}H_{12}$”, *Advanced Energy Materials* **8** (15), 1703422 (2018).
- **M. Bercx**, B. Partoens, D. Lamoen, “Quantitative modeling of secondary electron emission from slow-ion bombardment on semiconductors”, *Physical Review B* **99**, 085413 (2019).
- **M. Bercx**, L. Slap, B. Partoens and D. Lamoen, “First-Principles investigation of the oxygen framework of Li-rich battery cathodes”, *MRS Advances* **4**, pp. 813-820 (2019).
- C. Callaert, **M. Bercx**, D. Lamoen and J. Hadermann, “Interstitial defects in the Van der Waals gap of Bi2Se3”, *Acta Crystallographica Section B* **75**, pp. 717-732 (2019).
- K. Choudhary, **M. Bercx**, J. Jiang, R. Pachter, D. Lamoen and F. Tavazza, “Accelerated Discovery of Efficient Solar-cell Materials using Quantum and Machine-learning Methods”, *Chemistry of Materials* **31**, pp. 5900-5908 (2019).
- M. Jørgensen, P. Shea, A. Tomich, J. Varley, **M. Bercx**, S. Lovera, R. Černý, W. Zhou, T. J. Udovic, V. Lavallo, T. Jensen, B. Wood, V. Stavila, “Understanding Superionic Conductivity in Lithium and Sodium Salts of Weakly Coordinating Closo-Hexahalocarbaborate Anions”, *Chemistry of Materials* **32**, pp. 1475-1487 (2020).
- A. Paulus, M. Hendrickx, **M. Bercx**, O. Karakulina, M. Kirsanova, D. Lamoen, A. M. Abakumov, J. Hadermann, M. K. Van Bael, and A. Hardy, “An in-depth study of Sn substitution in Li-rich/Mn-rich NMC as a cathode material for Li-ion batteries”, *submitted to Journal of Materials Chemistry A.*

B2 - Book chapter

- **M. Bercx**, R. Saniz, B. Partoens, D. Lamoen, “Exceeding the Shockley-Queisser limit within the detailed balance framework”, In: Many-Body Approaches at Different Scales, ed. by G.G.N. Angilella and C. Amovilli; (New York : Springer, 2018), ISBN 978-3-319-72373-0 (2018).

In preparation

- **M. Bercx**, B. Partoens, D. Lamoen, “Modeling secondary electron emission from slow-ion bombardment on metals by including plasmonic effects”, *in preparation.*

Conference Contributions

- **Talk**: “First-principles analysis of the efficiency of photovoltaic layers for CuAu-like chalcogenides and silicon”, at the *International Conference on Advanced Materials Modelling*, September 5-7 2016 in Rennes, France.
- **Poster**: “First-Principles Investigation of Local Cation-Anion Interactions in Polyborane Solid Electrolytes”, at the *Computational Chemistry and Materials Science Summer Institute*, June 5 - August 11 2017 in Lawrence Livermore National Laboratory, USA.
- **Talk**: “Exceeding the Shockley-Queisser limit within the detailed balance framework”, at the *Deutsche Physikalische Gesellschaft Spring Meeting of the Condensed Matter Section*, March 11-16 2018 in Berlin, Germany.
- **Talk**: “First-principles modeling of secondary electron emission from slow ion bombardment”, at the *Deutsche Physikalische Gesellschaft Spring Meeting of the Condensed Matter Section*, March 11-16 2018 in Berlin, Germany.
- **Poster**: “Exceeding the Shockley-Queisser limit within the detailed balance framework”, at the *Belgian Physical Society meeting*, April 11 2018 in Antwerp, Belgium.
- **Talk**: “First-Principles Investigation of the Stability of the Oxygen Framework of Li-Rich Battery Cathodes”, at the *Materials Research Society fall meeting*, November 25-30 2018 in Boston, USA.
- **Talk**: “First-Principles Investigation of the Stability of the Oxygen Framework of Li-Rich Battery Cathodes”, at the *International Conference on Advanced Materials Modelling*, June 26-29 2019 in Rennes, France.
- **Poster**: “First-Principles Investigation of the Stability of the Oxygen Framework of Li-Rich Battery Cathodes”, at the *Big Data Summer School*, September 9-13 2019 in Platja d’Aro, Spain.