PhD-FSTM-2025-052
The Faculty of Science, Technology and Medicine

DISSERTATION

Defence held on 24 April 2025 in Luxembourg

to obtain the degree of

DOCTEUR DE L'UNIVERSITÉ DU LUXEMBOURG

EN PHYSIQUE

by

Henry Phillip FRIED
Born on 17 June 1992 in Berlin, (Germany)

Ab-Initio and Machine-Learning Methods to Study Defects in Complex Materials

Dissertation defence committee

Prof. Dr. Ludger WIRTZ, dissertation supervisor
Université du Luxembourg

Prof. Dr. Hannu-Pekka KOMSA
University of Oulu

Prof. Dr. Patrick RINKE
Technische Universität München

Prof. Dr. Susanne SIEBENTRITT, Chair
Université du Luxembourg

Prof. Dr. Florian LIBISCH, Vice Chair
Technische Universität Wien

Affidavit

I hereby confirm that the PhD thesis entitled “Ab-Initio and Machine-Learning Methods to Study Defects in Complex Materials" has been written independently and without any other sources than cited.

Writefull and ChatGPT were solely used to edit text passages written by me to correct for grammar and spelling mistakes and to improve the readability of the text. AI tools were not used to generate original content in any way.

Luxembourg, _____ _____

Name

Acknowledgments

During the past years, I have received constant support from many people around me. Before expressing my gratitude, I write to note that I am equally grateful to everyone who assisted me in any capacity during this experience, without assigning any order of importance.

I would like to thank my supervisor, Ludger Wirtz, not only for his guidance and support both within and beyond the University over the past years, but also for encouraging me to come to the Grand Duchy of Luxembourg, which has broadened my horizon.

Equally important for my choice to come here was Daniel Barragan-Yani. I am thankful for his introduction to the world of defects and for all the guidance and moral support over the past years.

As part of the theoretical solid-state group, my gratitude goes to all its current and former members, Muralidhar Nalabothula, Aseem Kshirsagar, Sven Reichardt, Max Sinner and Sahar Pakdel. Not only did the coffee breaks, but also the discussion about scientific and daily topics contributed to where I am now. I also thank Adamantia Galani for her outstanding support with everything that came along.

Next, I would like to thank the University of Luxembourg and its staff for their support and the resources provided during my time here. I would also like to thank the Fonds National de la Recherche (FNR) for funding my PhD (PRIDE17/12246511/PACE) and my fellow PhD students who were part of this doctoral training unit, Poorani Gnanasambandan, Ricardo Poeira, Longfei Song, Alfredo Blázquez Martínez, Saba Tabean, Taowen Wang and Himanshu Phirke, who shared all the difficulties and who were dedicated to make our

small conference to the positive experience that it was. My gratitude also goes to Phillip Dale for his efforts in managing our meetings and our appointments, and Alex Redinger for his support during teaching and scientific discussions. I am also thankful for my time in Vienna, where I would like to highlight Christoph Schattauer's commitment to making it as pleasant and interesting as possible.

I would also like to express my sincere thanks to the CET committee members, Florian Libisch and Susanne Siebentritt, for their support and critical assessment over the years. My gratitude also goes to the other committee members, Hannu-Pekka Komsa and Patrick Rinke, for taking the time to read and evaluate my thesis, and for visiting us for my defense. I am grateful for all your feedback.

I am also thankful for Nora Schwartz, Michel Leruth, Cathy Wagner, Kirsti Hagemann and Hannah Milbers, with whom I felt at home during my time in Luxembourg. I would also like to thank Daniel Richter and Kevin Britz who gave me a warm welcome in Luxembourg; Paul Bourgue, for the support and the swimming sessions; Gustavo Morales for the shared joy in music.

Special thanks go to my Berlin flatmates Mohamed Ali Abdulmalik, Marcus Hohlfeld and Wojtek Szczepanik who always offered me a warm and welcoming place to stay during my visits; to my study partners at TU Berlin, who are since good friends Ramtin Dastgheib Shirazi, Nikolas Rixen, Peter Meinhold and Evangelos Louris; and to the Atomkerne Stephan Kamrad, Louisa Birkner and Leonie Köster who I can call good friends since high school, including, of course, the other Kern Arthur Gogohia for always supporting me. I am also grateful for my visitors in Luxembourg and Vienna that have not been named yet, Vincent Blaschke, Anna Fröchtenicht, Pauline Lüth, Henriette Langner, Alex Wagner, Rocío Berdaguer, Agnieszka Dabrowska, Klaudia, Julian and Elia Mannarini and of course all other friends. Although I have not mentioned every name, I hope you all know how much you mean to me and how much I enjoyed visiting Berlin. I would also like to thank my parents Ingo Fried and Christiane Ebel, their partners Michiyo Fried and Michael Schweiger, my siblings Lenny Fried and Lily Fried, and Christoph Lenze for their constant support.

Finally, I would like to thank my partner Michèle Peters for her love, for believing in me, and for highlighting my time outside of the university over the past years.

Abstract

Defects are present in all materials and are essential for many applications. However, defects can also reduce device performances, and understanding their properties is therefore crucial to further push the boundaries of solar-cell efficiency. In this context, chalcopyrites (Cu(In,Ga)Se_2/Cu(In,Ga)S_2) have emerged as an alternative to silicon-based solar cells. They are intrinsically doped and inherit a tunable band gap depending on the ratio of indium and gallium. This broad range of band gaps makes them a well-suited candidate for tandem-solar cells.

We want to understand the defect properties in $CuInS_2$ and $CuGaS_2$ by conducting a detailed study of the most common intrinsic defects in these materials. We use the Heyd, Scuseria and Ernzerhof (HSE) hybrid functional and investigate the impact of the two HSE parameters, α and ω, on the band gap and compliance with the general Koopmans' theorem. With the optimized HSE parameters, we calculate the thermodynamic charge transition levels to access the electronic properties. We further investigate the optical transition levels, by considering that optical transitions are vertical within a configuration coordinate diagram. Our results emphasize that for comparison with photoluminescence measurements, this difference (corresponding to a Franck-Condon shift) should not be neglected. We further calculate optical transitions with a Δ-SCF approach in which we constrain an electron to an unoccupied state enabling a charge neutral excited state.

A comparison with the available photoluminescence measurements in the literature reveals multiple candidates for the experimentally observed deep defect transitions. For indium-rich Cu(In,Ga)S_2, we find candidates for both deep defect transitions, namely

$Cu_{In/Ga}$, V_{In}, V_{Cu}, Cu_i and In_{Cu}. Furthermore, our results suggest three defects, V_{Cu}, Cu_{Ga}, and Cu_i as possible candidates that could be involved in a broad defect peak around 2.15 eV observed in measurements on $CuGaS_2$.

The complexity of these materials and the need for supercells to compute defect properties limit us to total energy evaluations. Thus, calculations of macroscopic properties of defects remains difficult. One possible solution is to use semi-empirical methods like tight-binding that are capable of calculating the properties of large super cells with many more atoms than density functional theory (DFT) is capable of.

Fitting such a semi-empirical tight-binding model to a pristine crystal via the band structure is common practice. However, the fit to the defective band structure become difficult because the increase of the cell in real space results in a folding in reciprocal space. We therefore present a workflow that uses the atom- and orbital-projected density of states instead. Additionally, introducing defects into the crystal perturbs its structure. The fit of the large number of different parameters for these defective systems presents a considerable challenge. We reduce the number of fitting parameters to a minimum by introducing two distance dependencies for the tight-binding parameters. This allows a neural network to learn the impact of defects on the host material within the tight-binding approximation.

We investigate carbon substitutions (monomer and dimer) in hexagonal boron nitride and demonstrate the capability of the workflow to predict tight-binding parameters for deep defect states. To address transferability of the tight-binding model, we show that it can describe defects under strained conditions and that it maintains DFT accuracy for different distances between neighboring defects. The method opens a path to understanding complicated defect landscapes using a computationally affordable semi-empirical approach without sacrificing accuracy, thereby enabling investigations of macroscopic defect properties.

List of Publications

1. **H. P. Fried**, D. Barragan-Yani, F. Libisch, and L. Wirtz. *A machine learning approach to predict tight-binding parameters for point defects via the projected density of states*. *Accepted in* npj Computational Materials.

2. **H. P. Fried**, D. Barragan-Yani, and L. Wirtz. *Electronic and optical properties of native point defects in $CuInS_2$ and $CuGaS_2$*. *Under review in* Physical Review Materials.

Contents

Abbreviations

AI Artificial intelligence

CCD Configurational coordinate diagram

CBM Conduction band minimum

cDFT Constrained DFT

CTL Charge transition level

DFT Density Functional Theory

DOS Density of States

gKT generalized Koopmans' theorem

HSE Heyd, Scuseria, and Ernzerhof

KS Kohn-Sham

LR Long-range

ML Machine learning

OTL Optical transition level

PDOS Projected Density of States

PL Photoluminescence

SR Short-range

TB Tight-binding

VBM Valence band maximum

ZPL Zero phonon line

Chapter 1

Introduction

The year 2024 marks an important milestone, as it is the first year in which the average global temperature exceeded 1.5℃ above preindustrial levels [1]. This surpasses the target outlined in the Paris Agreement, although the temperature remains below the upper threshold of 2℃ [2]. This small margin underscores the urgency of addressing climate change.

The risks of climate change have been recognized for more than a century. For example, in 1896, Swedish professor and Nobel laureate for chemistry in 1903, Svante Arrhenius, explored how carbon dioxide emissions from burning fossil fuels could increase global temperatures [3]. Today, there is a strong scientific consensus that the current warming is predominantly human-induced and driven by greenhouse gas emissions [4]. Reducing these emissions is one of the most critical challenges of our time.

In the fight against climate change, renewable energy is a cornerstone along with storage solutions and improved grid capacity, helping to reduce greenhouse gas emissions and their impact on global temperatures. Renewable energy can be generated through various means and solar cells, wind turbines, and hydropower are already widely used. In 2023, the European Union generated 24.5% of its energy from renewable sources [5]. Despite room for improvement, the number has more than doubled in the last 20 years, reflecting significant progress in the transition to cleaner energy.

1.1 Defects and solar cells

Solar cells are among the most important technologies for renewable energy, as they directly convert sunlight into usable energy. The concept of a standard silicon solar cell is based on defects that introduce additional charges to the crystal. These defects can either introduce a positive (p-doping) or a negative charge carrier (n-doping) to the system. Single-junction solar cells consist of two oppositely doped materials that form a p-n junction. Defects are therefore necessary to build standard solar cells. However, the efficiency of solar cells can be strongly decreased by certain impurities and defects within their structure.

In the following, we will introduce some limitations on solar cell efficiency and the critical role that defects play. For a complete description of the processes in solar cells and limitations, the interested reader is referred to these books [6, 7].

Undesired defect processes

Although defects that act as dopants are necessary for the construction of most solar cells, they can also reduce the performance of devices. Impurities might introduce opposite doping, which can lead to doping compensation and therefore to reduced doping concentrations. Another problem is deep defects that introduce defect states (so-called trap states) within the band gap that act as recombination centers. For materials like silicon, which have an indirect band gap, this is one of the main recombination mechanisms. A key task to increase the efficiencies of solar cells is therefore the identification of undesired defects.

Spectral losses

Other limiting factors for solar cells are spectral losses. Firstly, only light that has an energy higher than that of the band gap can be absorbed and contribute to the efficiency. Secondly, light that has an energy higher than the band gap will excite electron-hole pairs

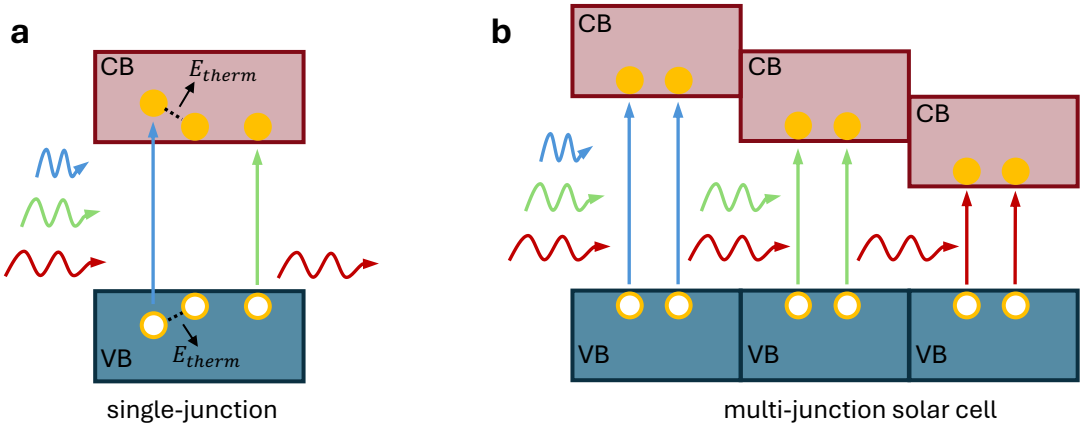

Figure 1.1: **a** Illustration of spectral losses in a single-junction solar cell. Photons with energy higher than the band gap (blue) result in thermal losses due to relaxation of the charge carrier to the band edges. Photons with energy below the band gap (red) are not able to excite an electron-hole pair and are not absorbed. **b** Illustration of a multi-junction solar cell. Photons with higher energy (blue) are absorbed in the first layer and do not contribute to thermal losses in the second layer. Photons with lower energy (green, red) are absorbed in the following layers.

below and above the respective band edges. The charge carriers relax down to the band edges, exciting phonons, and the excess energy is lost to thermal energy (Figure 1.1 **a**).

One approach to overcome spectral losses is to combine materials with different band gaps. These so-called multi-junction solar cells are capable of absorbing light at different wavelengths and are therefore addressing both causes of the discussed spectral losses (Figure 1.1 **b**). By stacking materials with a higher band gap on top of materials with a lower band gap, the photons that are not absorbed by the first layer can be absorbed by the next layers. Additionally, the photons with higher energy that would result in thermal losses have already been absorbed by a previous layer.

1.2 Cu(In,Ga)Se_2/Cu(In,Ga)S_2

Chalcopyrites like Cu(In,Ga)Se_2 and Cu(In,Ga)S_2 are native p-type semiconductors with a direct band gap. In contrast to the example of silicon, no extrinsic atoms are needed to dope them. In combination with other materials (n-type), commonly CdS, they are used as absorber materials in single-junction thin-film solar cells. Nevertheless, they are

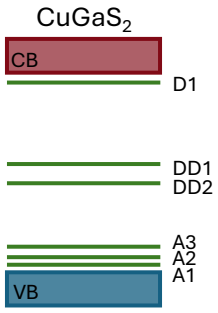

Figure 1.2: Illustration of defects observed in photoluminescence measurements in $CuGaS_2$ as reported by Adeleye et al. [10]. Four shallow defects, including three shallow acceptors and one shallow donor, as well as two deep defect-related transitions are observed.

also well suited for application in a multi-junction solar cell. The quaternary compounds $Cu(In,Ga)Se_2$ and $Cu(In,Ga)S_2$ (CIGS) have different band gaps that depend on the content of [In]/[Ga]. Theoretically, this allows for a band gap that continuously varies between the two limiting cases of having either only indium or gallium.

Selenide-based chalcopyrites have been studied extensively, both experimentally and theoretically [8], and a record efficiency of 23.6% has been achieved in 2024 [9]. The focus of recent research has been on sulfur-based CIGS as an alternative. As highlighted previously, a good understanding of the defect landscape, experimentally and theoretically, helps to improve the efficiencies in these materials.

In the first part of this thesis, we want to deepen our understanding of the defect landscapes in $CuInS_2$ and $CuGaS_2$. Both exhibit multiple defect-related transitions within photoluminescence (PL) measurements (see Figure 1.2 for CGS defects). Measurements in $CuGaS_2$ show two deep defect transitions that may act as recombination centers and affect the device performance [10–13]. One of the peaks shows a position dependence of the peak for different [Cu]/[Ga] ratios, and its origin remains unclear. A theoretical assessment of possible candidates for these peaks is still missing.

Complementary measurements on the quaternary compound $Cu(In,Ga)S_2$ with a high [In]/[Ga] ratio have been made by Shukla et al. [11], where they also observed two defect-related peaks. In addition, they have performed calculations of charge transition level for various defects in both compounds to identify possible candidates for one of the transitions.

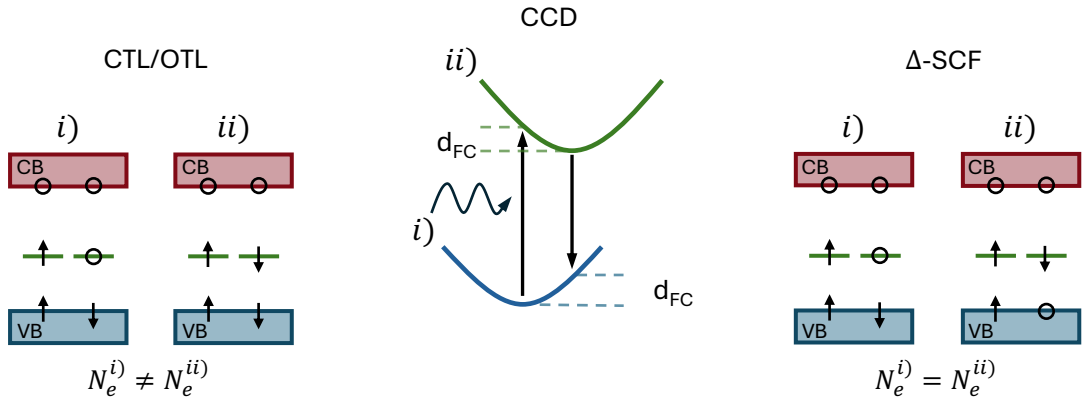

Figure 1.3: Illustration of methods for obtaining optical properties from total energy differences. Both optical transition level (OTL) and Δ-SCF approaches can be visualized with a configurational coordinate diagram (CCD). However, OTL compares the defect in two different charge states, whereas with Δ-SCF we can constrain an electron to a previously unoccupied state.

However, these transition levels are not always comparable to PL measurements. More precisely, we can include the impact of the lattice on the charging process by considering that the transitions are vertical within a configurational coordinate diagram (CCD) (Figure 1.3). This is similar to the Franck-Condon approximation for molecules, and, for the first time for chalcopyrite, we investigate these differences.

These optical transition levels still compare two systems with different charge states and, therefore, different number of electrons. However, in an optical process the number of electrons does not change. An alternative method to calculate the optical properties is the Δ-SCF approach in which we constrain an electron to a previously unoccupied Kohn-Sham state to model an excited state (Figure 1.3). We explore this methodology for defects in both sulfur-based chalcopyrites. Although this method is in principle more accurate, it is more difficult to converge these calculations given the complexity of the materials.

Due to this complexity, previous studies have been done on supercell sizes with 32 to 144 atoms. We increase the cell size to 288 atoms doubling the distance between defects in the c-direction with an effective increase of the distances of 1.48, to further boost our accuracy. Our results provided new detailed insights into the optical properties of defects in CIGS that have not yet been obtained. However, the computational demands of density

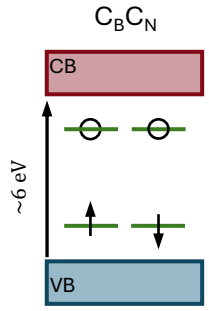

Figure 1.4: Illustration of the electronic structure of the carbon dimer in hBN.

functional theory pose significant challenges when dealing with larger systems.

To address these challenges, the second part of this thesis shifts toward method development, introducing a semi-empirical tight-binding approach designed to reduce computational costs while retaining accuracy in defect calculations including thousands of atoms. Given the inherent complexity of CIGS, we chose hexagonal boron nitride (hBN) as a simpler and well-characterized model system to validate this method.

1.3 hBN: Single-photon emission

The example of the solar cell demonstrated that defects can have a large impact on the electronic properties that we can use to build electronic devices. We want to highlight another application, namely single-photon emission, which can be used, e.g., in quantum sensing or quantum computing [14, 15].

Defect-related emission of single photons has been reported for different materials [16–18] and it is associated with deep defects that have localized electron states in the band gap (Figure 1.4). These small quantum systems can emit one single photon per excitation cycle. After excitation, an electron in the defect's excited state can decay back to the ground state through spontaneous emission, producing a single photon per excitation cycle.

One promising material in this context is monolayer hexagonal boron nitride (hBN), a material with a large band gap ($\approx$ 6 eV). The large band gap favors isolated mid-gap

states that can emit light in various wavelengths. Therefore, single-photon emission has been reported that ranges from visible to UV light [18–22]. Finding the defect involved in these transitions is an ongoing research. Some of them have been suggested to be intrinsic defects. However, there is also clear experimental and theoretical evidence that carbon atoms are involved at least in one of the emission at 4.1 eV [22–26] and are also suggested to be involved in other emission lines [25, 27, 28].

The relatively simple structure of hBN and its interesting defect properties make it an ideal candidate for developing and benchmarking our tight-binding method, paving the way for future extensions to more complex systems.

1.4 Tight-binding parameterization for defects

As a semi-empirical method, tight-binding enables calculations with many more atoms compared to *ab-initio* DFT; however finding a good parameterization for a defective crystal inherits multiple challenges closely related to the perturbation introduced by the defect. Ideally, we should address all to find the best parameterization, and considerable effort has been put into doing so [29, 30].

To accurately describe a defect, we must consider that the perturbation affects not only the local atomic environment but also the lattice structure, altering interatomic distances. In the case of a supercell (necessary to model defects), there are a lot of parameters that have to be fitted. Consequently, refitting every parameter and explicitly accounting for these changes quickly becomes impractical.

The supercell approach adds an additional layer of complexity for the fitting process itself. Generally, the tight-binding parameters are obtained by fitting the tight-binding band structure to *ab-initio* results. However, the supercell in real space results in the backfolding of the bands in reciprocal space (Figure 1.5). The addition of a defect further complicates this band structure, and unfolding to identify the correct eigenvalues becomes cumbersome.

In the second half of the thesis, we explore different approaches to solve the problems

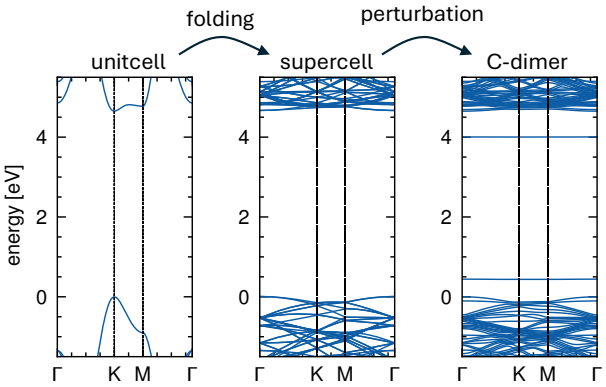

Figure 1.5: The necessity of supercells results in a folded band structure, which is difficult to disentangle. Adding a defect to the crystal complicates the problem. The presence of the defect perturbs the bands and can result in isolated electron states within the band gap.

related to defect tight-binding parameterizations. We will do so by describing carbon defects within hBN. It is straight-forward to introduce new tight-binding parameters that describe the carbon atoms. However, the impact on the host material is more delicate. A common way is to start with pristine parameters, which are perturbed by the presence of the defect. The impact on the onsite energies will be described by a Gaussian potential that varies the pristine onsite energies with a distance dependence to the defect. This is similar to the procedure of Lambin et al. [29], who were able to model nitrogen on carbon N_C subsitutional defect in graphene fairly well, by only adjusting the onsite energies. A completely opposite approach to describe defects was chosen by Schattauer et al. [30], who introduce a distance dependence between the hopping neighbors but neglect the impact on the onsite energies. They reduce the number of fitting parameters by discretizing the hopping matrix elements. We chose a different approach. Instead of fitting the parameters for a supercell, we make use of strain-dependent calculations of a pristine unit cell. This introduces a distance dependence between respective hopping neighbors.

We used the atom- and orbital-projected density of states instead of the folded band structure to fit the tight-binding model. Projecting the density of states onto a specific orbital disentangles the problem, while the projection on the site introduces a spatial resolution. However, the investigated carbon defects have a deep defect state within

a large band gap. Finding a good parameterization that gives rise to these states and simultaneously describes the bands correctly is not easy. We therefore employ a machine-learning algorithm that learns the mapping between tight-binding projected densities of states as input and tight-binding parameter as output. Since training is done within the tight-binding approximation, it is computationally inexpensive. Our approach opens a new path towards a comprehensive tight-binding description of more complex systems.

1.5 Structure of the thesis

The thesis is structured as follows. In Chapter 2, we will first introduce the methodology that helps us to understand defects and how we can model them theoretically. In Chapters 3 and 4, we discuss the main results and we will conclude the thesis in Chapter 5.

Chapter 2 starts by explaining theoretical methods of solid state physics that we use to describe defects, most importantly density functional theory and tight-binding. We will discuss defects in solids and how we can calculate their properties using the methods explained previously. Finally, we will introduce the concepts of machine learning, which enables to understand the architecture used later in the thesis.

In Chapter 3, we will apply the *ab-initio* methods to intrinsic point defects in $CuInS_2$ and $CuGaS_2$. We calculate the charge transition levels and discuss their connection to optical transition levels. Our results highlight that for deep defects we can not neglect the contribution of the lattice if we want to consider optical transitions. Additionally, we explore charge-neutral excitations with a Δ-SCF approach, in which we construct the excited state by forcing an electron to a previously unoccupied state. Our results suggest different defects that might be involved in optical transitions observed in photoluminescence measurements.

In Chapter 4, we introduce a novel path to obtain tight-binding parameters for defects. Instead of using the folded and intertwined band structure, we train a machine-learning model to learn the connection between projected density of states and tight-binding parameters. We simplify the problem by considering the defect as a perturbation to

the pristine crystal. This introduces different distance dependencies for both onsite and hopping parameters that reduce the number of fitting parameters to a minimum. This allows us to parameterize different carbon substitutional defects within hBN, including the famous carbon dimer, expected to be responsible for a single-photon emission peak at 4.1 eV. The tight-binding descriptions fitted to the projected density of states can reproduce the respective band structures. Furthermore, we find that the parameters can be transferred to defects with different distances between each other and that the tight-binding model can also describe the defects under strained conditions.

We will conclude the thesis in Chapter 5, where we summarize the thesis, before discussing the outlook.

Chapter 2

Methods

One of the goals of this thesis is to understand the properties of defects in crystals and how we can access them computationally. We will therefore first introduce some basic concepts of solid-state physics in Section 2.1. This gives us the tools to describe defects and the computational methods that we can use to calculate defect properties, which we will discuss in Section 2.2. Later in the thesis, we will use machine learning as a regression tool and we will introduce the basic concepts in Section 2.3.

2.1 Methods of solid state physics

Solid-state physics focuses on materials in their crystalline form. Because these materials consist of numerous periodically arranged atoms, providing a precise description is challenging. Nevertheless, various concepts and approximations enable us to study them. In this chapter, our goal is not to derive every concept in full detail but to provide a broad introduction to the topics necessary to understand and follow the thesis. Numerous well-written textbooks cover these topics in greater depth [31–34].

2.1.1 Many body Hamiltonian

The many-body Hamiltonian characterizing solids describes both electrons and nuclei. It includes the individual Hamiltonians for electrons ($\hat{\mathcal{H}}_e$) and nuclei ($\hat{\mathcal{H}}_n$), as well as an interaction Hamiltonian ($\hat{\mathcal{H}}_{n-e}$) that describes their interactions:

$$\hat{\mathcal{H}} = \underbrace{T_e + V_{ee}}_{\hat{\mathcal{H}}_e} \underbrace{+ T_n + V_{nn}}_{\hat{\mathcal{H}}_n} \underbrace{+ V_{ne}}_{\hat{\mathcal{H}}_{e-n}} \ . \tag{2.1}$$

Here, the first two Hamiltonians are defined by their respective kinetic (T) and potential (V) energies, while the electron-nuclei interaction is described solely by a Coulomb term. Together, they can be expressed as:

$$\hat{\mathcal{H}} = \underbrace{-\frac{\hbar^2}{2m_e}\sum_i \nabla_i^2 + \frac{1}{2}\sum_{i\neq j}\frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|}}_{\hat{\mathcal{H}}_e} \underbrace{-\sum_I \frac{\hbar^2}{2M_I}\nabla_I^2 + \frac{1}{2}\sum_{i,I}\frac{Z_I Z_J e^2}{|\mathbf{R}_I - \mathbf{R}_J|}}_{\hat{\mathcal{H}}_n} \underbrace{-\sum_{i,I}\frac{Z_I e^2}{|\mathbf{r}_i - \mathbf{R}_I|}}_{\hat{\mathcal{H}}_{n-e}} , \tag{2.2}$$

where electrons are denoted with lowercase subscripts and coordinates, while nuclei are indicated with uppercase subscripts and coordinates. This Hamiltonian represents a coupled differential equation and is not trivial to solve. Thus, we must introduce approximations to understand the complex behavior of solids with multiple nuclei and even more electrons.

The Born-Oppenheimer approximation provides a useful first simplification by separating the dynamics of electrons and nuclei [35]. In it, we can neglect the nuclear kinetic term, because it is proportional to $1/M_I$ and the nuclear mass M_I greatly exceeds the electron mass m_e. Consequently, we assume that the nuclei remain fixed at constant positions, so their Coulomb interactions remain constant and we can treat the electrons independently of the nuclei Hamiltonian.

Treating $\hat{\mathcal{H}}_n$ as constant simplifies the problem considerably; however the solution of

the Schrödinger equation remains a multi-electron wavefunction

$$\hat{\mathcal{H}}(r_1, r_2, ..., r_n)\psi(x_1, x_2, ..., x_N) = E\psi(x_1, x_2, ..., x_N) \ , \tag{2.3}$$

which depends on all electron positions and their interactions. Such wavefunctions are very large, making it unfeasible to solve these equations even with today's computational resources. Another simplification can be achieved by neglecting interactions between electrons. This allows the wavefunction to be factorized into a product of single-particle states:

$$\psi(x_1, x_2, ..., x_N) = \psi_1(x_1)\psi_2(x_2)...\psi_N(x_N) \ . \tag{2.4}$$

This approximation results in a solution in which all particles can occupy the same state, which is incorrect for electrons because they are fermions. Fermions must obey the Pauli exclusion principle, which states that two fermions can never occupy the same quantum state, implying that they must differ in at least one quantum number. This requirement can be enforced by using a Slater determinant:

$$\psi = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1(x_1) & \psi_2(x_1) & \cdots & \psi_N(x_1) \\ \psi_1(x_2) & \psi_2(x_1) & \cdots & \psi_N(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_1(x_N) & \psi_2(x_N) & \cdots & \psi_N(x_N) \end{vmatrix} \ , \tag{2.5}$$

which properly anti-symmetrizes the wave function.

Further approximations of the Hamiltonian are necessary to study solids. We will focus on two methods that are frequently used to investigate defects and are therefore particularly important. First, we will discuss density functional theory (DFT), a mean-field approach where the electron density is used instead of the wavefunction. As an *ab-initio*

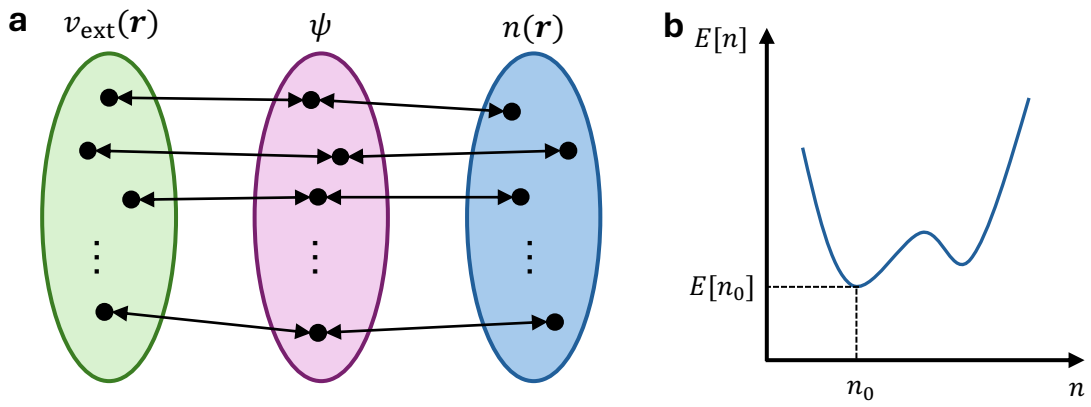

Figure 2.1: Illustration of the Hohenberg-Kohn theorem: **a** bijection between the wavefunction $\psi(\mathbf{r})$ and the external potential $v_{\text{ext}}(\mathbf{r})$ as well as the density $n(\mathbf{r})$. **b** The energy is expressed as a functional of the electron density instead of the wavefunction allowing for a minimization with respect to the density.

method, DFT is expected to yield very accurate results. Following this, we will introduce another important approximation, the tight-binding method, which employs a localized basis set and allows for the parameterization of the Hamiltonian. This parameterization enables calculations for systems containing many atoms.

2.1.2 Density functional theory

Density functional theory (DFT) is one of the most used electronic structure methods [36]. We begin by discussing the work of P. Hohenberg and W. Kohn [37] which laid the foundation for the Kohn-Sham equation [38].

Hohenberg-Kohn theorem

Similarly to the previous approximations, the goal of P. Hohenberg and W. Kohn was to simplify the many-body problem. Instead of using the complex, high-dimensional many-body wavefunction, they demonstrated that the low-dimensional density $n(\mathbf{r})$ can be used to describe the ground-state properties of the system. They achieved this through two proofs by reductio ad absurdum, which establish a bijective mapping among the external potential $v_{\text{ext}}(\mathbf{r})$, the wavefunction $\psi(\mathbf{r})$, and the density $n(\mathbf{r})$ (Figure 2.1 **a**) [37].

An immediate conclusion is that any observable can be expressed as a functional of the density or the external potential

$$\langle\psi|\,\hat{O}\,|\psi\rangle = O[\psi] = O[v_{\text{ext}}] = O[n]\ . \tag{2.6}$$

We can therefore evaluate observables of interest by using the density instead of the wavefunction. For example, we can minimize the total energy as a functional of the density, i.e., $\min_n E[n] = E_0$, This energy functional of the density can be expressed as

$$E[n] = \langle\psi|\,T_e + V_{ee} + V_{ext}\,|\psi\rangle \tag{2.7}$$

$$= \underbrace{T_e[n] + E_{ee}[n]}_{F_{\text{HK}}[n]} + \int d\mathbf{r}[V_{ext}(\mathbf{r})]n(\mathbf{r})\ , \tag{2.8}$$

where $F_{\text{HK}}[n]$ is the Hohenberg-Kohn or universal functional that is independent of the external potential. This can be rewritten to

$$F_{\text{HK}}[n] = E[n] - \int d\mathbf{r}[V_{ext}(\mathbf{r})]n(\mathbf{r})\ . \tag{2.9}$$

The two Hohenberg-Kohn theorems are illustrated in Figure 2.1 and read as:

1. The external potential v_{ext} (within a constant) is a unique functional of the density $n(\mathbf{r})$.

2. The functional $F_{\text{HK}}[n]$, if known, can be used to find the energy of the ground state by minimizing the three-dimensional density of the system.

These are two general statements that are not directly applicable because the functional $F_{\text{HK}}[n]$ is unknown. However, these two theorems laid the foundation for the work of W. Kohn and L. J. Sham [38].

Kohn-Sham DFT is based on the idea of an auxiliary system that replicates certain properties of the interacting many-body system. It assumes that a system of non-interacting electrons with the same density as the real system exists. By redefining the Hohenberg-Kohn functional (Equation 2.9), we can separate the energy into a calculable single-particle term, $E^H[n]$, and the exchange-correlation energy, $E^{\mathsf{XC}}[n]$, which accounts for all interaction contributions:

$$E[n] = \underbrace{T_s[n] + \int d\mathbf{r} V_{ext}(\mathbf{r}) n((\mathbf{r})) + E^H_{ee}[n]}_{E^H[n]} + \underbrace{T_e[n] - T_s[n] + E_{ee}[n] - E^H_{ee}[n]}_{E^{\mathsf{XC}}[n]} \ . \tag{2.10}$$

Here, $E^{\mathsf{XC}}[n]$ remains unknown. However, the differences in $E^{\mathsf{XC}}[n]$ are assumed to be small, since the single-particle energies and their many-body counterparts are expected to be similar. The minimization of the energy functional leads to a set of single-particle equations that describe the effective behavior of the electrons

$$\left[-\frac{\hbar^2}{2m_e}\nabla^2 + v_{ext}(\mathbf{r}) + \int d\mathbf{r}' \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + \frac{\delta E^{\mathsf{XC}}[n]}{\delta n(\mathbf{r})} \right] \psi_i(\mathbf{r}) = \lambda_i \psi_i(\mathbf{r}) \tag{2.11}$$

$$\left[-\frac{\hbar^2}{2m_e}\nabla^2 + \underbrace{v_{ext}(\mathbf{r}) + v^H[n] + v^{xc}[n]}_{v_{\mathsf{eff}}(\mathbf{r})} \right] \psi_i(\mathbf{r}) = \lambda_i \psi_i(\mathbf{r}) \ , \tag{2.12}$$

which is know as the Kohn-Sham equation, a Schrödinger like single-particle equation describing an auxiliary system with the same ground-state density as the many-body interacting system. The effective potential, $v_{\mathsf{eff}}(\mathbf{r})$, is often referred to as the Kohn-Sham potential, $v_{\mathsf{KS}}(\mathbf{r})$. In principle, this equation can be solved self-consistently to obtain the ground-state density of the system, and it should yield the exact solution provided that the exact exchange-correlation functional is known. Similar to the wavefunction of the real

system, we require that the Kohn-Sham orbitals are normalized

$$\int d\mathbf{r}\ \psi_i^*(\mathbf{r})\psi_i(\mathbf{r}) = 1\ . \tag{2.13}$$

The density of the auxiliary Kohn-Sham system is then defined in terms of the Kohn-Sham orbitals

$$n(\mathbf{r}) = \sum_i f_i |\psi_i(\mathbf{r})|^2\ , \tag{2.14}$$

where f_i is the occupation number.

We can use the Kohn-Sham equation to describe any observable. In the case of the total energy, we can rewrite the equation by multiplying the bra vector $\langle\psi_i(\mathbf{r})|$ for all occupied states, therefore expressing the Kohn-Sham eigenvalues in terms of the single-particle kinetic energy:

$$T_s[n] = \sum_i^{occ} \lambda_i - \int d\mathbf{r}\ n(\mathbf{r})V_{ext}(\mathbf{r}) - \iint d\mathbf{r}d\mathbf{r}' \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} - \sum_i^{occ} \langle\Psi_i(\mathbf{r})| \frac{\delta E_{xc}[n]}{\delta n(\mathbf{r})} |\Psi_i(\mathbf{r})\rangle\ . \tag{2.15}$$

We can substitute this into the total energy functional of the density in Equation 2.10

$$E[n] = \sum_i \lambda_i - \frac{1}{2}\int d\mathbf{r}d\mathbf{r}' \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} + E_{xc}[n] - \int d\mathbf{r}\ n(\mathbf{r})v_{xc}([n],\mathbf{r})\ . \tag{2.16}$$

One can immediately observe that the sum of the Kohn-Sham eigenvalues does not equal the total energy of the system. This is consistent with the definition of the auxiliary system, which is defined to inherit the same ground-state density. However, this does not imply that the eigenvalues themselves have a direct physical meaning. The differences arise from the double counting of the Hartree term and the contribution of the unknown exchange-correlation potential functional.

Semi-local exchange-correlation functionals

There are many approximations to describe the exchange-correlation functional, each varying in complexity and accuracy. W. Kohn and L. J. Sham provided the simplest

approximation along with the Kohn-Sham equation [38], the local density approximation (LDA). This functional is based on the Jellium model[1], whose solution is a homogeneous electron gas (HEG). For the HEG, analytical solutions exist for the exchange energy as a function of the density, $\epsilon_x^{\text{HEG}}(n) = -\frac{3}{4}(\frac{3}{\pi}n)^{\frac{1}{3}}$, thus providing a local description. The correlation energy, $\epsilon_c^{\text{HEG}}(n)$, can be obtained numerically. To obtain a global property, we integrate these local energy contributions over **r**

$$E_{xc}^{\text{LDA}}[n] = \int d\mathbf{r}\; n(\mathbf{r})\epsilon_{xc}^{\text{HEG}}(n(\mathbf{r}))\;. \tag{2.17}$$

The LDA can be further improved by including gradients and higher derivatives of the density and is known as the generalized gradient approximation (GGA):

$$E_{xc}^{\text{GGA}}[n\uparrow, n\downarrow] = \int d\mathbf{r}\; n(\mathbf{r})\epsilon_{xc}(n\uparrow, n\downarrow, \nabla n\uparrow, \nabla n\downarrow)\;. \tag{2.18}$$

One of the most widely used methods for constructing these functionals was proposed by J.P. Perdew, S. Burke, and M. Ernzerhof [39], and is known as the PBE functional. This functional will be used for all DFT calculations later in the thesis. However, both LDA and the improved GGA do not correct the nonphysical self-interaction error in the Hartree term.

Generalized Koomans' theorem

The eigenvalues of Kohn-Sham DFT represent single-particle states of an auxiliary system that shares the same ground-state density as the real system. Even if the exact exchange-correlation energy were known, the auxiliary system is not designed to yield electron addition and removal energies. There is, however, an exception for the highest occupied Kohn-Sham state, which defines the ionization energy (I) and electron affinity (A) of a

[1]The Jellium model approximates the solid by representing the nuclei as a smeared-out potential.

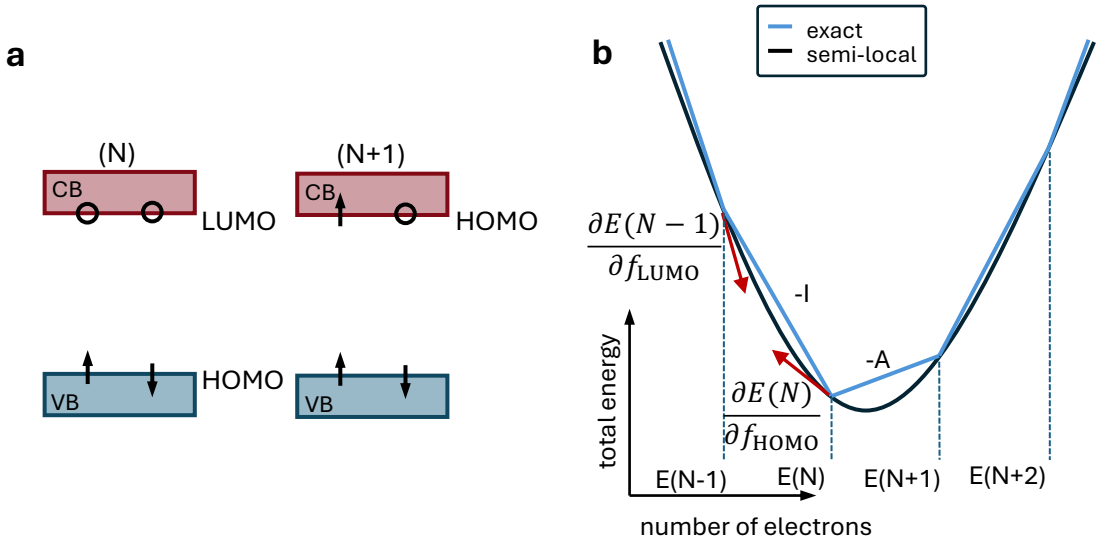

Figure 2.2: **a** HOMO and LUMO for systems with N and N+1 electrons. **b** Total energy dependence on the number of electrons. Exact DFT yields a piecewise linear function with derivative discontinuities at integer numbers, whereas semi-local DFT is a continuous function with a convex parabolic behavior. Inspired by Saboya [40].

system with N electrons [41, 42]. We can express both in terms of total energies

$$I = E(N-1) - E(N) \tag{2.19}$$

$$A = E(N) - E(N+1)\ , \tag{2.20}$$

where the parentheses indicate the number of electrons in the system.

Perdew et al. [41] showed that the exact exchange-correlation functional is a piecewise linear function with derivative discontinuities at integer electron numbers (see Figure 2.2). Janak's theorem [43] interprets the eigenvalues as the partial derivative of the total energy with respect to the fractional occupation of the orbitals:

$$\frac{\delta E}{\delta f_i} = \epsilon_i\ . \tag{2.21}$$

Thus, similar to Koopmans' theorem, we can interpret the fundamental band gap in terms of the ionization energy (I) and the electron affinity (A) using Kohn-Sham orbitals, and

compare it to the Kohn-Sham band gap [41, 42]:

$$E_g = I - A = -\epsilon^N_{\text{HOMO}} + \epsilon^{N+1}_{\text{HOMO}} \tag{2.22}$$

$$\epsilon^{\text{KS}}_g = \epsilon^N_{\text{LUMO}} - \epsilon^N_{\text{HOMO}} \tag{2.23}$$

$$\Rightarrow E_g = \epsilon^{\text{KS}}_g + \underbrace{\epsilon^{N+1}_{\text{HOMO}} - \epsilon^N_{\text{LUMO}}}_{\Delta_{xc}} \, . \tag{2.24}$$

In exact DFT, Δ_{xc} is referred to as the derivative discontinuity and must not be zero, due to a piecewise linear function that involves derivatives with varying slopes, when considering two different systems with different numbers of electrons (Figure 2.2).

However, if we consider the partial occupancy along one linear part (same ground state system), the partial derivatives are equal:

$$\frac{\delta E(N+1)}{\delta f}\Big|_{N+1-\eta} = \frac{\delta E(N)}{\delta f}\Big|_{N+\eta} \quad \Rightarrow \quad \epsilon^{N+1}_{\text{HOMO}} = \epsilon^N_{\text{LUMO}} \, , \tag{2.25}$$

Semi-local functionals are continuous and, therefore, lack the piecewise linear behavior and exhibit a convex curvature such that at any point:

$$\frac{\delta E(N+1)}{\delta f}\Big|_{N+1-\eta} > \frac{\delta E(N)}{\delta f}\Big|_{N+\eta} \quad \Rightarrow \quad \epsilon^{N+1}_{\text{HOMO}} > \epsilon^N_{\text{LUMO}} \, . \tag{2.26}$$

This means that convex behavior of the functional leads to a $\Delta_{xc} > 0$ and therefore the Kohn-Sham band gap underestimates the fundamental band gap.

In contrast to (semi-)local DFT, Hartree-Fock includes exact exchange energy and shows a concave behavior with respect to partial occupancy. A combination of both theories within hybrid functionals results in the possibility of correcting the band gap and getting closer to an exact functional.

Hybrid functionals

The basic idea of hybrid functionals is to combine the exact exchange energy from Hartree-Fock with a semi-local functional and partially reduce the self-interaction error and the

resulting problems. The general form is

$$E_{\mathrm{xc}}^{\mathrm{PBE0}} = \alpha E_{\mathrm{x}}^{\mathrm{HF}} + (1-\alpha) E_{\mathrm{x}}^{\mathrm{PBE}} + E_{\mathrm{c}}^{\mathrm{PBE}} \,, \tag{2.27}$$

where α is a mixing parameters. A widely used hybrid functional is the PBE0 with an $\alpha = 0.25$. Another important hybrid functional is introduced by Heyd, Scuseria and Ernzerhof (HSE) [44]:

$$E_{\mathrm{xc}}^{\mathrm{HSE}} = \alpha E_{\mathrm{x,SR}}^{\mathrm{HF}}(\omega) + (1-\alpha) E_{\mathrm{x,SR}}^{\mathrm{PBE}}(\omega) + E_{\mathrm{x,LR}}^{\mathrm{PBE}}(\omega) + E_{\mathrm{c}}^{\mathrm{PBE}} \,. \tag{2.28}$$

This additionally includes a separation into long and short range for the exchange energy via a range parameter ω which defines an error function. The most common parameter pair is $\alpha = 0.25$, similar to the PBE0, and $\omega = 0.2$ and is also known as the standard HSE or HSE06 functional.

These parameters can then be adjusted to match the band gap or any other physical property. One property is the previously discussed parabolicity of the semi-local functionals with partial removal or addition of electrons.

2.1.3 Tight-binding

So far, we have discussed various approximations to calculate the electronic properties of solids. Density functional theory (DFT) is a well-established, widely used *ab-initio* approach that yields very accurate results. However, it is computationally expensive and is not well suited for large atomic systems. Semi-empirical methods, such as the tight-binding formalism, provide a more computationally efficient alternative.

The tight-binding method approximates the full many-body Hamiltonian by projecting the electronic wavefunction onto a localized basis. The Hamiltonian is parameterized, and the parameters are often obtained by fitting to results from *ab-initio* calculations. This method enables calculations for systems with many more atoms, which is especially useful for defective systems that require large supercells.

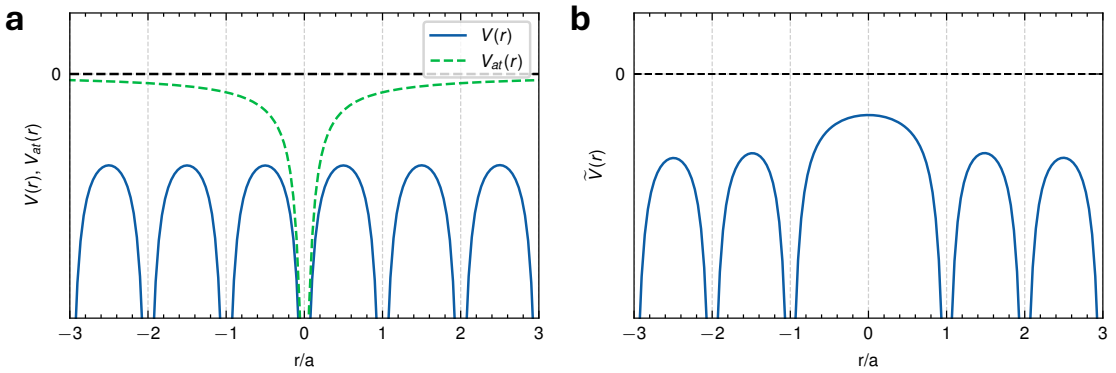

Figure 2.3: Illustration of potential energies important within the tight-binding approximation. **a** Crystal energy $V(r)$ including the contribution of an atom $V_{at}(r)$ at site 0. **b** Perturbation potential $\widetilde{V}(r)$ that impacts the states at site 0. Inspired by [33].

Tight-binding Hamiltonian

The main assumption of the tight-binding method is that deep, localized states in a solid have properties similar to those of electrons in isolated atoms. Solids form through the overlap of electron wavefunctions. However, this overlap primarily involves the higher-energy, delocalized states, while the deeper states remain confined within the atomic potentials (Figure 2.3). Consequently, these deeper states become part of the electronic structure of the solid but are perturbed by the potentials of neighboring atoms. We can therefore start with the unperturbed states $\hat{\mathcal{H}}_{at} = \frac{\hbar^2}{2m}\nabla^2 + V_{at}$:

$$\hat{\mathcal{H}}_{at}\phi^n_{at}(\mathbf{r}) = \epsilon^n_{at}\phi^n_{at}(\mathbf{r}) \; . \tag{2.29}$$

We assume that the wavefunctions ϕ^n_{at} are localized and can be expanded as Bloch waves to account for crystal symmetry:

$$\psi^n_{\mathbf{k}}(\mathbf{r}) = \frac{1}{\sqrt{N}} \sum_{\mathbf{R}} e^{i\mathbf{k}\cdot\mathbf{R}} \phi^n_{at}(\mathbf{r} - \mathbf{R}) \; . \tag{2.30}$$

Here N is the number of atoms, $\mathbf{k}$ is the wave vector, and n is the band index. These Bloch waves are eigenstates of the crystal Hamiltonian:

$$\hat{\mathcal{H}}\psi^n_{\mathbf{k}}(\mathbf{r}) = \left[\frac{\hbar^2}{2m_e}\nabla^2 + V(\mathbf{r})\right]\psi^n_{\mathbf{k}}(\mathbf{r}) = E_n(k)\psi^n_{\mathbf{k}}(\mathbf{r}) \ . \tag{2.31}$$

The external potential $V(\mathbf{r})$ is expressed as a sum of atomic potentials, $V(\mathbf{r}) = V_{at}(\mathbf{r} - \mathbf{R}) + \sum_{\mathbf{R}'\neq\mathbf{R}} V_{at}(\mathbf{r} - \mathbf{R}')$, so that the Hamiltonian can be rewritten by separating the isolated atomic Hamiltonian from the crystal potential contributions

$$\hat{\mathcal{H}} = \hat{\mathcal{H}}_{at} + \tilde{V}(\mathbf{r} - \mathbf{R}') \ , \tag{2.32}$$

with $\tilde{V}(\mathbf{r} - \mathbf{R}') = \sum_{\mathbf{R}'\neq\mathbf{R}} V_{at}(\mathbf{r} - \mathbf{R}')$. We can evaluate the Hamiltonian:

$$\begin{aligned}
\langle\psi^n_{\mathbf{k}}(\mathbf{r})|\,\hat{\mathcal{H}}\,|\psi^n_{\mathbf{k}}(\mathbf{r})\rangle =& \frac{1}{N}\sum_{\mathbf{R},\mathbf{R}'} e^{i\mathbf{k}\cdot(\mathbf{R}-\mathbf{R}')}\,\langle\phi^n_{at}(\mathbf{r}-\mathbf{R}')|\,\hat{\mathcal{H}}\,|\phi^n_{at}(\mathbf{r}-\mathbf{R})\rangle \\
=& \underbrace{\langle\phi^n_{at}(\mathbf{r})|\,\hat{\mathcal{H}}_{at}\,|\phi^n_{at}(\mathbf{r})\rangle}_{\epsilon^n_{at}} + && (2.33)\\
& \underbrace{\langle\phi^n_{at}(\mathbf{r})|\,\tilde{V}(\mathbf{r})\,|\phi^n_{at}(\mathbf{r})\rangle}_{\beta_n} + && (2.34)\\
& \sum_{\mathbf{R}'\neq 0} e^{-i\mathbf{k}\cdot\mathbf{R}'}\,\underbrace{\langle\phi^n_{at}(\mathbf{r}-\mathbf{R}')|\,\tilde{V}(\mathbf{r}-\mathbf{R}')\,|\phi^n_{at}(\mathbf{r})\rangle}_{\lambda(\mathbf{R}')} \ . && (2.35)
\end{aligned}$$

Here, ϵ^n_{at} represents the eigenvalues of the isolated atom, while β_n is correcting the atomic eigenvalues due to the influence of the potentials from neighboring atoms. The final term, $\lambda(\mathbf{R}')$, is the hopping or transfer integral, which quantifies the energy required for an electron to hop to an orbital at site $\mathbf{R}'$. Since the atomic orbitals are highly localized and decay rapidly, we can restrict the sum to the n-th nearest neighbor. This localization allows the parameterization to rely on only a few parameters. The energy dispersion of the n-th band is given by

$$E^n_{\mathbf{k}} = \epsilon^n_{at} + \frac{\beta_n + \sum_{\mathbf{R}\neq 0} e^{-i\mathbf{k}\cdot\mathbf{R}}\lambda(\mathbf{R})}{1 \ + \sum_{\mathbf{R}\neq 0} e^{-i\mathbf{k}\cdot\mathbf{R}}\alpha(\mathbf{R})} \ , \tag{2.36}$$

where $\alpha(\mathbf{R}) = \langle \phi_{at}^n(\mathbf{r}-\mathbf{R})|\phi_{at}^n(\mathbf{r})\rangle$ accounts for the non-orthogonality between atomic orbitals at different sites (note that ϵ_{at}^n are not true crystal eigenvalues, hence the orbitals are not orthogonal). Wannier functions or a linear combination of atomic orbitals (LCAO) can be used to obtain orthogonal tight-binding.

In systems with multiple orbitals and atoms per unit cell, this formulation takes on a matrix form:

$$\hat{\mathcal{H}} = \begin{bmatrix} \epsilon_n + \sum_{R_i \neq 0} t_{nn} e^{i\mathbf{k}\cdot\mathbf{R}_i} & \sum_i t_{nm,i} e^{i\mathbf{k}\cdot\delta_i} \\ \sum_i t_{mn,i} e^{-i\mathbf{k}\cdot\delta_i} & \epsilon_m + \sum_{R \neq 0} t_{mm} e^{i\mathbf{k}\cdot\mathbf{R}_i} \end{bmatrix}. \tag{2.37}$$

In standard notation, ϵ_n is used for onsite energy, and t signifies the hopping matrix elements. $\mathbf{R}_i$ indicates lattice vector displacements for identical atoms, and δ_i represents displacements between distinct atoms.

For a supercell containing many atoms, the Hamiltonian matrix enlarges; however, in the absence of perturbations, its entries remain the same as those for a single unit cell. When a defect is introduced, all these parameters can be perturbed, making the fitting process cumbersome. We will discuss approaches to address this issue when we cover defects.

2.1.4 Example: Graphene

We introduce graphene as a simple example to illustrate the methodology. The unit cell of graphene consists of two carbon atoms arranged in a hexagonal structure. Its p_z orbitals form a cone around the high-symmetry point K (Figure 2.4 **a**). The simplest description includes only the first nearest-neighbor hopping, so the Hamiltonian is given by:

$$\hat{\mathcal{H}} = \begin{bmatrix} 0 & \sum_i t e^{i\mathbf{k}\cdot\delta_i} \\ \sum_i t e^{-i\mathbf{k}\cdot\delta_i} & 0 \end{bmatrix}. \tag{2.38}$$

Without loss of generality, we can set the carbon atoms' onsite energies to zero and fit the hopping terms to best reproduce the *ab-initio* band structure. A hopping matrix element

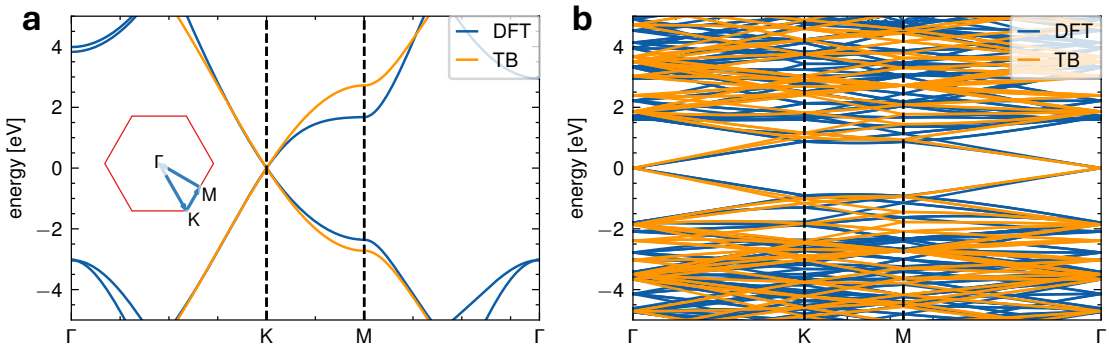

Figure 2.4: **a** DFT (blue) and tight-binding (orange) band structures of graphene. The tight-binding model includes first nearest-neighbor hopping terms and describes the system well for the cone around K-point. The sub figure on the left illustrates the first Brillouin zone and the k-path chosen for the band structure. **b** DFT (blue) and tight-binding (orange) band structures of a $9 \times 9 \times 1$ graphene supercell. The increased lattice vector reduces the Brillouin zone which results in a backfolding of the bands.

of $t = -2.72$ [29] reproduces the region around K-point fairly well (Figure 2.4). Although capturing the asymmetry between the two bands requires including hopping beyond the first nearest neighbors, this approach is sufficient for now and will later serve as the basis for modeling defects.

The defects will later be described using a supercell approach. A supercell features an increased lattice constant, which reduces the first Brillouin zone and leads to band backfolding (Figure 2.4 **b**). Although the number of parameters remains unchanged for the pristine case, fitting is typically performed on the unit cell because disentangling the folded bands to extract the correct eigenvalues is challenging.

2.1.5 Density of states

Another important quantity, we will use in this thesis is the density of states (DOS), that is the number of states for a given energy and it calculates as follows

$$D(E) = \frac{1}{V} \sum_{n,k} \delta(E - E_n(\mathbf{k})) \ , \tag{2.39}$$

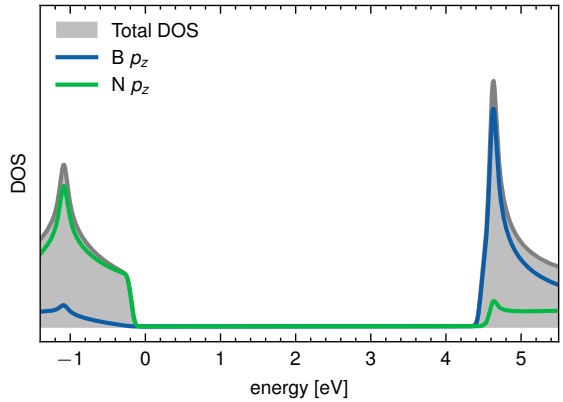

Figure 2.5: Total density of states $D(E)$ and the contribution of the atom- and orbital-projected densities of states $D_{i,l}(E)$ for the p_z orbitals of boron (blue) and nitrogen (green) in hBN.

where $\delta(E - E_n(\mathbf{k}))$ is a broadened Dirac-delta function, n is the band index and the sum performed is over a dense grid of wave vectors **k**. The DOS is a global property, as it includes all the eigenvalues of the system. However, we can project the DOS onto specific orbitals and atomic sites. The atom- and orbital-projected density of state (PDOS) is expressed as

$$D_{i,l}(E) = \frac{1}{V} \sum_{n,\mathbf{k}} |\langle \phi_{i,l} | \psi_n(\mathbf{k}) \rangle|^2 \delta(E - E_n(\mathbf{k})) \,, \tag{2.40}$$

where l is the orbital index and i the atomic site (Figure 2.5). This adds a spatial resolution to the problem which we can use to evaluate, for example, a defect-related orbital at specific sites.

In addition, the density of states gives us information about the localization of a defect state. If a state is localized in real space it is delocalized in the reciprocal space, which results in an undispersive band and therefore a narrow peak in the density of states.

2.2 Point defects in semiconductors

One of the reasons semiconductors have become popular is the ability to tune their properties through defects, enabling important applications. However, defects can also introduce unwanted effects that can reduce the performance of the device. Understanding

the properties of defects is therefore not only an interesting topic but also crucial for improving these devices. Before discussing how to model defects, let us first classify them, beginning with shallow and deep defects. Shallow defects are associated with defect states that lie energetically close to the respective band edges. As a general rule of thumb, room temperature is sufficient to efficiently ionize these shallow defects. Although k_BT is approximately 30 meV, the temperature-dependent electron distribution (described by the Fermi-Dirac distribution) allows electrons to populate states above 30 meV.

Generally, shallow defects are desirable in devices where doping is crucial because they provide additional charge carriers to the crystal. However, defects with opposite doping can also reduce the efficiency due to compensation. These defects can be described using the effective mass theory [45]. Consequently, their impact on the lattice is minimal. They are often associated with substitutional atoms, which disturb the lattice only slightly while introducing extra charges into the system.

Compared to shallow defects, deep defects or centers are localized, and their states are located deep within the band gap. Defect states within the band gap can act as recombination centers which, for example, counteract carrier separation, as discussed in the introduction. Although deep defects are generally undesirable in solar cells, some of them exhibit single-photon emission. Due to their localized nature, the tight-binding approach is a natural choice for describing deep defects. However, as we have discussed and will further see, a parameterization for defects introduces multiple difficulties. After further categorizing the defects, we will introduce different types of defects and discuss how to calculate their properties using different methods.

2.2.1 Defect types

Defects can be classified into several types, depending on their dimensionality and the number of atoms involved. In this thesis, we focus on point defects and defect complexes, which involve one or a small number of atoms, respectively. Defects can also extend along multiple atomic rows, forming so-called line defects (e.g., dislocations); however, these will not be discussed in detail as they fall outside the scope of this thesis.

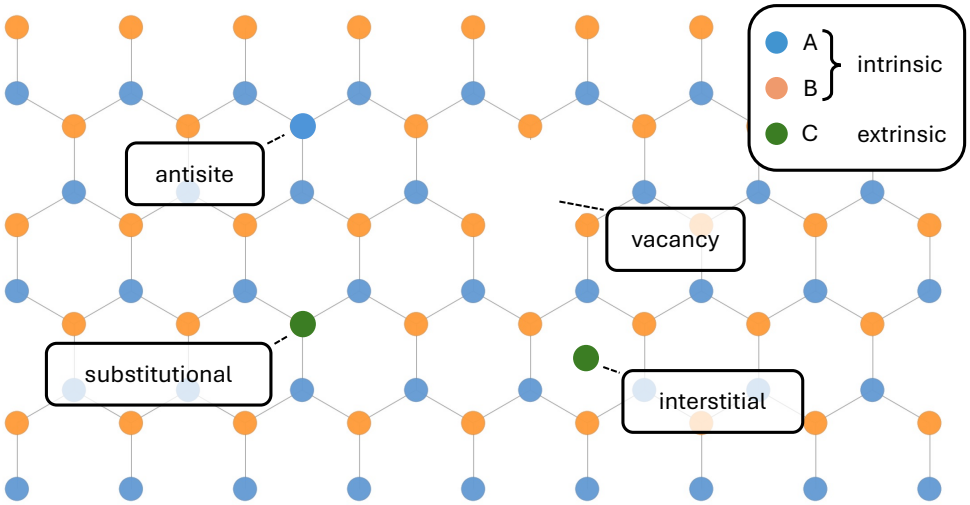

Figure 2.6: Two-dimensional hexagonal lattice consisting of atom A and B with different defect types. Antisites are atoms from the host material that are at a different lattice sites and a vacancy is missing atom. Both of them are intrinsic defects. Interstitial atoms are atoms that are at a different position than the defined by the lattice and they can either be intrinsic or extrinsic. Substitutional defects are atoms from a different element (C) positioned at a lattice site. They are extrinsic defects.

Point defects can be further classified as either intrinsic or extrinsic. Intrinsic (native) defects involve only atoms that are already present within the host material, such as vacancies and antisite defects, both illustrated in Figure 2.6. A vacancy (denoted by V_A) refers to a missing atom at lattice site A, while an antisite defect occurs when an atom occupies the lattice site of another type (e.g., A_B). Extrinsic defects, by contrast, involve foreign atoms introduced into the host material. These include substitutional defects (denoted as C_A, element C substituting for element A), and interstitial defects, where atoms (intrinsic or extrinsic) occupy positions between lattice sites (denoted by C_i).

2.2.2 Computational descriptions of defects: Density functional theory

When studying point defects or defect complexes, we must increase the cell size of our system. Keeping the same number of atoms as in the unit cell would effectively produce a new material as a result of the periodic boundary conditions. These enlarged cells, called supercells, need to be large enough to properly capture the impact of the defect on

surrounding host atoms. However, a larger number of atoms and consequently electrons increase the complexity of the problem, leading to greater computational cost. Although computationally less expensive methods, such as the effective mass theory or the tight-binding method, can also be employed to study defects, these methods are appropriate only for certain defect types. Therefore, identifying the type of defect is a crucial first step prior to conducting further investigations.

Density functional theory is applicable regardless of the defect type and used to calculate defect properties for various applications [46, 47]. While it may be tempting to directly associate single-particle Kohn-Sham states with actual defect states, we must keep in mind that these Kohn-Sham states are artificial constructs designed to reproduce only the ground-state density of the interacting electron system. Hence, the physically meaningful quantities obtained from standard DFT calculations are the ground-state electron density and the corresponding total energy of the defective system. Using this ground-state energy, various approaches enable us to explore defect properties, providing a solid basis for comparison with experimental results.

Formation energies

One of the key quantities used to describe defect properties is the defect formation energy. It provides insight into the stability of defects in a given system. Specifically, the formation energy of a defect X with charge state q is given by [46–50]:

$$E^{\mathsf{f}}[X^q] = E_{\mathsf{tot}}[X^q] - E_{\mathsf{tot}}[\mathsf{bulk}] + \sum_i n_i \mu_i + q E_F + E_{\mathsf{corr}} \, . \tag{2.41}$$

Here, the term $E_{\mathsf{tot}}[X^q] - E_{\mathsf{tot}}[\mathsf{bulk}]$ represents the total energy difference between the relaxed defective supercell and the pristine supercell of equivalent size. To form the defect, we either add or remove n_i atoms, for which a chemical potential μ_i is introduced, mimicking the conditions of a reservoir and reflecting varying chemical environments during growth. This allows the comparison of calculated formation energies with experimental data across different growth conditions. Defects may also accumulate charges q from

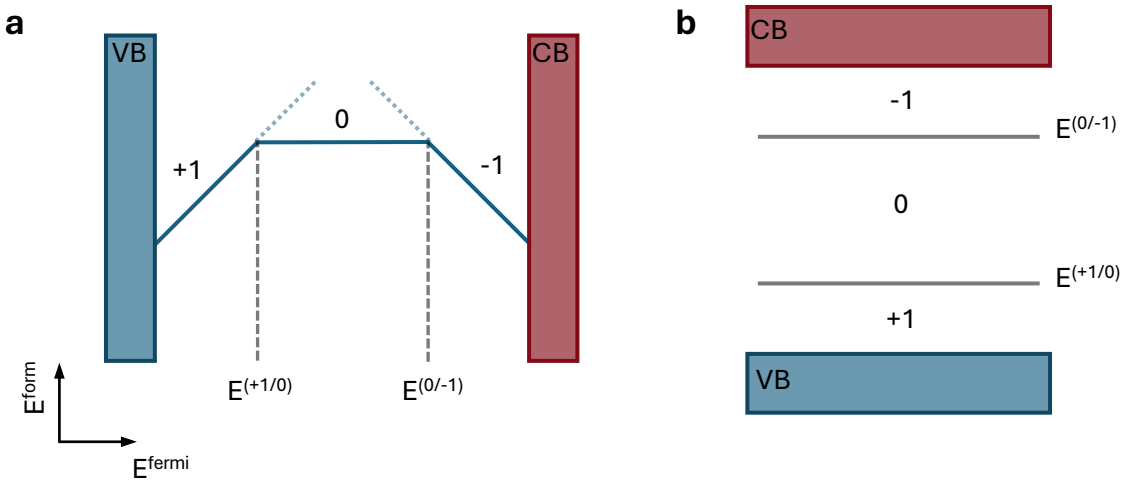

Figure 2.7: Illustration of the charge transition level. **a** Shows Example for a formation energy vs. fermi energy plot of an amphoteric defect. The formation energies of charged defects have a slope due to the qE_f term in equation 2.41. Inspired by Freysoldt et al. [58]. **b** Charge transition level plot for an amphoteric defect if the formation energy is not relevant.

the crystal; the corresponding electron reservoir in this case is represented by the Fermi energy E_F. The correction term E_{corr} addresses finite-size effects, which result from not achieving the dilute limit.

In recent years, considerable effort has been devoted to developing accurate correction schemes for finite-size effects [51–57]. In this thesis, we employ the correction scheme introduced by Freysoldt, Neugebauer, and Van de Walle [51, 53], as well as its extension to vertical transitions [56, 57].

Charge transition levels

The total energy gives access to another important concept: the charge transition level (CTL), also known as the thermodynamic transition level. As previously discussed, defects may exist in various charge states and each may significantly influence the surrounding lattice structure. By comparing the formation energies of two different charge states as a function of the Fermi energy, we can determine the specific energy at which a defect transitions from one stable charge state to another. Shifts in the Fermi energy alter the

availability of electrons, thus influencing the stability of defects' charge state. The specific Fermi energy at which two defect charge states have identical formation energies defines the charge transition level. Thus, CTLs provide insight into the stability of the charge states of defects under given doping conditions.

This concept can be illustrated by plotting the formation energy as a function of the Fermi energy (see Figure 2.7 **a**) for various defect charge states. Due to the explicit linear dependence of the formation energy on the charge in term qE_F, different charge states exhibit distinct slopes, leading to intersections in their respective energy lines. These intersection points represent the energies at which two charge states are equally stable and are known as charge transition levels. Specifically, the CTL between two charge states q_1 and q_2 is calculated as follows [46, 47, 50]:

$$\epsilon(q_1/q_2) = \frac{E^f[X^{q_1}, E_F = 0] - E^f[X^{q_2}, E_F = 0]}{q_2 - q_1} \,. \tag{2.42}$$

Here, $E^f[X^q, E_F = 0]$ is evaluated at the valence band maximum, serving as the reference energy. If the exact formation energies are not of primary interest, it is sufficient to plot only the charge transition levels, as shown in Figure 2.7 **b**. These charge transition levels should not be confused with electron states; instead, they indicate the specific positions of the Fermi level at which transitions between different defect charge states occur.

Using the concept of charge transition levels, we can further classify defects as donors, acceptors, or amphoteric. Acceptors have empty defect states that can accept electrons from the host crystal, therefore becoming negatively charged. In contrast, donors provide electrons to the host, resulting in positive charge states. Amphoteric defects exhibit both behaviors, acting either as donors or acceptors, depending on the position of the Fermi energy (Figure 2.7 **a**). By analyzing the CTLs, we can determine which charge states become energetically favorable at different Fermi energy positions. Moreover, the position of the charge transition levels relative to the band edges indicates whether the defects are shallow or deep: deep defects have transition levels farther within the band gap, while shallow defects are charged for all Fermi energies or have transition levels close to the

band.

Thus, charge transition levels provide valuable information about defect properties. However, caution is required when comparing these thermodynamic levels directly to results from optical experiments, even though such comparisons are tempting, especially when presented graphically as in Figure 2.7 **b**. However, photoluminescence measurements involve charge neutral transitions, and these transitions occur without full atomic relaxation. In the following, we will discuss the conditions under which charge transition levels can provide information to optical measurements.

Optical transition levels

Although many-body methods would ideally be employed to accurately determine the optical properties of defects, their application is currently limited to simpler systems. As a result, we often rely on total-energy calculations within the single-particle approximation. For thermodynamic transition levels, both charge states were considered in their fully relaxed, ground-state atomic configurations. This scenario corresponds well to experiments in which the defect has sufficient time (on the order of phonon relaxation timescales) to achieve structural equilibrium after changing its charge state, such as deep-level transient spectroscopy. In contrast, when analyzing optical properties, electron excitations occur almost instantaneously relative to ionic motion, leaving the lattice no time to relax. Consequently, we must account for this distinction when comparing our calculations with optical experiments.

The coupling between the electronic and lattice structure can be intuitively represented using a configuration coordinate diagram (CCD). In a CCD, the complex multidimensional atomic motion around a defect is simplified to a single-configuration coordinate (see Figure 2.8 **b**). The thermodynamic transition level corresponds to the Fermi energy at which the two charge states have equal energies at their respective equilibrium positions, indicated by the minima of their potential energy curves. As an initial approximation, these potential-energy curves are commonly described using a harmonic (parabolic) approximation near equilibrium.

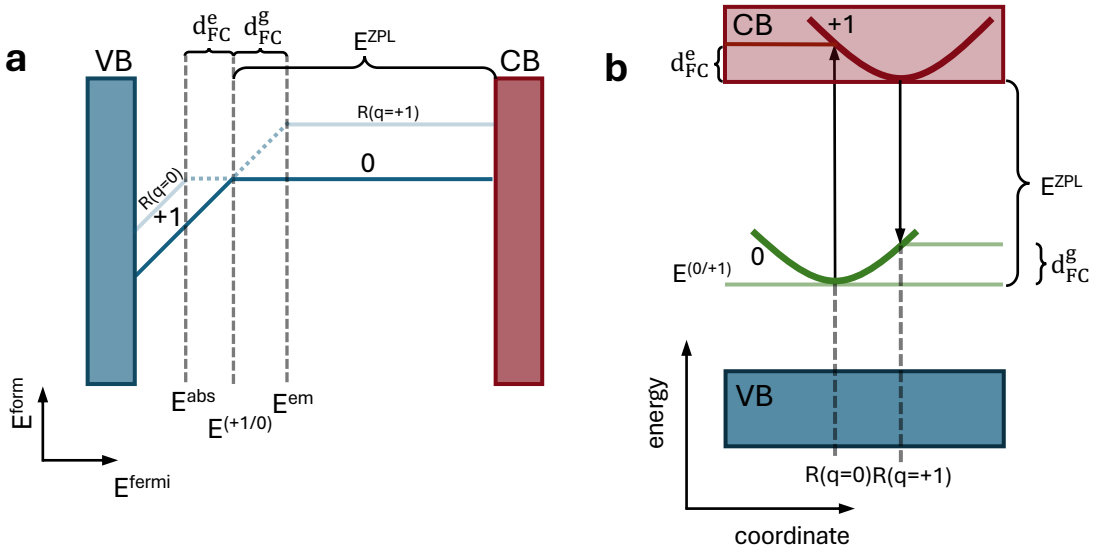

Figure 2.8: Illustration of the optical transition level. **a** Example for a formation energy vs. fermi energy plot for a charge transition $E^{(+1/0)}$. Indicated in light blue are the formation energies with the atomic ground state of the other charge state, which results in a higher energy. Keeping one atomic configuration therefore leads to different transition level, called optical transition level. We can identify Franck-Condon shifts d_{FC} and a zero-phonon line E^{ZPL}. Inspired by Freysoldt et al. [58]. **b** configuration coordinate diagram embedded between the valence band and conduction band. The example shows a transition from a defect in its ground state 0 to the conduction band, which charges it +1. We observe a Franck-Condon shift, due to the different atomic equilibrium of the charged state. The same is observed for the emission from the conduction band to the defect.

For optical transitions, we assume that the atomic configuration remains fixed during the electronic excitation, since the electronic transition is much faster than any atomic movement. Within the CCD, this corresponds to a vertical transition, analogous to the Franck-Condon approximation used in molecular systems. The energy difference between the optical (vertical) and the thermodynamic transition levels is therefore referred to as the Franck-Condon shift, denoted by d_{FC}.

We now consider the complete optical cycle for a defect using the corresponding configuration coordinate diagram. Without loss of generality, we assume that the defect is initially its neutral state. When the electron absorbs a photon, the electron is excited from the defect state into the conduction band at an energy E_{abs}. The local environment around the defect effectively becomes positively charged, although its atomic configuration remains that of the original neutral state.

In this excited state, the positively charged defect can relax to its new equilibrium atomic configuration corresponding to the +1 charge state. The energy difference between the initial (neutral) and the relaxed positively charged configurations defines the excited-state Franck-Condon shift, denoted by d^e_{FC}. The absorption energy can therefore be expressed as:

$$E_{\text{abs}} = E_{\text{therm}} + d^e_{\text{FC}} \, . \tag{2.43}$$

Subsequently, an electron from the conduction band may recombine with the defect, occupying the now-empty defect state and emitting light at energy (E_{em}). This recombination again changes the defect's local charge state back to neutral. Analogously to absorption, the atomic configuration immediately after emission is not yet relaxed, resulting in another energy shift compared to the thermodynamic transition level, called the ground-state Franck-Condon shift (d^g_{FC}):

$$E_{\text{em}} = E_{\text{therm}} - d^g_{\text{FC}} \, . \tag{2.44}$$

In practice, comparing calculated absorption and emission processes with experimental data requires identifying the most stable charge state of the defect, as this defines the ground-state geometry. Additionally, one must distinguish between two distinct types of charge transitions. Transitions to the defect (such as absorption from the valence band maximum, or emission from the conduction band minimum) involve adding an electron to the defect, while transitions from the defect (such as absorption to the conduction band, or emission to the valence band) correspond to electron removal. Therefore, we must carefully examine the electronic structure to determine which transitions are physically allowed. For instance, emission from the conduction band minimum to the defect state requires an empty defect-related Kohn-Sham state within the band gap. In contrast, emission from the defect state to the valence band requires that the defect-related state within the band gap is occupied.

Optical transition levels allow for an improved comparison with optical measurements. This is particularly important for deep defects, whose localized wavefunctions significantly impact the atomic configuration, resulting in large Franck-Condon shifts. In contrast,

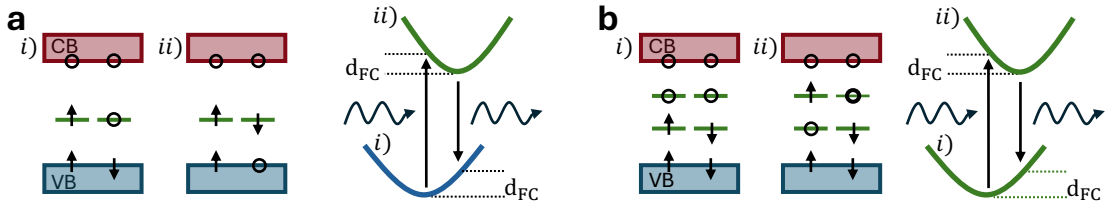

Figure 2.9: Illustration of the concept of the Δ-SCF approach. **a** Transition in which the excitation is well defined by a single Slater determinant by constraining the spin down electron from the valence band maximum into the defect state. **b** Transition in which either the spin up or the spin down electron can be promoted to the upper spin degenerate defect state.

shallow defects have a small impact on the surrounding lattice, leading to smaller Franck-Condon shifts, which makes their charge transition levels more directly comparable to experiments. However, we have to keep in mind that this method compares the total energies of systems with a different number of electrons, whereas experimental excitations are charge-neutral processes involving electron-hole interactions. The best agreement is therefore expected in cases with weak electron-hole coupling or transitions where charges rapidly separate from the defect site. An alternative total-energy-based method addressing some of these limitations will be introduced in the following section.

Δ-SCF approach

Both previously discussed approaches are based on transitions between different charge states, meaning that comparisons involve systems with varying electron counts. To better approximate the charge-neutral excitations observed experimentally, we can employ the Δ-SCF method (also known as constrained DFT). This method describes excited states using single Slater determinants, in which an electron is explicitly constrained to occupy a previously unoccupied Kohn-Sham state.

This approach is straightforward for singlet excitations, where only a single transition is possible (Figure 2.9 **a**), because the excited state can be clearly described by a single Slater determinant, allowing a direct calculation of its total energy within DFT. The situation becomes more complex for spin-degenerate cases or highly correlated systems, e.g. a defect-to-defect transitions similar to the carbon dimer (2.9 **b**). In these cases, the

excited state must be represented as an antisymmetric combination (linear combination) of spin-up and spin-down excitations, requiring careful consideration beyond a single Slater determinant [23]:

$$|S\rangle = \frac{1}{\sqrt{2}}\Big(|v_{\downarrow}d_{\uparrow}\rangle - |v_{\uparrow}d_{\downarrow}\rangle\Big) \,, \tag{2.45}$$

This excited state is not a single Slater determinant and we have to find another expression that allows us to calculate the total energy with. Mackoit-Sinkevičienė et al. [23] demonstrate this for the case of the carbon dimer, where they are able to find a good agreement with experimentally observed signatures. By constructing a mixed singlet-triplet state, one can calculate the excitation energy via:

$$E(S) = 2E(S/T) - E(T) \,, \tag{2.46}$$

where both $|S/T\rangle = |v_{\downarrow}d_{\uparrow}\rangle$ and $|T\rangle = |v_{\uparrow}d_{\uparrow}\rangle$ are expressed by single Slater determinants and with the assumption that both have the same geometric ground state (see Appendix A for details).

Similarly to the OTL, we can describe the process with a configuration coordinate diagram with different Franck-Condon shifts for emission and absorption. By construction, this approach includes to some degree electron-hole interaction, because the electron density reflects explicitly the missing electron (hole) in the lower Kohn-Sham orbital. However, we need well-defined Kohn-Sham states and the calculations can be more difficult to converge. For complex materials like $CuInS_2$ and $CuGaS_2$, we will reach the limits of constrained DFT calculations for the supercells that we use.

2.2.3 Computational description of defects: Tight-binding

An alternative method to model defects is the tight-binding approach. However, obtaining an accurate tight-binding description for defective supercells is challenging. We have illustrated some of the difficulties in Figure 2.10 and, ideally, we should address all of them to find the most accurate parameterization.

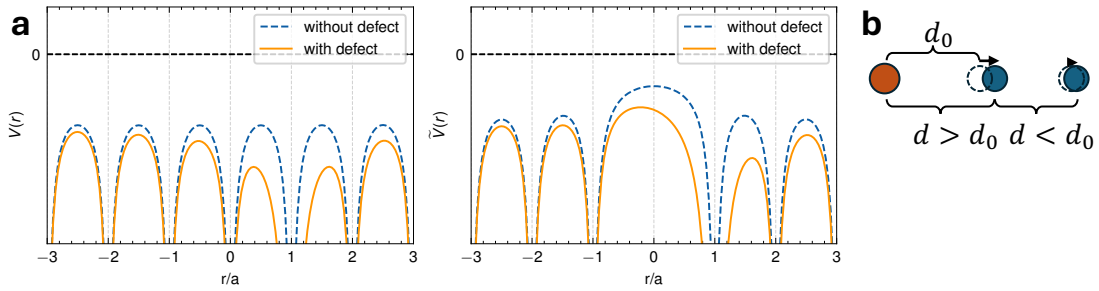

Figure 2.10: Illustration of difficulties for tight-binding parameterizations. **a** Impact on the tight-binding Hamiltonian. Left plot: Crystal potential energy for a pristine case (dotted blue) and with a defect (orange) at site 1. Right plot: Crystal potential without the atomic potential at site 0 ($\tilde{V}(r) = V(r) - V_{at}$). This is responsible for a shift of energies at the respective site and influences the transfer energy. A defect at a neighboring position impacts this shift. **b** A defect (red) has different properties than the host atoms and impacts the chemical environment in its vicinity. This impacts the geometry in its vicinity and results in different distances between hopping neighbors.

In Section 2.1.3, we have introduced the tight-binding Hamiltonian as:

$$\hat{\mathcal{H}} = \hat{\mathcal{H}}_{at} + \tilde{V}(\mathbf{r} - \mathbf{R}') \; . \tag{2.47}$$

Here $\tilde{V}(\mathbf{r} - \mathbf{R}')$ describes the perturbation of the crystal to the atomic Hamiltonian. Introducing a defect at a neighboring site affects the crystal potential in the vicinity and therefore $\tilde{V}(\mathbf{r} - \mathbf{R}')$ (Figure 2.10 **a**). This directly influences the onsite and transfer energies, both determined by this potential.

In addition, we need to consider that the defect has an impact on the geometry of the host crystal. This is reflected in altered distances between the neighbors (Figure 2.10 **b**) and therefore in the hopping energies. This effect can be long ranged, which means that in the case of a supercell many hopping parameters need to be adjusted. Furthermore, if we want to describe substitutional defects, it is essential to introduce new parameters that define the different chemical environment of the defect.

Due to these difficulties, creating a parameterization for defective supercells is still a topic of active research. One step forward has been done by Schattauer et al. [30]. By discretizing hopping distances and grouping parameters within defined thresholds, they introduce a distance dependence for the parameters which reduces the number

of fitting parameters while maintaining accuracy. To address the complexity of multiple tight-binding parameters and folded band structures, they employed a machine-learning approach. Different neural networks were trained on generated tight-binding data to learn the relationship between band structures and hopping parameters, allowing them to later optimize parameters directly from *ab-initio* band structures. The method was successfully applied to vacancy defects in graphene. For the extension to transition-metal dichalcogenides, they noted that capturing mid-gap states required additional adjustments to onsite energies for nearest neighbors [59].

An alternative approach from the pre-machine-learning era emphasizes adjustments to the onsite energies. Lambin et al. [29] used the simple tight-binding model that we have seen in Section 2.1.4 and modeled nitrogen defects in graphene by introducing distance-dependent onsite parameters. A Gaussian function centered at the defect site

$$\epsilon(d, \sigma, U) = \epsilon_C - |U| \cdot e^{-\frac{d^2}{2\sigma^2}} , \tag{2.48}$$

successfully captured the projected density of states and band structures. This model will serve as an example to explain the idea of how we will utilize the machine learning to obtain tight-binding parameters.

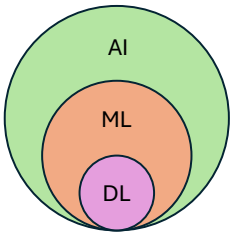

Figure 2.11: Illustration of different artificial intelligence (AI) applications. Machine learning (ML) and deep learning (DL) are both subcategories of AI, and DL is additionally a subcategory of ML.

2.3 Concepts of machine learning

The aim of an artificial neural network is to process information in the form of data and to understand its patterns. Over the past years, there has been significant advancement in the field of artificial intelligence (AI) which led to the development of many useful applications. Large language models (LLMs) have recently brought AI into everyday use, showcasing impressive abilities in translation, proofreading, and text generation [60]. Additionally, computer vision finds applications in robotics and the diagnosis of conditions like breast cancer [61–63]. Its importance in physics has also increased, as acknowledged by the most recent Nobel Prize, which was awarded to John J. Hopfield and Geoffrey Hinton for their contributions to the field [64].

Although there are many more examples in which AI offers a clear advantage, it also presents ethical concerns. As in most technological revolutions, there is always misuse, including discrimination, disinformation, scams, and more, some of which are listed on the website `Awful-ai` [65]. Another point of criticism is that the many resources needed for large neural networks shift the power further to big tech companies. Despite these concerns, it remains an interesting topic and a useful tool that we will leverage in this thesis and we will discuss the basic concepts in the following. Let us first define some of the frequently encountered terms: artificial intelligence (AI), machine learning (ML), and deep learning (DL).

In fact, deep learning is a subset of machine learning, which in turn is part of AI (Figure 2.11). This means that when we discuss deep neural networks, they can all be used

equivalently. AI includes basic algorithms, such as simple if-else conditions, aiming to replicate human behavior. Machine learning is a bit more specific and involves algorithms that learn from data, including tasks such as regression or classification. Deep learning, on the other hand, uses a large amount of data to train a deep neural network. Another key distinction is between supervised and unsupervised learning. With supervised learning, the data inputs and outputs are labeled, which is the framework we will employ when implementing a deep neural network for a regression task later on.

In the upcoming section, we discuss the general concept of deep neural networks before introducing two more different architecture types. This discussion is closely related to the neural network that we will use later. For further details, readers can consult several books that provide both introductory and advanced insights [66–69]. I also recommend the contributions of Grant Sanderson (also known as 3Blue1Brown), whose visualizations are outstanding and help clarifying the basic concepts [70].

2.3.1 Neural networks

The name 'artificial neural network' originates from its analogy to the nervous system. In it, neurons are interconnected, transmitting information through so-called synapses, where the signals are processed via electrical and chemical mechanisms [71]. In 1943 McCulloch and Pitts laid the foundation for a mathematical construct that models this biological process [72]. They introduced a threshold logic unit, which takes multiple weighted inputs and outputs either 0 or 1, similar to a neuron which either fires or does not fire to another neuron. This was further developed into a perceptron by Rosenblatt et al. who combined multiples of these into a network in 1957 [73, 74]. Minsky and Papert argued in 1969 that in order to solve more complex problems like the XOR logic gate, it needs multilayer perceptrons with non-linear activation functions [75]. It was only in 1986, that Rumelhart et al. reignited the spark by the introduction of back propagation[2][78]. These concepts still form the basis for many of today's artificial neural networks architectures and we will discuss them briefly.

[2]Comprehensive History can be found in [76, 77]

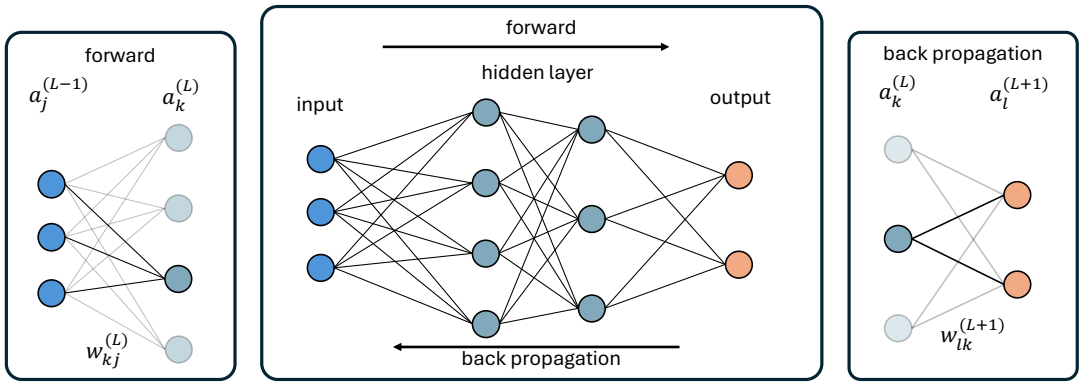

Figure 2.12: Illustration of a multi-layer perceptron. The network consists of fully connected linear layers. Left penal: Illustration of the forward process to clarify the sub- and superscripts. Right penal: Illustration of the back propagation to clarify the sub- and superscript.

Forward

A multilayer perceptron (MLP) consists of multiple fully connected linear layers where each layer is built of multiple neurons (Figure 2.12). Each neuron processes the output from a previous layer a_j^{L-1}:

$$a_k^L = \sigma\Big(\underbrace{\sum_j w_{jk} \cdot a_j^{L-1} + b_j}_{z^{(L)}}\Big) \,, \tag{2.49}$$

where it is weighted with w_{jk}. The superscript L refers to the layer, whereas the subscripts k, j denote the respective neuron in each layer. Additionally, a bias b can be added to the output. The weighted input and the bias are wrapped in an activation function σ which introduces non-linearity, which is essential to fit non-linear problems. The inner function of σ is denoted as $z^{(L)}$.

The number of layers and the number of neurons per layer is arbitrary, and Hornik et al. [79] have proven that a neural network can fit any continuous function by a sufficiently high number of parameters. This is known as the *universal approximation theorem*.

Training

We have understood how data are processed by each neuron from input to output, with the *forward* process that involves multiple matrix multiplications according to Equation 2.49. In order to train the network, we need to evaluate the performance by defining a metric for the neural network. In case of supervised learning, we know the expected value (label) for each input, and we can define a loss or cost function that compares it to the output of the neural network. One of the standard loss/cost functions is the mean-squared error (MSE):

$$\mathcal{C} = \frac{1}{\mathsf{N}} \sum_j (a_j^L - y_j)^2 \,, \tag{2.50}$$

where a_j^L is the j-th output of the final layer and y_j the corresponding label. In general, one is free to define specific loss functions for a given task.

We train the neural network by minimizing the loss through adjustments to each weight and bias. The cost function, $\mathcal{C}(\theta)$, is a high-dimensional function because it includes all adjustable parameters, $\theta = (w, b)$. Finding the minimum of a function involves the derivative or, in the case of a many dimension, the negative gradient $-\nabla\mathcal{C}(\theta)$. The negative gradient points in the direction of steepest descent and can therefore be used by optimization algorithms like gradient descent (sometimes referred to as steepest descent):

$$\theta = \theta - \eta \nabla \mathcal{C}(\theta) \,. \tag{2.51}$$

Here η is the learning rate, a parameter that defines the step size for the direction of descent. Since all inputs (data points) must be evaluated before updating the weights and biases, a more common way is to batch the input data and evaluate the mean for each of these batches (subsets) of size B. This is called stochastic gradient descent (SGD), which evaluates the average of each batch $\mathcal{C}^B(\theta) = \frac{1}{B} \sum_i^B \mathcal{C}_i(\theta)$. Both the learning rate and the batch size are hyper parameters that can be adjusted for a given problem and need to be monitored for optimal training.

Backpropagation

The gradients required for the optimization algorithms are obtained via backpropagation. As the name suggests, this process reverses the forward pass and allows us to calculate the partial derivatives of the adjustable parameters θ.

The cost depends on the output of the final layer a^L, which relies on the previous layer (L-1) and, more importantly, on its derivatives. To describe how a change in the weight or bias in layer (L) affects the cost, we can use the chain rule:

$$\frac{\partial \mathcal{C}}{\partial \theta^L} = \frac{\partial a^L}{\partial \theta^L} \frac{\partial \mathcal{C}}{\partial a^L} \tag{2.52}$$

$$\frac{\partial \mathcal{C}}{\partial \theta^L} = \frac{\partial z^L}{\partial \theta^L} \frac{\partial a^L}{\partial z^L} \frac{\partial \mathcal{C}}{\partial a^L} \ , \quad \theta = (w_{kj}, b_k) \ , \tag{2.53}$$

where we make use of the inner function $z^{(L)}$. If we want to update the parameters deeper into the network, layer L is not necessarily the final layer; its value depends on subsequent layers (L+n). Since the update for θ depends on $\frac{\partial \mathcal{C}}{\partial a^L}$, we need an expression for the derivative of a^L that includes the activations from layer (L+1). To achieve this, we can once again apply the chain rule to the activation function of neurons in layer a_l^{L+1}:

$$\frac{\partial \mathcal{C}}{\partial a_k^L} = \sum_k \frac{\partial z_l^{L+1}}{\partial a_k^L} \frac{\partial a_l^{L+1}}{\partial z_l^{L+1}} \frac{\partial \mathcal{C}}{\partial a_l^{L+1}} \ . \tag{2.54}$$

Both Equation 2.52 and Equation 2.54 enable the recursive calculation of partial derivatives for all weights and biases based on the previous layers (Figure 2.12).

In practice, these gradients can vanish or diverge. Activation functions such as *sigmoid* or *tanh* have small derivatives, which can cause the weight updates in earlier layers to be small due to the multiplication of these derivatives. One can either use more robust activation functions that do not suffer from this problem[3] or employ batch normalization. Batch normalization ensures that the activations in each layer remain within a reasonable

[3]Today, rectified linear units (ReLU) are the state-of-the-art activation functions used in multilayer perceptrons.

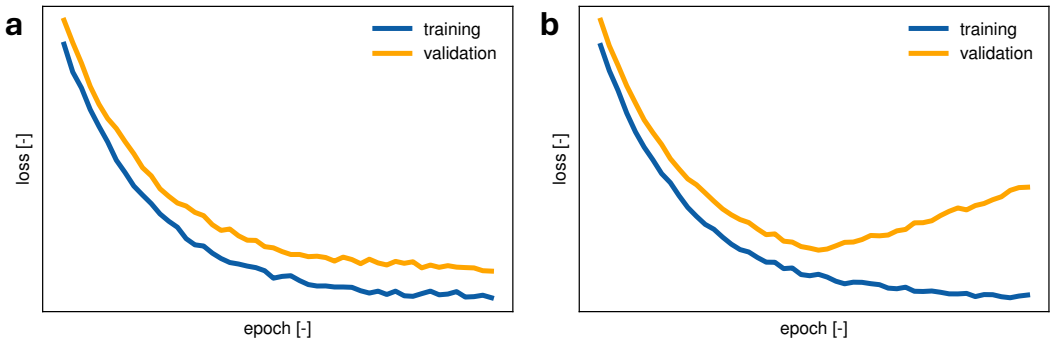

Figure 2.13: **a** Example training and validation loss curves for a successful training. **b** Example training and validation loss curves for overfitting.

range. Equally important is ensuring that the neural network generalizes well to unseen data and therefore avoid overfitting.

Overfitting

We can monitor training by dividing the dataset into training, validation, and test sets. The neural network is trained on a subset of the data (training set). After each epoch which involves processing all batches, the network is evaluated on unseen data (validation set) to assess its performance. The network tends to generalize better if both the training and validation losses decrease throughout the training and follow a parallel trend (Figure 2.13 **a**). In contrast, if the training loss continues to drop while the validation loss remains constant or increases, this indicates overfitting, as the network performs well only on familiar data (Figure 2.13 **b**). We can apply regularization methods like dropout or weight decay to avoid overfitting.

The MLP is a basic network that serves as a good starting point, though it is often combined with more complex architectures. In the following sections, we will introduce two other architecture types that we also applied in our neural network.

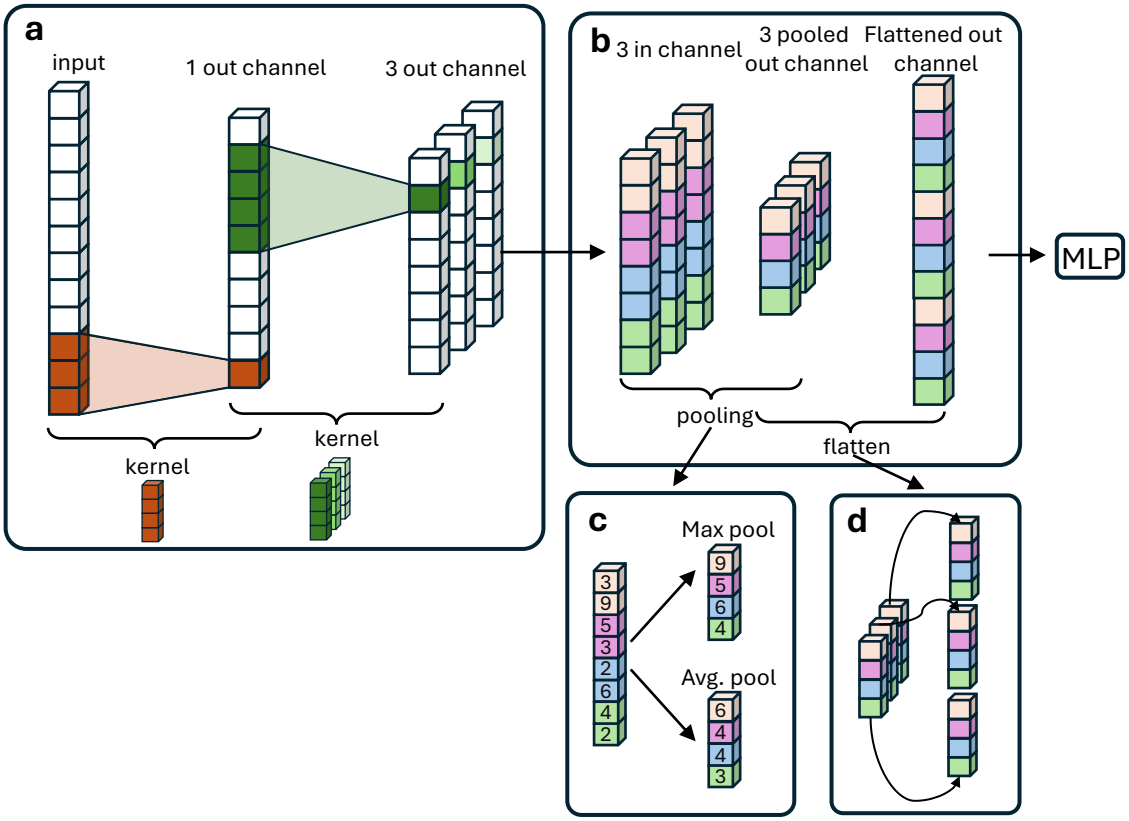

Figure 2.14: **a** Illustration of the forward process in a 1D convolutional layer. Kernel strides over the input and convolutes kernel and input. A convolutional layer can have multiple output channel, that have separate weights that can be trained. **b** The output channels are pooled before being flattened. The flattened output can be fed into any other part of the architecture like a MLP. **c** Depiction of two common pooling examples. **d** Depiction of the flattening process. Dimensionality is reduced to 1D again to act as an input for the next layer. Inspired by Schattauer [59].

Convolutional neural networks

Convolutional neural networks (CNNs) are well-suited for data with high local correlations, such as images or sound. They were first introduced by LeCun et al. [80], who set a new standard for image recognition by training on a predecessor of the famous MNIST[4] data set for handwritten digits [81]. Today, CNNs are applied to a variety of tasks, including 1D problems[5].

Although convolutional neural networks are often demonstrated using 2D examples,

[4]The MNIST data set is still one of the textbook examples of a classification task.
[5]A comprehensive overview of different applications can be found in [82]

we will concentrate on 1D inputs, given that a 1D problem will be addressed later in this thesis. A convolutional layer is designed to capture local correlations by mapping specific segments of data to subsequent layers [83]. This is achieved by using a *kernel* or *filter* to scan the input. The kernel performs a stepwise convolution of the input layer, which can include multiple kernels (see Figure 2.14 **a**). After convolution, it is common to pool the output before passing it to the next layer, thereby reducing the dimensionality [84] a standard step since the number of output channels may increase (see Figure 2.14 **b**). Common pooling methods include max pooling or average pooling, as depicted in Figure 2.14 **c**. After the final convolutional layer, the output is flattened into a 1D vector for further processing by subsequent layers.

Scaled Dot-product Attention Mechanism

Another important architecture type are transformer neural networks, which are used in many of today's large language models. A key component of a transformer is the self-attention mechanism which was designed to capture long-range dependencies while reducing computational cost and maximizing efficiency [85]. In our model, we use a self-attention mechanism to capture long-range dependencies between defect states and bands. To understand the self-attention mechanism we will again use the example of a 1D input vector, but it is also applicable for higher-dimensional inputs.

The idea is to divide the input into query, key, and value, by weighting each of the entry of the PDOS with fully connected linear layers to obtain three matrices ($\mathbf{Q}$, $\mathbf{K}$, $\mathbf{V}$). For the case of our 1D density of states $D(\mathbf{E})$, where the energy is restricted to a fixed interval $\mathbf{E} \in [E_0, E_n]$, this results in

$$\begin{aligned} \mathbf{q}_i &= \mathbf{D}(E_i)\mathbf{W}_Q \\ \mathbf{k}_i &= \mathbf{D}(E_i)\mathbf{W}_K \\ \mathbf{v}_i &= \mathbf{D}(E_i)\mathbf{W}_V \ , \end{aligned} \tag{2.55}$$

where the weight matrices ($\mathbf{W}_Q, \mathbf{W}_K, \mathbf{W}_V$) have trainable parameters. The queries and

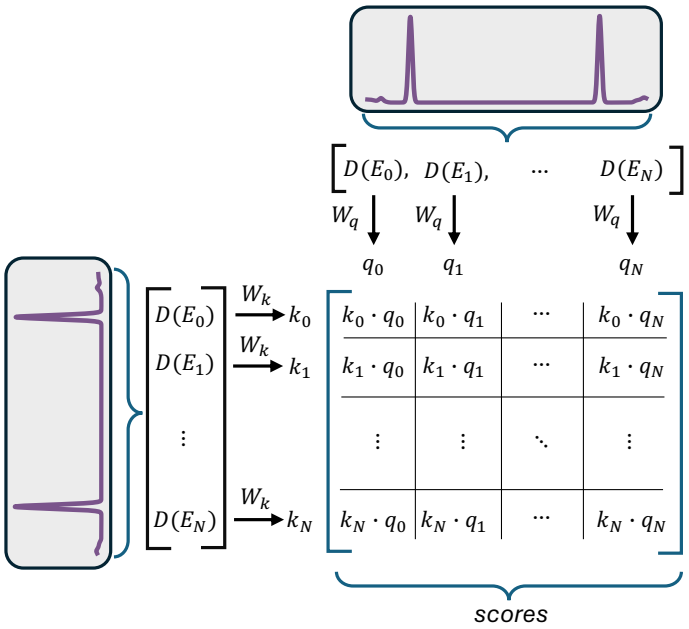

Figure 2.15: Illustration of the score matrix for an example of a 1D projected density of states. To obtain the query matrix Q and key matrix K, the entries $D(E_i)$ are processed by two linear layers with weights W_q and W_k, respectively. A matrix multiplication of the Q and K results in the attention weight matrix. If a softmax is applied this results in a probability distribution which captures long-range dependencies between the entries.

keys have the dimension d_k, and the values have the dimension d_v for each respective PDOS entry.

A matrix multiplication between the **Q** and **K** enables the model to capture the dependencies between their individual entries. If two entries align well, the corresponding entry in the matrix is high; otherwise, it is low. This process, which determines how one entry 'attends to' another, gives the mechanism its name attention. The entries of the so-called score matrix can be any number between $-\infty$ to ∞. A softmax function squishes these numbers between 0 and 1 such that each column can be associated with a probability distribution and the scaling factor $\sqrt{d_k}$ stabilizes the training. The score matrix that now represents the probability of how much two entries attend to each other is then multiplied by the values to obtain the final output. The self-attention can be expressed in a compact

form [85].

$$\text{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{softmax}(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d_k}})\mathbf{V} \ . \tag{2.56}$$

The network can be complicated by including multiple scaled dot-product attention heads, which is called multi-head attention. A multilayer perceptron typically follows self-attention and in a transformer architecture multiple of these are stacked after one another with input also from the previous self-attention layers.

2.3.2 Machine learning example: Tight-binding parameterization of N_C

In this thesis, machine learning will be used as a regression tool to obtain the tight-binding parameters from projected densities of states. The idea is to learn the correspondence between PDOS as the input and tight-binding parameters as the output (Figure 2.16.**a**). After the training process, the neural network predicts tight-binding parameters for a PDOS independent of the method.

We have previously introduced the nitrogen substitutional in graphene. Graphene is naturally well suited because it has a very simple tight-binding description and no band gap. This means that there are eigenvalues around the Fermi energy and therefore PDOS values.

We will employ the method previously introduced by Lambin et al. [29]. Here, only the onsite energies are varied by a distance-dependent perturbation with respect to the defect

$$\epsilon(d, \sigma, U) = \epsilon_C - |U| \cdot e^{-\frac{d^2}{2\sigma^2}} \ , \tag{2.57}$$

where ϵ_C is the pristine on-site energy, U the height of the perturbation potential, σ the standard deviation which is proportional to the width of the potential. Lambin et al. only varied the U parameter and kept $\sigma = 0.15$ as reported in [86] to obtain good agreement between tight-binding and DFT PDOS.

We will fit both parameters that define the Gaussian to find a tight-binding model that can reproduce the DFT PDOS. The pristine tight-binding model is described in Section 2.1.4. The parameters are based on a simple single-orbital model, including the first

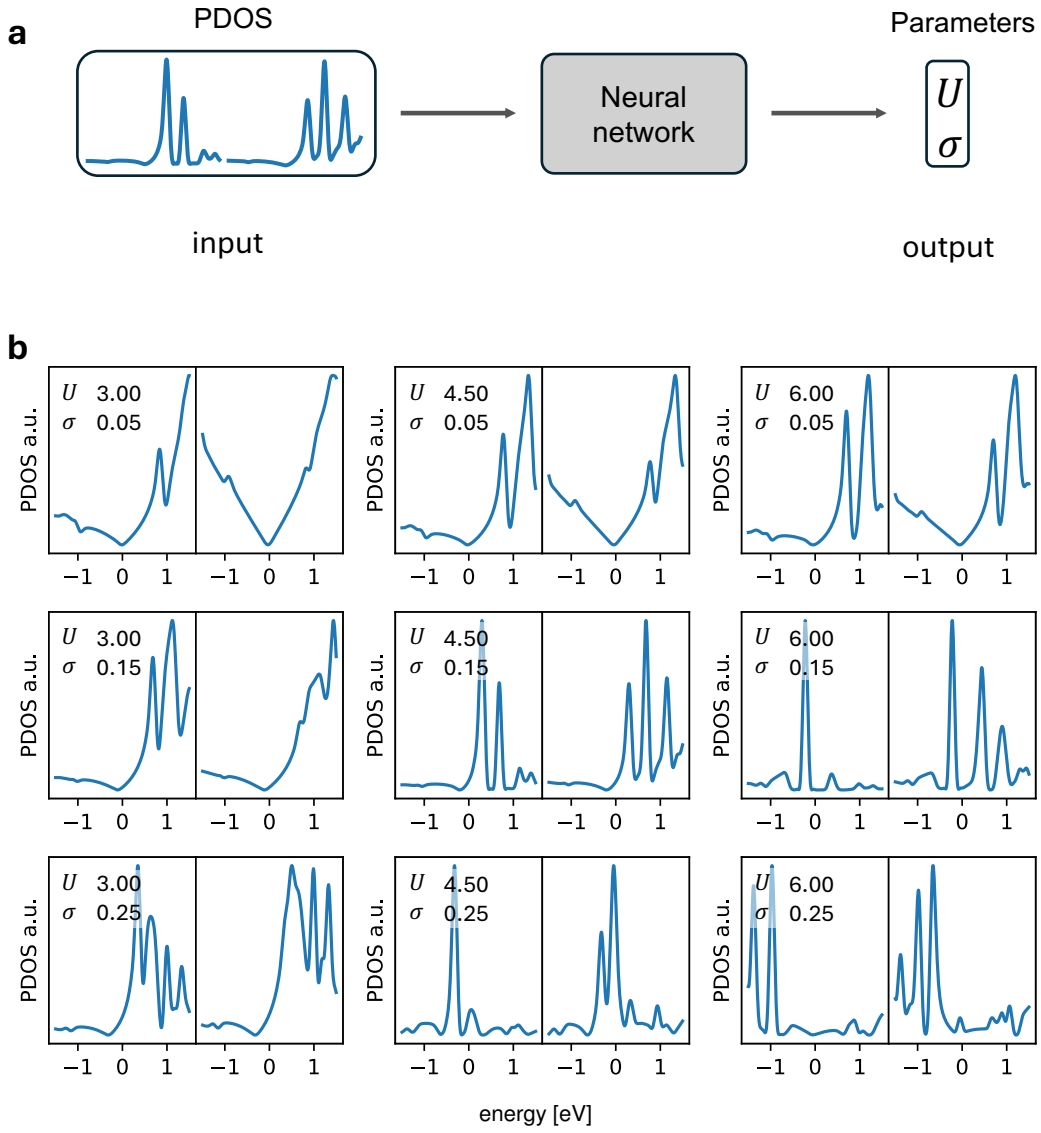

Figure 2.16: **a** Concept of the training process. The neural network maps the projected density of states (input) on the tight-binding parameters (output). **b** Example data points for the N_C represented by tight-binding PDOS for different perturbations defined by U and σ. PDOS is at defect site (left panels) and first nearest neighbor (right panels).

nearest-neighbor hopping parameter ($t = -2.72$) and the pristine on-site energy is set to 0. Both U and σ vary within a range according to the literature ($U \in \pm[3, 6], \sigma \in [0.05, 0.25]$). The data set is generated only by tight-binding calculations with different U and σ. We have illustrated the data set in Figure 2.16 **b**, where the corresponding PDOSs are plotted.

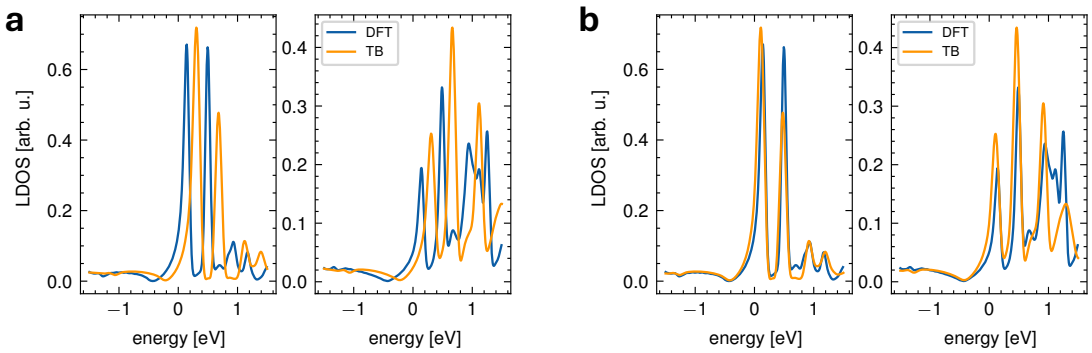

Figure 2.17: Projected densities of states for the N_C. **a** DFT (blue) and tight-binding PDOS (orange) calculated from the predicted parameters. **b** DFT (blue) and tight-binding PDOS (orange), where the tight-binding PDOS is shifted to align at the cone.

Since the data set is labeled, we have supervised learning, where we have the tight-binding PDOS as the input and the parameter pair of the Gaussian function as an output. For the training, we used the PDOS on the defect site and its first nearest neighbor.

After training, we present a DFT PDOS to the network to predict tight-binding parameters that reproduce the DFT result. For N_C, we find that even though the results were shifted, they reproduced the *ab-initio* results well. In Figure 2.17 **a** we observe that the predicted PDOS does not align correctly with the input.

However, we believe that this might be related to the simplicity of the defective model, which is not able to properly shift the Dirac cone to the right energy. By aligning the two PDOS (Figure 2.17 **b**), we can see that although the energetic positions were not correct, the machine learning network was able to predict parameters that reproduce the features.

Chapter 3

Electronic and optical properties of native point defects in $CuInS_2$ and $CuGaS_2$

This chapter is partly based on the article 'Electronic and optical properties of native point defects in $CuInS_2$ and $CuGaS_2$', currently under review in Physical Review Materials [87].

The following chapter will be an application of the previously discussed total energy evaluations for defects by means of hybrid functionals. We will first investigate the parameters of the hybrid functional, to align with the experimental band gap and the generalized Koopmans' theorem. We will extend the supercells by including 288 atoms compared to the previous studies with supercells of 144 atoms (216 for certain defects), which is an increase of minimum distances between defects by a factor of 1.48.

Many experimental results on defects in chalcopyrites are based on photoluminescence measurements. However, previous theoretical studies were based on the evaluation of the formation energy and the resulting charge transition levels. We want to improve

the comparability with these measurements and investigate the differences between the charge transition levels and optical transition levels. We consider that the electron excitation during an optical process is on a different time scale compared to that of ionization processes. The results highlight that this cannot be neglected for deep defects. Finally, we explore charge-neutral excitations with the Δ-SCF approach and compare the results with experiments.

3.1 Introduction

We have already discussed that defects are essential for many devices, including transistors, solid-state quantum sensing, or solar cells [14, 15, 88–91] and we have seen that impurities can also drastically reduce the performance of the device. For example, in solar cells they can cause unwanted doping compensation or have deep defects states that act as recombination centers, thus reducing the efficiency of the device [92]. To successfully design the next generation of devices, we need a clear understanding of the physics of promising new materials and their defect properties.

In the context of new materials for photovoltaic applications, the family of chalcopyrites has been growing in recent years as promising low- and wide-band gap absorber materials within solar cells [8, 11, 93]. Most recently, $CuInS_2$ and $CuGaS_2$ have received attention as its quaternary alloy Cu(In,Ga)S_2 allows for the adjustment of the band gap depending on the [In]/[Ga] content, making them suitable for application in tandem solar cells. To date, an efficiency of over 16.1% has been demonstrated for Cu(In,Ga)S_2 (CIGS)-based solar cells [94]. A fairly complete assignment of defect levels has been achieved for quaternary Cu(In,Ga)Se_2 compounds [8]; however, similar efforts for Cu(In,Ga)S_2 are still ongoing.

Photoluminescence measurements are an experimental method that is frequently used to investigate defects [6, 95]. In the quaternary compound Cu(In,Ga)S_2, these measurements reveal two optical transitions associated with deep defects around 1.35 eV (D1) and 1.1 eV (D2), respectively [11]. Photoluminescence measurements on $CuGaS_2$ also show two deep defect transitions at approximately 2.15 eV (DD1) and 1.85 eV

(DD2) [10, 13, 96]. Moreover, the DD1 position of the peak exhibits a dependence on the [Cu]/[Ga] ratio. So far, only Cu_{III} has been proposed as a candidate defect for the D2 peak [11], while it has been hypothesized that more then one defects is involved in the DD1 peak [10, 96]. Therefore, additional theoretical analysis is necessary to gain a better understanding of the defects in these materials.

State-of-the-art theoretical approaches mainly consist of density functional theory (DFT) calculations of formation energies and the resulting charge transition levels [6, 7, 46, 47, 50]. For these DFT calculations, the choice of the right exchange-correlation functional is important. Vidal et al. [97] highlight this through self-consistent GW calculations of the band structure of chalcopyrites, where they also show that hybrid functionals can give a satisfactory description of the electronic properties of pristine chalcopyrites. Consequently, most recent studies on the electronic properties of defects in these compounds are based on hybrid functionals, using medium-sized periodic supercells [11, 98–111].

Most of the defect-related studies on sulfur-based CIGS are considering either indium or gallium chalcopyrites. For $CuInS_2$ the most prominent intrinsic point defects have been investigated by Chen et al. [103], who also include the O_S substitutional as a possible defect in non-vacuum growth conditions. Other studies focus on defect complexes; for example, Yang et al. [109] studied intrinsic defect complexes, while Xiang et al. [110] investigated the impact of phosphorus on the doping type. The stability of the different intrinsic defects for different phases of $CuGaS_2$ was studied by Bailey et al. [98], where the authors demonstrated the native p-type character of the material. Sn-doped $CuGaS_2$ has been studied by Han et al. [111, 112] while Pohl et al. [99] investigated V_{Cu} for four chalcopyrites. Only Shukla et al. [11] considered both compounds and calculated the charge transition levels of a wide range of defects. They propose that the Cu_{In} and Cu_{Ga} antisites are potential defects for the peak D2 (1.1 eV) observed in indium-rich CIGS. However, to strengthen their claim, optimization of the HSE parameters (e.g., using generalized Koopmans' theorem) and a discussion of the limitations of charge transition levels as well as of their connection to optical transitions via lattice relaxation are missing.

The aim of this chapter is to provide a complete and accurate study of the defect

landscape in both $CuInS_2$ and $CuGaS_2$ by means of density functional theory within the Heyd, Scuseria and Ernzerhof (HSE) hybrid functional scheme. For both chalcopyrites, the standard parameter pair still drastically underestimates the band gap. As a result, in previous studies, the mixing parameter [108–110, 112] or the range parameter [99, 103] was adjusted to achieve a band gap close to the experimental value. Only Han et al. [111] adjust both parameters. In addition to the band gap, they used parameters that best describe the Kohn-Sham orbitals of Sn_{Ga} within the generalized Koopmans' theorem (gKT). However, we are interested in the broad defect landscape of both materials. Therefore, we find the optimal parameter pair for the pristine crystals by investigating the parabolicity of partial removal of an electron to align with the gKT. This is a similar approach to Deák et al. [113] who calculate the parabolicity for the neutral V_O in $\beta-Ga_2O_3$ to fulfill the gKT. To further enhance the accuracy delivered in our study, we increase the supercell size to 288 atoms to reduce the defect-defect interaction and to provide an updated insight on the charge transition levels of the most common defects. Compared to the previously used supercell of 144 atoms [11] which corresponds to $3\times3\times1$ conventional unit cells (of 16 atoms each), the use of a $3\times3\times2$ supercell with twice the number of atoms increases the minimum distance between defects by a factor of 1.48 which makes a clear difference for the resulting charge transition levels (CTL).

Moreover, we also deliver a more accurate evaluation of optical transitions associated with defects. This is achieved by considering optical transition levels (OTL) within the single-particle picture. They can provide us with a more detailed insight into the experimentally observed optical transitions, compared to what CTL alone can deliver. As expected intuitively, for deeper, more localized defects, the lattice contribution to their spectroscopic signature is not negligible. Calculations of the OTLs of the respective defect in their expected most stable charge state under p-type conditions reveal multiple candidates for the transitions observed in photoluminescence measurements. We identify V_{Cu}, V_{In} and Cu_i as potential defects for the D1 transition. The results for Cu_{In} antisite reveal an emission, which aligns well with the D2 peak observed in photoluminescence measurements. This supports previous suggestions that the Cu_{III} defect is a suitable candidate.

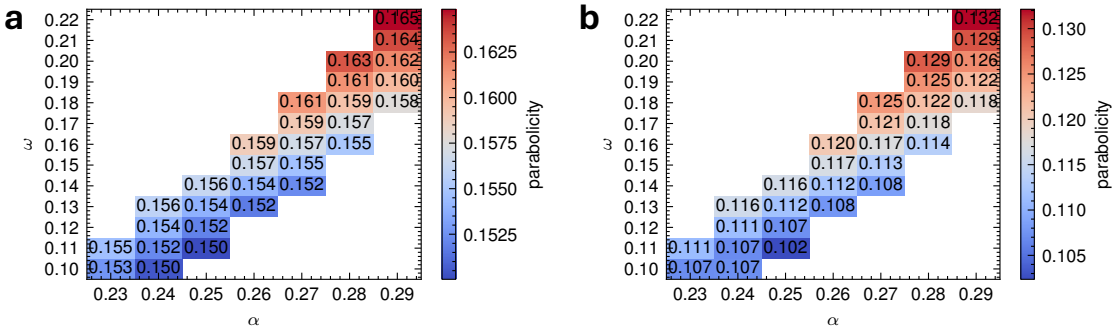

Figure 3.1: Parabolicity results for $CuInS_2$ and $CuGaS_2$ for parameter pairs that are in the range of previous band gap calculations.

Furthermore, our results indicate V_{Cu}, Cu_{Ga} and Cu_i as possible defects responsible for the DD1 transition in $CuGaS_2$. The closest candidate for the DD2 transition in our results is the V_{Ga}. Finally, we are able to converge Δ-SCF calculations for two deep defects, In_{Cu} and Ga_{Cu}. The calculations reveal In_{Cu} as an additional candidate for the D1 transition in indium-rich $Cu(In,Ga)S_2$.

The chapter is structured as follows. In Section 3.2 we present the study on the hybrid functional parameter. This is followed by the charge transition levels of the studied point defects and a discussion of the impact of the lattice relaxation via optical transition levels. Finally, we will show Δ-SCF results for some of the defects, before we compare these results with photoluminescence measurements.

3.2 Results

3.2.1 Hybrid functional parameterization

In Section 2.2.2, we have discussed hybrid functionals, such as the one introduced by Heyd, Scuseria and Ernzerhof [44]. They include to some degree the exact exchange-energy of the Hartree-Fock theory and they partially overcome the well-known problem of the band gap underestimation associated with local and semi-local exchange-correlation functionals. The HSE exchange-correlation energy is given by Equation 2.28 and includes a mixing

parameter, α, and a range parameter, ω, for the separation of the exchange-energy into a long- and a short-range part.

The standard HSE06 with a parameter pair of α = 0.25 and ω = 0.2 Å is well established for many materials. However, when used for chalcopyrites, the band gap is still underestimated. To study defects in these materials, it is therefore crucial to find a pair of HSE parameters that accurately describes the electronic structure of the system. Consequently, it is necessary to examine two properties of the system to identify the most suitable pair of parameters. Preselection is done for parameter pairs around the standard HSE06 (α: 0.2-0.3, ω: 0.1-0.3 Å), by comparing the band gaps with the experimental data. Here, a sparse Monkhorst-Pack grid was used to find the trends and the calculation the number of parameter pair candidates substantially.

The parameters that fall into that range are then further investigated to be compliant with the gKT as discussed in Chapter 2.1.2. The exact exchange-correlation functional exhibits a piecewise linear behavior, featuring derivative discontinuities at integer electron numbers. Therefore, to improve the performance of the hybrid functional, we identify parameter pairs that best approximate linearity.

Each defect has different properties, and ideally one should analyze the impact of the parameter pair on the electronic behavior of the respective defects individually. However, to make a comparison for a broad range of defects, we carried out the search for the best parameter pair for the pristine crystals. We calculate the total energies in five steps from N to N-1 electrons for pristine $CuInS_2$ and $CuGaS_2$ (Figure 3.1). We observe that the linearity is best described around the standard HSE mixing parameter $\alpha = 0.25$. Subsequent band

	this work	exp. data
$CuInS_2$	1.56	1.53[a], 1.55[b]
$CuGaS_2$	2.45	2.43[b], 2.53[a]

[a] Measurements carried out at 2° K [114]
[b] Measurements carried out at room temperature [115].

Table 3.1: Theoretical an experimental band gap energies in eV. The band gaps are calculated with parameters $\alpha = 0.25$ and $\omega = 0.12$.

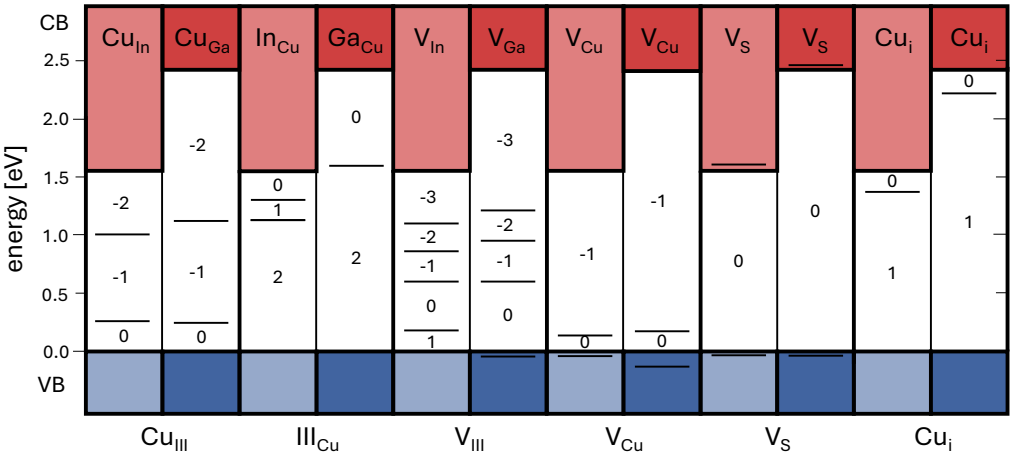

Figure 3.2: Charge transition levels of the studied intrinsic defects in $CuInS_2$ and $CuGaS_2$. Defects are grouped by types. Antisites, the group three vacancies and the copper vacancies show deep charge transition levels. V_S has no transition level within the band gap. For both compounds, the Cu_i has a shows a charge transition level close to the conduction band.

gap calculations were performed using the same Monkhorst-Pack grid density used for supercell calculations, and the best agreement with experimental data is obtained for $\omega = 0.12$. The corresponding band gaps for both compounds can be found in table 3.1.

3.2.2 Charge transition levels

In Section 2.2.2, we have introduced formation energies together with charge transition levels. They provide us with information about the types of defects or their most stable charge states, depending on the Fermi energy. Therefore, we have calculated the charge transition levels of the most prominent intrinsic point defects for both $CuInS_2$ and $CuGaS_2$. The results are shown in Figure 3.2 and in Table 3.2.

Antisite defects

Indium and gallium, belonging to the third group (III), have two more valence electrons than copper; this means that when the latter sits in a group-III site it should behave as an acceptor. Both Cu_{In} and Cu_{Ga} have two charge transition levels (CTL) from 0 to -2 at similar positions with respect to the valence band maximum, confirming the acceptor

$CuInS_2$			**$CuGaS_2$**		
Cu_{In}	(0/-1)	0.26	Cu_{Ga}	(0/-1)	0.25
	(-1/-2)	1.01		(-1/-2)	1.13
In_{Cu}	(+2/+1)	1.14	Ga_{Cu}	(+2/0)	1.61
	(+1/0)	1.31			
V_{In}	(+1/0)	0.18	V_{Ga}	(+1/0)	-0.04
	(0/-1)	0.60		(0/-1)	0.61
	(-1/-2)	0.86		(-1/-2)	0.96
	(-2/-3)	1.10		(-2/-3)	1.22
V_{Cu}	(+1/0)	-0.04	V_{Cu}	(+1/0)	-0.13
	(0/-1)	0.13		(0/-1)	0.18
V_S	(+1/0)	-0.04	V_S	(+1/0)	-0.03
	(0/-1)	1.61		(0/-1)	2.48
Cu_i	(+1/0)	1.37	Cu_i	(+1/0)	2.24

Table 3.2: Calculated charge transition levels for the considered intrinsic point defects for $CuInS_2$ and $CuGaS_2$. Energies are in eV.

behavior.

In the case of gallium and indium sitting on the copper site, our results show a transition from +2 to 0 when increasing the Fermi energy, which is characteristic for a donor. The In_{Cu} has two distinct CTLs, whereas the Ga_{Cu} has only one CTL from +2 to 0.

The comparison of our results with previous studies is not straightforward, as they were carried out with different supercell sizes and different HSE parameters. Despite this, the results identify similar CTLs, however, at different positions for some of the defects. A notable difference is in the results of Xiang et al. [110] who find Cu_{In} to be a shallow acceptor without any neutral charge state within the band gap.

Furthermore, for In_{Cu}, Chen et al. [103] find a relatively shallow 2/1 CTL, while Yang et al. [109] and Xiang et al. [110] do not observe any charge transition level with a stable In_{Cu}^{+2} charge state. In contrast, the other studies reveal a deeper position of the 2/1 CTL.

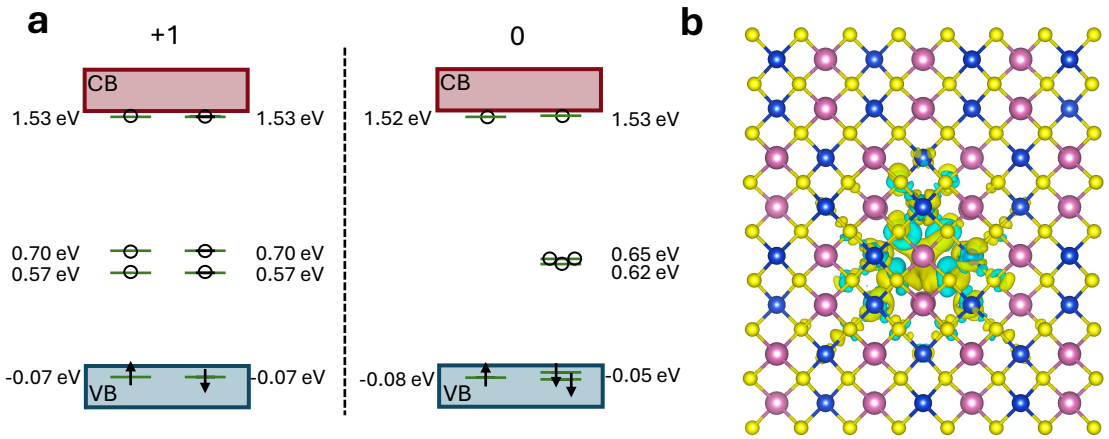

Figure 3.3: **a** Kohn-Sham eigenvalues of the V_{In} for +1 (left) and 0 (right) for their respective ground-state geometries. **b** Charge difference for +1 and 0 with the geometry of the neutral charge state.

Vacancy defects

Both group III (In and Ga) vacancies show three similar charge transition levels ranging from 0 to -3 at similar positions with respect to the valence band. Although this results in donor behavior for the V_{Ga}, V_{In} is an amphoteric defect with a CTL $\epsilon(1/0)$ at 0.18 eV. The V_{Cu} exhibits a single CTL (0/-1) which is also positioned close to the valence band for both compounds. In contrast to all the other defects investigated in this study, the sulfur vacancy does not exhibit any charge transition level within the band gap, and the neutral charge state is the most stable.

The most significant difference from previous studies is the amphoteric behavior of V_{In}, a phenomenon absent in the others [11, 103, 109]. However, the analysis of the charge difference for +1 and 0 shows a very localized behavior of the charge, indicating a stable charge state (Figure 3.3). The results of Pohl et al. [99] indicate a shallow V_{Cu} without any charge transition levels in both compounds, while the other studies that include V_{Cu} have a $\epsilon(0/-1)$ transition level relatively close to the valence band. Another contrast is the $\epsilon(+2/0)$ for the V_S deep within the band gap in the study by Chen et al. [103], a defect that has not been studied by the other articles.

Copper interstitial

To calculate the properties of the copper interstitial, we assume a behavior analogous to that previously observed in selenide-based chalcopyrites, where the atomic ground state is located at the center of the cation octahedron [116]. We find that the Cu_i defect exhibits a charge transition level $\epsilon(+1/0)$ at similar positions relative to the conduction band in both compounds. In contrast, Yang et al [109] and Xiang et al. [110] report that Cu_i acts as a shallow donor without exhibiting any distinct charge transition levels.

3.2.3 Optical transition levels

We have discussed the charge transition level for common defects in both $CuInS_2$ and $CuGa_2$. These charge transition levels are sometimes directly compared with photoluminescence measurements. However, a direct comparison of CTLs with PL peaks can be justified for shallow defects (small impact on the lattice) within host materials with weak electron-hole interaction.

We can reach higher accuracy by considering that optical (de-)excitation is much faster than lattice relaxation. The central idea of this so-called optical transition level (OTL) is that for optical processes the defect geometries do not change instantaneously [46, 47]. The key advantage of considering vertical transitions is therefore the possibility of capturing the response of the lattice via the Franck-Condon shift.

To apply this methodology for the two chalcopyrites, we need to consider the most stable charge states of the respective defects. $CuInS_2$ and $CuGaS_2$ are native p-type semiconductors with the Fermi level located in proximity to the valence band [98, 110]. This determines the most stable charge states of the defects and the starting point for subsequent investigations, including optical transition levels.

For transitions from the conduction band to the defects, empty Kohn-Sham states within the band gap are necessary. If we want to describe an emission process from the defect to the valence band, the defect needs to have an occupied state within the band gap after excitation.

$CuInS_2$				**$CuGaS_2$**			
acceptors	ZPL	OTL	dFC		ZPL	OTL	dFC
Cu_{In}^{em} $(0 \rightarrow -1)$	1.30	1.20	0.10	Cu_{Ga}^{em} $(0 \rightarrow -1)$	2.20	2.15	0.05
V_{In}^{em} $(+1 \rightarrow 0)$	1.38	1.35	0.03	V_{Ga}^{em} $(0 \rightarrow -1)$	1.84	1.70	0.14
V_{Cu}^{em} $(0 \rightarrow -1)$	1.43	1.38	0.05	V_{Cu}^{em} $(0 \rightarrow -1)$	2.27	2.23	0.04
donors	CTL	OTL	dFC		CTL	OTL	dFC
In_{Cu}^{abs} $(+2 \rightarrow +1)$	1.14	1.49	0.35	Ga_{Cu}^{abs} $(+2 \rightarrow +1)$	1.61	2.27	0.66
In_{Cu}^{em} $(+1 \rightarrow +2)$	1.14	0.65	0.49	Ga_{Cu}^{em} $(+1 \rightarrow +2)$	1.61	0.51	1.10
Cu_i^{abs} $(+1 \rightarrow 0)$	1.37	1.36	0.00	Cu_i^{abs} $(+1 \rightarrow 0)$	2.24	2.24	0.00
Cu_i^{em} $(0 \rightarrow +1)$	1.37	1.34	0.03	Cu_i^{em} $(0 \rightarrow +1)$	2.24	2.19	0.05

Table 3.3: Calculated optical transition level for the expected charge states of the studied defects under p-type conditions. Transition to the acceptor defects (charging them with an additional electron) are referenced to the CBM and compared to the ZPL. The transition for the donors are referenced to the VBM, as the defects have to be charged first (abs) to have an occupied Kohn-Sham state. After the charging they can transition back to the initial charge state (em). The Franck-Condon shift is denoted as d_{FC}. All energies are in eV.

The emission energies obtained from optical transition levels, along with the corresponding Franck-Condon shifts, are listed in Table 3.3. The first number in parentheses relates to the ground state of the initial charge state. The donor defects III_{Cu} antisite and the Cu_i interstitial do not have an empty Kohn-Sham state in the band gap and need to be charged first to show an emission line.

In order to estimate the predictive power of CTL when used to study the optical properties of defects, the correlation between the position of CTL and the corresponding d_{FC} is visualized in figure 3.4. As expected, the shifts are small for defects with a CTL close to the VBM. In particular, only CTL above 0.2 eV exhibits a significant d_{FC} of more than 50 meV.

Our results highlight that if we want to compare charge transition levels to optical measurements, the Franck-Condon shifts cannot be neglected. Considering the weak electron-hole interaction in the studied chalcopyrites [117], it is reasonable to compare the OTLs with optical measurements. Furthermore, for defects with CTLs below 0.2 eV a direct comparison with photoluminescence measurements seems to be sufficient.

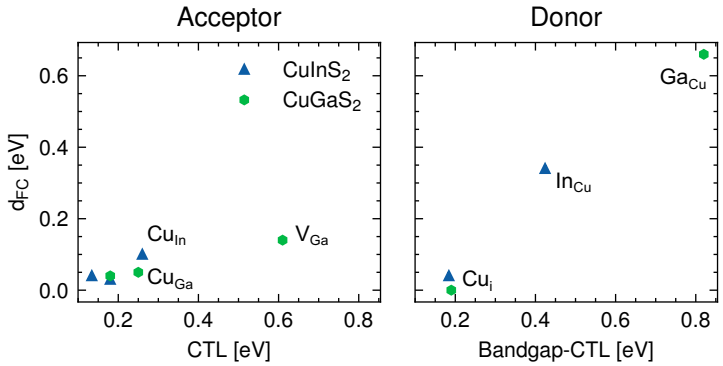

Figure 3.4: Franck-Condon shift dependence on the position of the CTL with respect to the corresponding band. Deeper CTL result in a bigger impact on the lattice in the charging process and therefore a larger shift. The investigated defects show a significant d_{FC} for CTL above approximately 0.2 eV.

3.2.4 Δ-SCF

Both previous approaches included a charging of the defect. However, there is a method in which we can conserve the number of electrons in the (de-)excitation process. This can be achieved with the Δ-SCF (self-consistent field) approach, sometimes referred to as constrained DFT (cDFT), which we have discussed in Section 2.2.2. Here, an excited state can be constructed in terms of single Slater determinants, where an electron is constrained into a specific unoccupied Kohn-Sham state. Similarly to the OTL, in this approach we can relax the system in the excited state and use a CCD to analyze and obtain the ZPL and the transition energies for absorption or emission for the defect.

The optical transition levels that we discussed require the same or even fewer computational resources than the charge transition levels. This means that they are always accessible when CTLs are. However, constrained DFT calculations include a more complicated electronic structure with an extra unoccupied state (hole) in the ground state density, which makes convergence more challenging. An additional difficulty encountered in chalcopyrites is the lack of clearly defined Kohn-Sham states for some of the defects, which in some cases become resonant states once occupied. Therefore, we have only been able to converge the calculations for the two antisite donor defects so far, which will be discussed below.

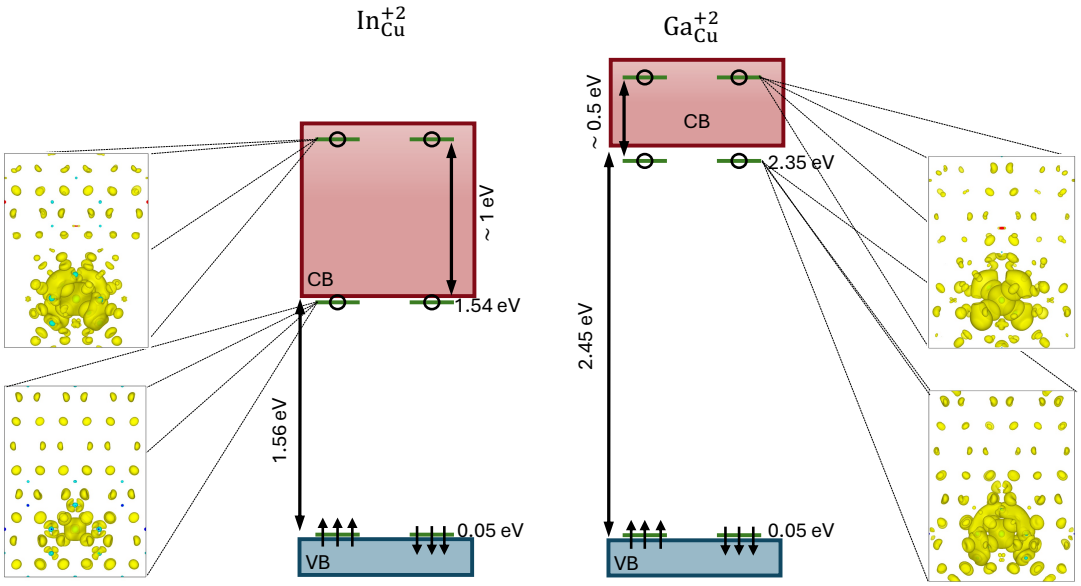

Figure 3.5: Electronic structure of the +2 charge state ground state of the deep donors In_{Cu} and Ga_{Cu}. Valence band maximum and conduction band minimum are referenced to the bulk. The defect state in the indium compound lies deeper in the continuum than for the gallium compound. Left and right side figures display the charge densities of the two lowest Kohn-Sham states. Both states show some localized features of the defect state. The atoms are removed for better visibility.

In_{Cu} and Ga_{Cu}

The electronic structure of both deep donors (In_{Cu} and Ga_{Cu}) is similar. In the p-type regime considered, the defects are two times positively charged, and we observe two Kohn-Sham states that show features of a localized defect state. The main difference between the two defects lies in the depth of the upper state within the conduction band.

Because we have a spin-degenerate excited state, we calculate its total energies similar to Mackoit-Sinkeviciene et al [23], who introduced an expression for the total energy in terms of single Slater determinants (see Section 2.2.2). We constrain the electron to the lower state and investigate if this leads to its localization at the defect site.

Following the configurational coordinate diagram, we observe a distinct behavior for both deep donors, In_{Cu} and Ga_{Cu}. Although the excited state prior to relaxation (absorption) exhibits a charge density similar to that of the lower states, the behavior of the relaxed states differs. For In_{Cu}, the geometry optimization does not further localize the electron

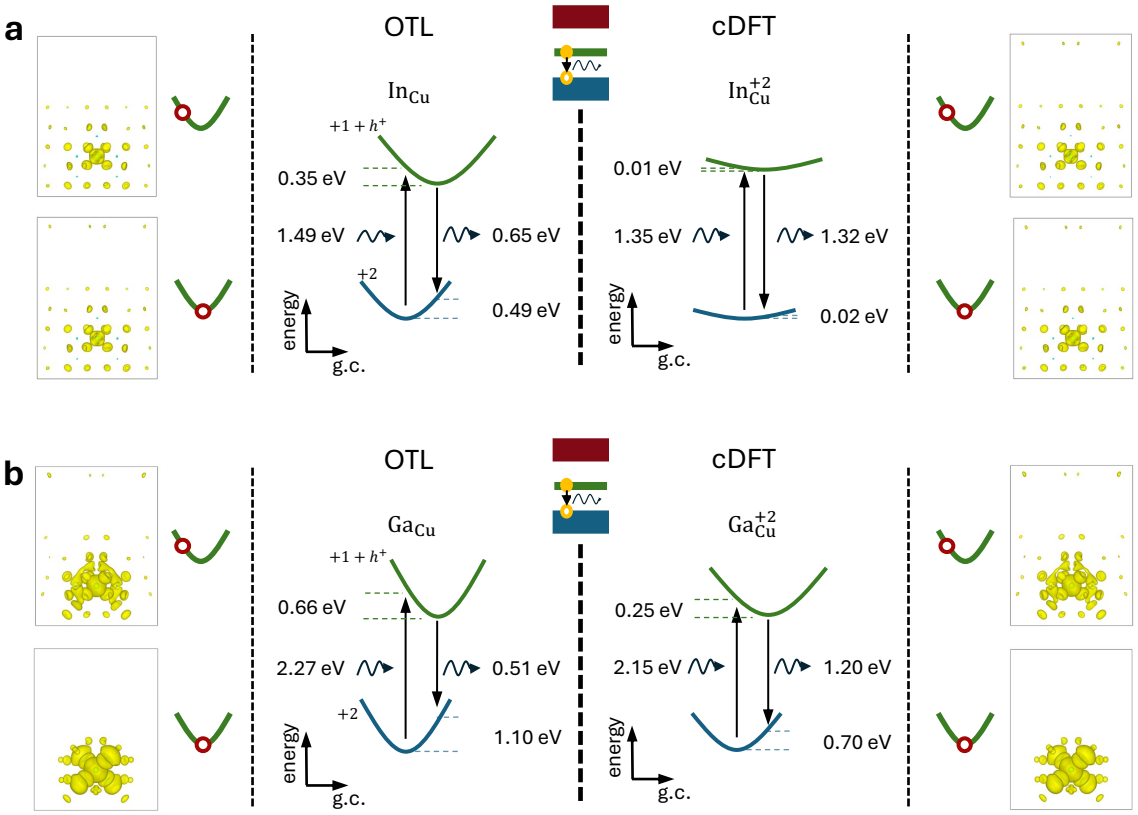

Figure 3.6: Optical transition level and constrained DFT results for In_{Cu}^{+2} **a** and Ga_{Cu}^{+2} **b**. Left and right site panels displays the respective charge densities of the relaxed and unrelaxed excited states. The atoms are removed for better visibility.

at the defect site, resulting in a negligible impact on the atomic configuration. In contrast, for Ga_{Cu}, relaxation leads to the electron localizing at the defect site. This is reflected in the Franck-Condon shift, which is observed only for Ga_{Cu}, indicating that the localized electron affects the surrounding geometry.

Our current explanation for the distinct behavior of the defects is the different depths of the upper state. Once occupied, the Kohn-Sham states of both defects are shifted downward. However, in the indium compound, the second state is deeper in the continuum, and the shift might not be enough to result in an isolated defect state in the band gap after absorption. Subsequently, this relaxation fails to further localize the electron. Nevertheless, we have no clear evidence to support this and a future study should try to identify the

reason.

If we compare the results to the optical transition level, we can clearly see a similarity of the charge densities. We observe the same behavior, where the Ga_{Cu} is localized, while the In_{Cu} appears to have features from the conduction band minimum. However, the emission lines for both defects differ, and we have no explanation at this stage.

3.2.5 Connection to experiments

Photoluminescence measurements performed on Cu(Ga,In)S_2 reveal two defect-induced transitions at around 1.35 (D1) and 1.1 eV (D2) [11]. The measurements were performed on samples with a [Ga]/[Ga+In] ratio ranging from 0.12 to 0.18 and a band gap of approximately 1.58 eV. The D2 transition has been attributed to the Cu_{III} antisites, as they exhibit a CTL at this energy and the photoluminescence peak becomes more prominent under copper-rich growth conditions.

The present study investigates intrinsic defects in the two limiting chemical compositions of quaternary chalcopyrite. The calculated charge transition levels for the indium-based chalcopyrite reveal five potential candidates for the observed spectroscopic signatures. In particular, the Cu_{In} exhibits a charge transition level ϵ^{CTL}(0/-1) 0.26 eV above the VBM. Our OTL calculations yield a transition from the conduction band at 1.20 eV, closely matching the experimentally observed transition D2 close to 1.1 eV. The second candidate is the In_{Cu} defect, which exhibits a charge transition level at 1.13 eV above the valence band maximum, suggesting that it could be responsible for the D2 transition. However, analysis of optical transition levels indicates a substantial Franck-Condon shift of more than 0.35 eV for both absorption and emission lines, which rule out this defect (Figure 3.7 **a**).

Although the optical transition level rules the In_{Cu} out for the D2 transition, when applying the Δ-SCF approach, we observe a transition in the range for the peak at 1.35 eV (D1). Other candidates for this transition (1.35 eV) include the V_{In} and V_{Cu} vacancies, as well as Cu_i, which exhibit optical transition levels close to the experimentally observed peak. However, since this peak is more prominent under copper-rich conditions, V_{Cu} is

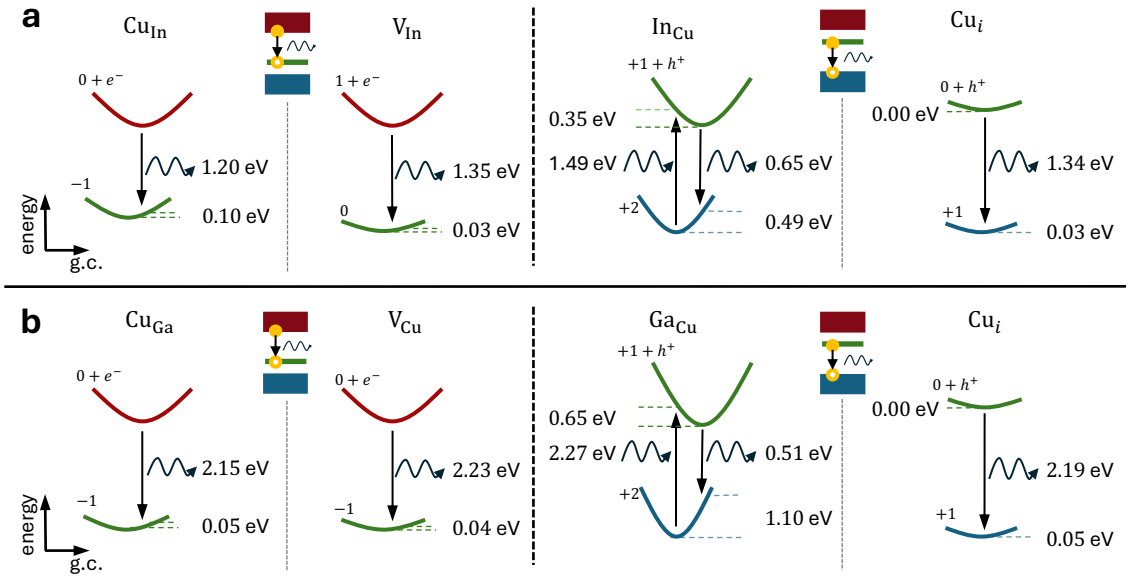

Figure 3.7: Sketch of the optical/vertical transition levels within a configuration coordinate diagram for some of the defects of interest. The color code of the parabolas describes the position of the electron and is consistent with the inset pictures (CB = red, defect = green, VB = blue). Left corresponds to transitions from the conduction band to the defect, while the right transitions are for transitions between defect and valence band, where the defect has to be charged prior to the emission. Defects in **a** are in $CuInS_2$ and **b** in $CuGaS_2$.

less likely, leaving V_{In} and Cu_i as more plausible candidates.

Studies on $CuGaS_2$ show two deep defect transitions around 2.12-2.24 eV (DD1) and 1.80-1.85 (DD2) [10, 13, 96]. Adeleye et al. [10] observe that DD1 dominates under Cu-deficient growth conditions. The same peak shows a shift to lower energies for increased copper content in the measurements by Botha et al. [96] and Eberhardt et al. [13]. Due to this shift and the broad feature of the peak, they concluded that more than one defect might be responsible for the transition.

We find that three defects exhibit an optical transition level that aligns with the measurements for the DD1 peak in $CuGaS_2$: V_{Cu}, Cu_{Ga} and Cu_i (Figure 3.7 **b**). Since the peak is more prominent under Cu-deficient conditions, it is more likely that a transition to the V_{Cu} is responsible for the peak. However, the experimentally observed shift of the DD1 peak to lower energies in Cu-rich environments might indicate that another defect is involved in the transition. Here, either Cu_{Ga} or Cu_i would be good candidates. Our findings

align well with the hypothesis that more than one defect is involved in this broad peak.

Finally, our results show only one possible candidate in the range of the DD2, namely the V_{Ga}. However, Adeleye et al. [10] mention that reducing [Cu]/[Ga] increases the intensity of the DD1, making it unlikely to be a V_{Ga}.

3.3 Discussion

In summary, we carried out a thorough search to find the best HSE parameter pair to fit the experimental band gap and to comply with the generalized Koopmans' theorem. We found a good agreement for both studied compounds for α = 0.25 and ω = 0.12 Å with respect to the experimental band gap, and investigated the pristine crystals to be compliant with the generalized Koopmans' theorem. This aligns with the results by Pohl [116] who varied ω to reproduce the experimental band gap for different chalcopyrites compounds ($CuInSe_2$ and $CuGaSe_2$). Han et al. [111] find a different parameter pair for $CuGaS_2$. However, they optimized the parameters for the theoretical band gap and find a parameter pair that best describes the Sn_{Ga}. Our approach allows for a comparison of a wide range of defects. Additionally, we increased the supercell size to 288 atoms compared to previous studies which used smaller cells with a maximum of 144 atoms [11].

We further calculated the formation energies and the corresponding charge transition levels for the most common intrinsic defects, namely Cu_{III}, III_{Cu}, V_{III}, V_{Cu}, V_S, and Cu_i, where III stands for group three elements: indium and gallium. We discuss the differences with previous studies, which is mainly an amphoteric behavior of the V_{In}.

The relevance of both compounds as a p-type absorber material for solar cells motivated the investigation of the optical properties. We investigate the difference between different single-particle methods used to obtain optical properties. Previous reports on these materials either use the single-particle Kohn-Sham states or the charge transition levels directly to explain the origin of the deep defect signals.

In the present work, we consider that optical transitions are vertical transitions within the configuration coordinate diagram, thereby accounting for the different timescales

associated with electron excitation and ionic relaxation. To the best of our knowledge, this approach has not yet been applied to chalcopyrites. We therefore shed new light on the optical properties of the defects, making them more directly comparable to photoluminescence measurements. In agreement with theory, our results reveal that for deeper charge transition levels, the contribution of ionic relaxation is significant, as evidenced by a Franck-Condon shift of several hundred meV. We also observe that for charge transition levels below 0.2 eV the shift is relatively small, making them suitable for initial comparisons with photoluminescence measurements.

Specifically, we observe that the Cu_{In} antisite in $CuInS_2$ exhibits an optical transition level in close proximity to the peak, labeled as D2, in photoluminescence measurements performed on indium-rich $Cu(In,Ga)S_2$. As a limiting case for the quaternary compound, this suggests Cu_{III} as a good candidate, supporting previous results. The second deep defect-related transition (D1) might be associated with V_{In}, V_{Cu} or Cu_i, as they exhibit an optical transition level comparable to the experimental data. Additional Δ-SCF calculation reveal also the In_{Cu} antisite as a candidate for the D1 transition. Furthermore, we identified strong candidates for a deep transition around 2.15 eV in $CuGaS_2$. The defect peak shifts to lower energies with increased copper content, which aligns well with the calculated optical transition levels of V_{Cu}, Cu_{Ga} and Cu_i. Our results could explain the hypothesis that multiple defects are involved in this peak.

Chapter 4

A machine learning approach to predict tight-binding parameters for point defects via the projected density of states

This chapter is partly based on the article 'A machine learning approach to predict tight-binding parameters for point defects via the projected density of states' accepted in npj Computational Materials [118].

In this chapter we discuss whether the projected density of states is sufficient to obtain tight-binding parameters for defective crystals. We will increase the complexity by considering substitutional defects in hBN, compared to the example of graphene introduced in the methods section. The sparseness in the PDOS and the very localized defect states in the band gap increase the difficulty of the problem. To describe the defects, we will adapt the perturbative model for the onsite energy and extent the model by investigating the impact

on the hopping parameters of the pristine crystal.

4.1 Introduction

The calculation of electronic properties with density functional theory using the super cell approach is well established [46, 47, 119–122]. In the previous section, we investigated two chalcopyrites, $CuInS_2$ and $CuGaS_2$, with DFT total energy calculations. The total energy gave us access to formation energies, charge transition levels and even optical transitions, and we were able to suggest multiple defects as possible candidates for peaks in photoluminescence measurements. However, it was not feasible to apply many body perturbation methods for chalcopyrites, although it has been used for simpler systems with localized defects [123–131]. Furthermore, studying the impact of defects on transport, mechanical, or excitonic properties of complex materials requires a large number of atoms. As discussed, this challenges pure *ab-initio* approaches and suggests the use of complementary semi-empirical methods, like tight-binding, to arrive at an accurate, yet affordable, atomistic modeling approach.

Recently, there has been a growing interest in applying machine learning techniques to the scalability of electronic structure calculations. The predictive power of neural networks can utilize information obtained by DFT on smaller systems to construct Hamiltonians for larger systems (e.g. by obeying the system symmetries) [132–134].

Another use of machine learning is in the context of semi-empirical tight-binding, where it has emerged as a powerful regression tool. For example, Wang et al. [135] used machine learning for a tight-binding parameterization where the neurons of the network act as matrix elements. Nakhaee et al. [136] fitted Slater-Koster parameters for BiTeCl by optimizing the parameters through machine learning. Soccodato et al. [137] and Gu et al. [138] obtain distance-dependent tight-binding (TB) parameters by applying local changes to the structures. However, all of the mentioned methods either require training data sets based on many DFT calculations, or they emphasize that a small cell is necessary to achieve a reasonable TB fit.

These efforts are complemented by the use of machine learning within the the self-consistent charge density-functional tight-binding method which aim to accurately perform geometry relaxations [139, 140]. The goal of such general tight-binding approaches is the ability to predict, for a given set of atomic positions, the corresponding tight-binding Hamiltonian and the resulting energy and forces.

For the description of point defects in their dilute limit, large super cells in real space are necessary. This results in a large density of intertwined bands in reciprocal space. Disentangling the bands (in the sense of making sure that the right tight-binding band is fitted to the right *ab-initio* band) becomes challenging. In recent years, quantitative parameterizations of the tight-binding model for defects, fitting to *ab-initio* calculations have been attempted. One way is a transformation from the (delocalized) DFT Bloch states to localized Wannier functions [141]. This approach remains challenging in the presence of many entangled bands even though some progress has been made by Fontana et al. [142]. Enforcing the localization of orbitals with the help of additional "spread functionals", they were able to obtain a Wannier fit of the electronic bands of the NV^- center in diamond in a 64 atom supercell.

One way to overcome the problem of band disentanglement in large supercells is by evaluating the Green's function $G(\mathbf{r}, \mathbf{r}'; E)$ of the perturbed system (i.e., the defective material). The imaginary part $\Im$ of its diagonal elements is proportional to the local density of states $\rho(\mathbf{r}; E)$:

$$-\frac{1}{\pi}\Im G(\mathbf{r}, \mathbf{r}; E) = \rho(\mathbf{r}; E) \ . \tag{4.1}$$

This enables the qualitative study of defective systems and has been widely applied for defects in graphene [29, 143–147]. In particular, nitrogen substitution in graphene has been studied by Lambin et al. [29] by introducing a dependence of the onsite energies on the distance from the defect. They optimize the tight-binding parameters with the local density of states.

In this chapter, we follow the idea of using the local (i.e., atom- and orbital-projected) density of states for the description of defects. In contrast to earlier work, we make it

quantitative by optimizing a small (but physically relevant) set of defect-related parameters through machine learning. We note that a fitting by machine learning of (total) electronic densities of states for various periodic systems (including defects) has been performed within the TBMaLT toolkit of DFTB [148, 149]. In our work, we introduce the fit to the atom- and orbital-projected density of states as the essential ingredient to obtain not only a reliable result for the (global and local) densities of states but also for the electronic band structure, a topic that was mentioned in Ref. [149] to be addressed in future studies.

Considering the balance between accuracy and transferability, we choose an approach different from training on a large set of DFT calculations. We rather create a large training data set directly by using our lean tight-binding model and varying its parameters in a certain range around the values for the pristine material. With the trained neural network, we then efficiently parameterize a defect using only *one single DFT calculation* for this defect.

A similar approach in this direction has been made by Schattauer et al. who used hopping parameters obtained from a pristine Wannierization as a starting point to describe different vacancy defects in graphene [30]. They have trained a machine learning model with band structures obtained by a distant-dependent Wannier parameters and showed the capability of the model to describe a vacancy in a transition metal dichalcogenide monolayer via the variation of pristine Slater-Koster parameters found in the literature. However, this approach is still based on the band structure, which becomes increasingly challenging to interpret for larger supercells due to the expected complex band folding (Figure 4.1). For larger cells, a model based on real-space instead of reciprocal-space data seems useful.

In this chapter, we describe a systematic way to obtain an intuitive, accurate, and transferable set of tight-binding parameters for the electronic structure of defects. We present a procedure to obtain such a model with the help of machine learning. We start with the tight-binding parameters of the pristine material. The defect (in this work we consider only substitutional atoms) is then treated as a perturbation of the pristine material and enters the TB model in four ways: (i) it has a different onsite energy and a

modified hopping parameter to its neighbors' nearest neighbors, mimicking the difference in chemical properties of the substitutional atom. (ii) In the vicinity of the defect, the shift of onsite energies is modeled by a Gaussian dependence on the distance from the defect. (iii) An extra hopping parameter between the nearest neighbor of the defect is introduced to capture the impact on the pristine parameters in the close vicinity. (iv) The defect may lead to a distortion of the surrounding lattice. This distortion is taken into account through the distance dependence of the tight-binding parameters of the pristine material.

Our approach keeps the number of fitting parameters for the defective cell to a minimum. As standard fitting tools fail to consistently predict a good parameter set, we employ a machine learning algorithm for this task. The machine learning model is trained with a set of projected densities of states calculated in the tight-binding approximation. Consequently, no expensive DFT calculations are involved in the training. Once trained, the models are given the projected density of states from DFT and provide the corresponding tight-binding parameters that properly reproduce both the projected density of states and the DFT band structures of supercells with defects.

To demonstrate our method, we study three cases of substitutional defects in monolayer hexagonal boron nitride (hBN): the carbon substitution of a boron atom (C_B), the carbon substitution of a nitrogen atom (C_N) and the carbon dimer substitution (C_BC_N). For single carbon substitutions, there are many-body calculations of the band structures and optical absorption levels of the defects [130, 131]. The carbon dimer is the main candidate for the 4.1 eV peak observed in the PL-spectra, as shown by constrained hybrid DFT calculations [23]. Semi-empirical models, even at a tight-binding level, have not yet been made available for these defects. Thus, in addition to serving as benchmark calculations for the development of our method, our results will enable the exploration of defects in hBN, which is a promising host for defect-related single-photon emission [19].

This chapter is structured as follows: We first discuss the problems of fitting the defect parameters in 4.2. This is followed by an explanation of the general workflow 4.3, in which we also discuss the architecture of our neural network. We will then introduce the tight-binding Hamiltonians, discuss the impact of defects, and apply descriptions of the

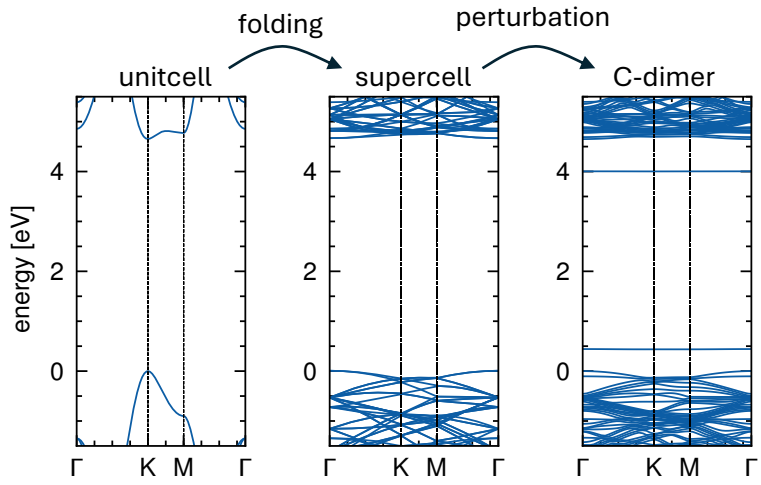

Figure 4.1: DFT band structures along high symmetry points of a hBN unit (2 atoms), a supercell (162 atoms) and a defective supercell (162 atoms) including a carbon dimer. Increasing the lattice leads to folding of the band structure. The defect perturbs the bands and introduces states within the band gap.

perturbation of the defect for hBN in 4.5. Finally, we will benchmark our results with carbon defects in hBN by means of the projected density of states (PDOS), the band structure, lattice deformation and distances between the defects in 4.6. We conclude the chapter with a discussion in 4.7.

4.2 The band structure problem

To properly describe a defect in an otherwise perfect crystal, using periodic boundary conditions, the defect is usually embedded within a sufficiently large supercell of the host material. The construction of a supercell leads to the back-folding of the bands of the primitive unit cell into the correspondingly reduced 1st Brillouin zone of the supercell (Figure 4.1). The introduction of a defect (here a carbon dimer substitution) then leads to a perturbation of the valence and conduction bands and, depending on the defect, to the appearance of one or more localized (i.e., weakly dispersing) states in the band gap. The description of such a system using a tight-binding Hamiltonian obtained by

jointly fitting the defect bands and the perturbed valence/conduction bands becomes a formidable task. Directly fitting the TB bands to the respective DFT bands (i.e., minimizing $|\epsilon_{nk}^{DFT} - \epsilon_{nk}^{TB}|^2$ for all n and k) requires their disentanglement: the respective eigenvalues have to be from the same band n at the same k. Tweaking tight-binding parameters to shift the "wrong" eigenstate to match a DFT energy will actually worsen the overall fit. The disentanglement of the bands becomes cumbersome quite quickly with increasing supercell size, and Wannierization reaches its limits already for defective 2D materials [30].

Using instead the PDOS offers several advantages. Most importantly, the key problem of the disentanglement of the bands is overcome by projecting the density of states onto the orbitals and atomic sites. Furthermore, the PDOS at the defect site and its close neighbors converges quickly as a function of supercell size. Even in the limit of large and complex defect supercells (where DFT calculations are restricted to a single k-point), the PDOS is still accessible and reliable.

The PDOS has significant advantages for the fitting process, but the perturbation of the host material by the defect also influences the tight-binding Hamiltonian. We will later see that finding such a tight-binding parameterization for defects remains difficult even when using the PDOS instead of the band structure. Before we discuss this in detail, let us first introduce the workflow, where we use a neural network to predict tight-binding parameter from the PDOS.

4.3 Workflow

The work flow is shown in Figure 4.2. The starting point are optimized geometries of the pristine cell and the defect supercell, calculated using density functional theory (DFT). The pristine tight-binding model is obtained by fitting to the *ab-initio* band structure in the primitive cell. A distance dependence of the parameters can be obtained by refitting to *ab-initio* calculations of strained unit cells. Once the pristine tight-binding parameters have been established, we use the defect geometry obtained from a DFT relaxation to build the

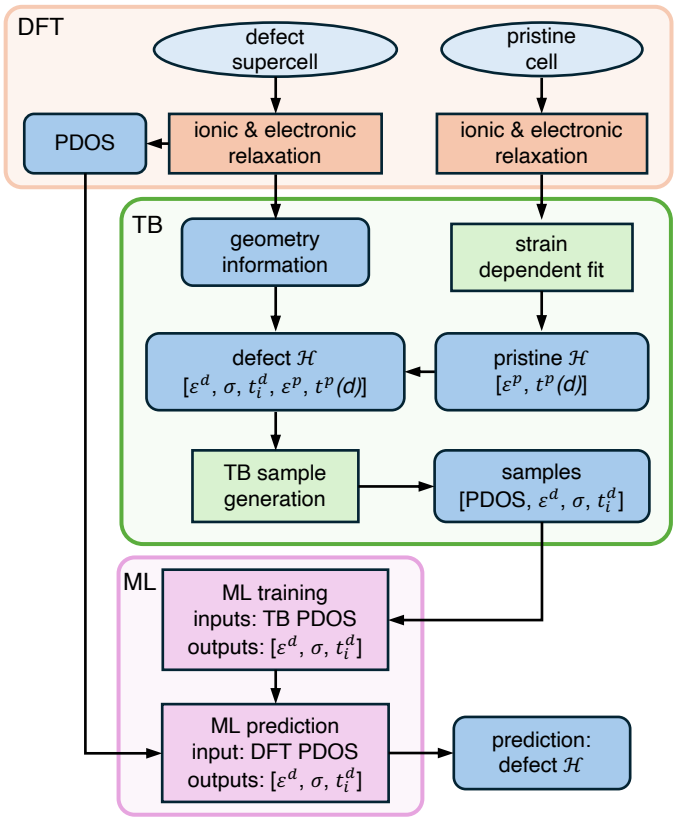

Figure 4.2: Workflow to obtain tight-binding parameters for defects. Oval nodes indicate the starting points, cornered nodes represent steps including calculations and rounded nodes represent data. DFT related calculations are displayed in orange, TB in green and ML in purple. Although some similarities are present to Schattauer at al. [30], instead of refitting the pristine parameters, we treat the defect as a perturbation to reduce the number of fitting parameters. Furthermore, we introduce an additional step for the generation of the pristine model, where we calculate a distance (strain) dependence of the TB parameters for the long-range description (see. 4.5.2). Last but not least, instead of using the band structure, we use the PDOS to avoid the cumbersome disentanglement of the bands.

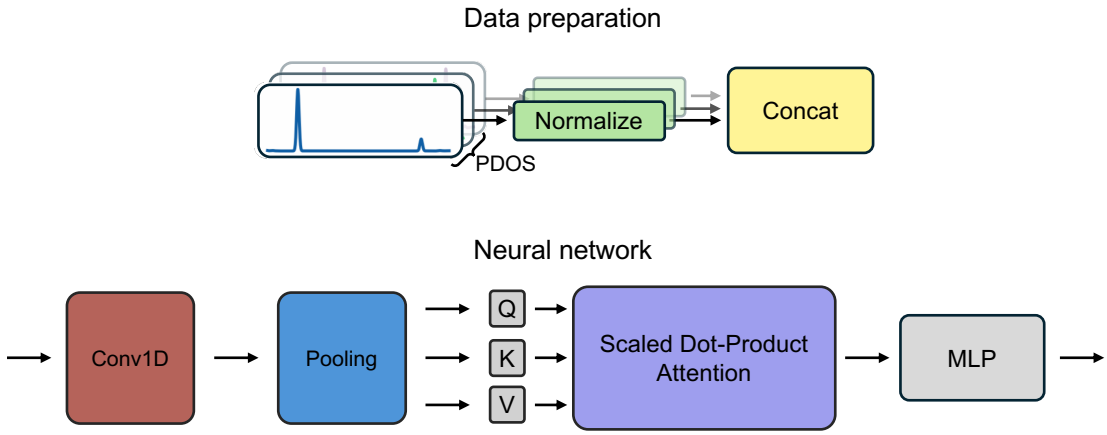

Figure 4.3: Upper part: Scheme of the data preparation. PDOS $D(\mathbf{r}_i; E)$ at different sites $\mathbf{r}_i$ is normalized separately. Lower part: Scheme of the neural network, displaying the order of the different parts of the network.

defect Hamiltonian as a perturbation of the strain-dependent pristine Hamiltonian.

Random sampling of the tight-binding parameters defining this Hamiltonian generates the data sets on which we train, and the neural network learns the mapping between the tight-binding PDOS (input) and the tight-binding parameters (output). A key benefit is that the training data is solely acquired within the tight-binding framework, making it computationally inexpensive. After the training, a DFT PDOS is used to predict tight-binding parameters that accurately describe the defect Hamiltonian. Before discussing the details of the tight-binding model, we will introduce the machine learning approach to obtain the parameters.

Machine learning architecture

The key task in our approach is to find tight-binding parameters that give rise to localized defect states within the band gap and simultaneously describe the bands properly. Using standard fitting tools to directly fit the tight-binding PDOS to DFT provides a first hint on the complexity of this task (Appendix C.3). Without prior knowledge of the defect parameters, standard fitting tools struggle to find the localized peak and simultaneously describe the bands properly. Another problem is the scalability of this procedure. While with a small

number of parameters, one can still manually fit the parameters, this becomes increasingly difficult the more parameters need to be fitted.

In recent years, state-of-the-art fitting tools have been complemented by machine-learning algorithms, enabling the exploration of a complex parameter space with many parameters. It is the goal of this chapter to make this connection between machine learning and fitting TB parameters via the PDOS: The neural network learns the mapping between the PDOS and the parameters used for the generation, minimizing the chance of being stuck at a local minimum and making sure that the defect state and the bands are properly described.

The PDOS $D(\mathbf{r}; E)$ as a fitting property has multiple advantages but it also inherits some difficulties. The considered defects are deep defects within hBN, a material with a large band gap. For such a material, $D(\mathbf{r}; E)$ is mostly zero within the band gap, with narrow defect peaks with big contributions at the defect site, resulting in few activations in the first layer.

To prepare the data for the training, we first simplify the problem by considering $D(\mathbf{r}; E)$ only at a small set of points $\mathbf{r}_i$ and only within a restricted energy range around the band gap to train only on the 1D PDOS data. Additionally, we include the sum $\sum_i D(\mathbf{r}_i; E)$ for the training since we get this additional information for free. We then normalize each $D_i(E) = D(\mathbf{r}_i; E)$ (between 0 and 1) separately as normalization of the data is essential for the training and it enables access of all features equally at every site (Figure 4.3).

The separately normalized $D_i(E)$ are then stacked before entering a 1D convolution (conv1D) layer. The aim is to capture the correlated nature of the PDOS, but also to compress the data and simultaneously reduce the fraction of zeros. Here, the stride defines how much the data is squished while the kernel size determines the broadening of the PDOS features. The output of the conv1D layer is pooled along the channel direction and averages the outputs for different kernels to maintain the features of the input.

The pooling of the conv1D is followed by a scaled dot-product attention layer as proposed by Vaswani et al. [85]. The aim is to include long-range dependencies between the defect states and bands, which we have discussed in detail in Section 2.3.1.

The final component of the network architecture is a multi-layer perceptron (MLP), which consists of fully connected linear layers. For each layer the ReLu function is used as an activation function. The architecture of the network is depicted in Figure 4.3. The final output is evaluated by the standard mean-square error (MSE):

$$\text{loss} = \frac{1}{\text{N}} \sum_{n} (y_n - \hat{y}_n)^2 \,, \tag{4.2}$$

where $\hat{y}_n$ is the generated and y_n the expected output of the network. We tested to include the Gaussian relation for ϵ^{defect} and σ for a given d, but we found no improvement for the final results.

This network is trained using data produced with the tight-binding approximation. The data set consists of PDOSs, serving as input, and the corresponding labels $[\epsilon^d, \sigma, t_i^d]$ as output. After training, the network can predict the tight-binding parameters from DFT PDOS obtained from a single calculation.

The network has been developed in parallel with the tight-binding model. Since such architectures solved the parameterization problem at hand, we did not increase its complexity further. For more complex materials and more parameters to be fitted, this can be adjusted. Possible adjustments include further separation of the training at different levels of the neural network. For example, a separation of the conv1D layer with padding in reflection mode enables more data from the bands. One can also use multi-head attention or include more scaled dot-product attention mechanism similar to a transformer neural network.

4.4 Pristine tight-binding Hamiltonian

In section 2.1.3 we discussed the tight-binding approximation and the impact of a defect in section 2.2.3. Let us revisit the pristine tight-binding Hamiltonian once more to understand how we want to treat the perturbation of a defect. The tight-binding Hamiltonian is an approximation of the full many-body Hamiltonian projected on the localized atomic orbital

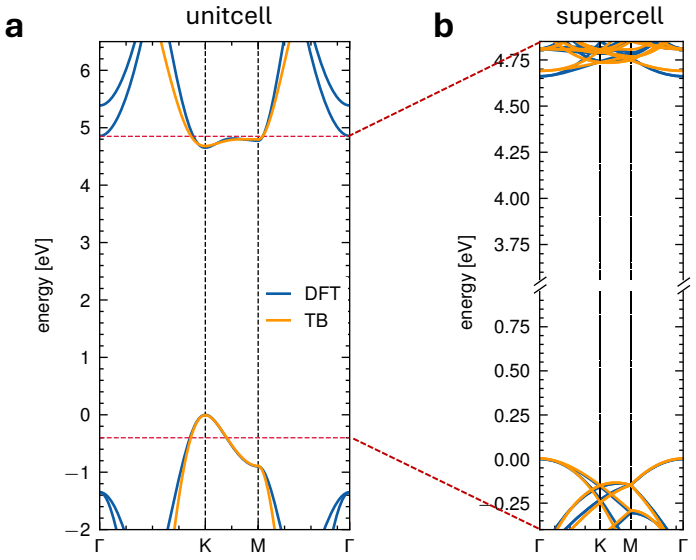

Figure 4.4: **a** DFT (blue) and the TB (orange) band structure of pristine hBN. The fit was done to best reproduce the DFT band structure along $\overline{MK}$. Red dotted line indicates range for later PDOS calculations. **b** Zoomed in band structures of the pristine supercell, displaying the difference observed for the conduction band minimum. The energy scale has been cut.

basis $|n\rangle$ located at site *n*. The Hamiltonian can be formulated as follows

$$\hat{\mathcal{H}}_0 = \sum_n \epsilon_n^{\text{prist}} |n\rangle\langle n| + \sum_{n,m} |n\rangle t_{nm} \langle m| \ , \tag{4.3}$$

where diagonal matrix entries $\langle n| \hat{\mathcal{H}}_0 |n\rangle$ are described with the onsite energies $\epsilon_n^{\text{prist}}$, whereas the off diagonal elements, $\langle n| \hat{\mathcal{H}}_0 |m\rangle$, are referred to as hopping matrix elements, t_{nm}, from site m to n. Due to the localized nature of the orbitals (fast decay), the hopping parameters can be restricted to the n-th nearest neighbor (usually $n \leqslant 3$), limiting at the same time the number of hopping parameters and the accuracy of the model. The Hamiltonian can be obtained by a fit of the parameters to experiments or *ab-initio* calculations.

In the case of a monolayer of hBN, we employ a pristine model including the p_z orbitals of both N and B with nonzero hopping parameters up to the 3^{rd} nearest neighbors. Furthermore, we use different 2^{nd} nearest neighbor parameters for N-N and B-B hoppings, $\text{t}_{2^{nd}}^{\text{NN}}$ and $\text{t}_{2^{nd}}^{\text{BB}}$, in order to reproduce the asymmetry in the highest valence and lowest conduction

	ϵ_N	ϵ_B	$t^{NB}_{1^{st}}$	$t^{NN}_{2^{nd}}$	$t^{BB}_{2^{nd}}$	$t^{NB}_{3^{rd}}$
	0.0	2.37	-3.12	0.002	-0.77	-0.35
α			-2.66	0.18	0.05	-0.19
β			5.86	-0.08	0.97	0.55

Table 4.1: The table contains the pristine tight-binding parameter, including the two onsite energies and four hopping parameters. The second nearest neighbor hopping is regarded separately for nitrogen and boron. The parameters α and β refer to the quadratic fit for the distant dependent hopping parameters.

band. Together with the two onsite energies, a total of 6 tight-binding parameters have been obtained and are shown in table 4.1. To best describe the valence band maximum and the conduction band minimum, these have been fitted to an *ab-initio* band structure along $\overline{KM}$ (Figure 4.4).

4.5 Defect tight-binding model

We have discussed multiple difficulties for a defective tight-binding Hamiltonian in detail in section 2.2.3. The use of a supercell increases the size of the matrix, which is not a problem in the pristine case. Here, the matrix is represented by the same parameters as for the pristine unit cell as discussed in section 2.1.3. However, in the more general case of a defective crystal, it is naturally expected that the perturbation introduced by the defect significantly influences the physics of the system, including changes in the positions of the neighboring atoms. Both affect the parameters of the defective supercell. In principle, all site energies and the parameters for all possible hoppings between nearest neighbors within the supercell (that are included in the model) would need to be refitted. In the case of a defective 162 atom hBN supercell, we would be facing more than 1000 parameters (hopping neighbor $n \leqslant 3$). Although this number can be reduced by symmetry [30], in the present work we aim for a minimal parameter set.

For the example of graphene, we obtained reasonably good results simply by fitting a distance dependence for the onsite energies without discussing the impact on the hopping

terms. For hBN, our aim is to include a description of this impact.

A substitutional defect introduces different chemical properties that we have to account for by new parameters. A new onsite parameter ϵ^{defect} is complemented by three hopping parameter to its nearest neighbors similar to the pristine model. This is sufficient because all respective nearest neighbors have the same distance to the defect as there is no Jahn-Teller distortion present for the investigated defects. Although this introduction of four additional TB parameters for the defect site is straightforward, it is not sufficient because the defect also has a subtle and delicate influence on the chemical properties of its surroundings that are not captured by these parameters.

Instead of re-fitting all hoppings and site energies in the defect supercell, we limit the number of additional fitting parameters to a minimum. The onsite energies of each atom are shifted as a function of its distance from the defect site, similar to that of graphene. Additionally, the hopping between pristine neighbors may vary due to geometric deformation. This is taken into account by making the corresponding hopping parameters dependent on the hopping distance.

4.5.1 Distance dependence of the onsite parameters

For the onsite energies, we employ a dependence on the distance from the defect. We use a Gaussian function to describe the perturbation potential introduced by the defect. The function is defined as follows:

$$\epsilon(\epsilon^{\text{defect}}, \sigma, d) = \epsilon^{\text{prist}} + \left[\epsilon^{\text{defect}} - \epsilon^{\text{prist}}\right] \cdot e^{-\frac{d^2}{2\sigma^2}} \ , \tag{4.4}$$

where d is the distance to the defect site and σ is the variance, which is directly proportional to the width of the potential. The dependency is modeled such that the height of the potential is represented by the difference between the onsite energy of the defect (ϵ^{defect}) and the host atom (ϵ^{prist}). This is a necessary adaptation from the previously distance dependence for graphene, since we have different pristine onsite energies and want the height of the potential to be linked to the onsite energy of the defect (ϵ^{defect}). This

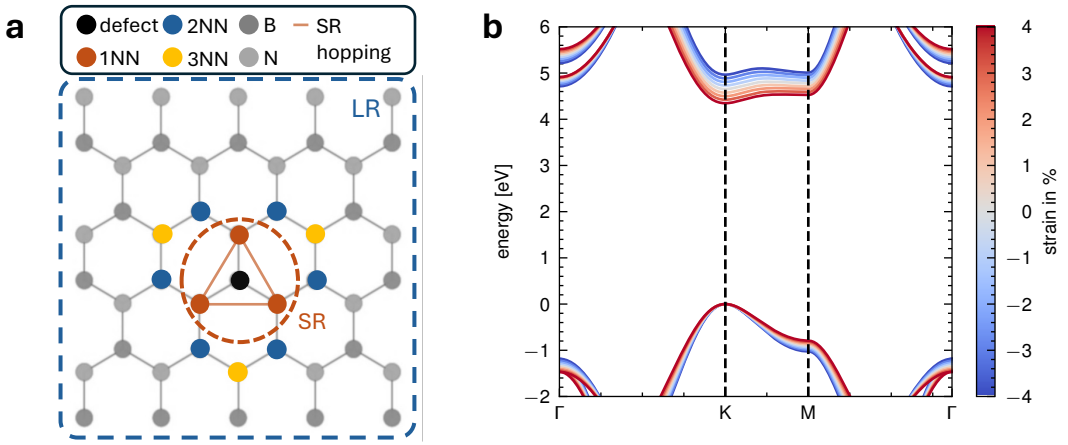

Figure 4.5: **a** Defect and its nearest neighbors and the adaptations used for the pristine hopping parameters. The short-range (SR) impact on the hopping parameters is described with an additional parameter between the nearest neighbors (1NN) of the defect. The long-range impact effects all other hopping parameters via a distance dependence between the hopping neighbors. **b** Strain dependent DFT band structures for hBN. Band structures are aligned at the valence band maximum and the k-axis is aligned for better comparability.

results in a total 5 parameters for the defect itself including its perturbation on the onsite parameters (height and width of the Gaussian function) and the three hopping matrix elements. Different adaptations for the pristine hopping terms t_{nm} will be discussed in the following.

4.5.2 Distance dependence of the hopping parameters

For the defect's influence on the hopping parameters, we distinguish between short-range and long-range effects. The first model contribution accounts for the local change of the chemical environment introduced by the defect. This is done by means of an additional hopping parameter between its nearest neighbors and is referred to as the short-range (SR) model (Figure 4.5 **a**). Because no Jahn-Teller distortion is present for the studied monomers, only one extra parameter needs to be fitted for the defective supercell.

The second contribution to the model accounts for the long-range (LR) component of the perturbation, and it is an expansion of the pristine tight-binding model by a distance dependence between the hopping neighbors. With it we aim at capturing the perturbation

on the geometry of the host crystal (Figure 4.5 **a**). The idea originates from the fact that tight-binding is well suited to model deformation even on large scales. Similarly to the perturbation of the defect, the deformation affects the distances between the neighboring atoms. The biggest impact is therefore expected to be on the hopping matrix elements. W. A. Harrison [32] introduced a scaling $t_0(d) = t_0 \left(\frac{d_0}{d}\right)^2$, which has proven to be a crude approximation [150–152]. Improvement can be gained by refitting the exponential to the respective hopping matrix elements.

The four hopping parameters were therefore fitted to strain-dependent pristine band structures, keeping the onsite energies fixed for later adjustments. For the defects studied and at the DFT level, a maximum change of the distances of the host atoms of 3.7% is observed for the C_N and 3.5% for the C_B. A biaxial strain within a range of 4% is therefore sufficient to capture the distance dependence for which the band structure was still well described.

We fit a distance dependence by obtaining strain-dependent band structures (Figure 4.5 **b**) for which we only refit the hopping parameters. In addition to the power function, we test three models to properly describe their dependencies

$$\begin{aligned} t(d) &= t \left(\frac{d_0}{d}\right)^{\alpha} \\ t(d) &= \alpha \cdot (d - d_{prist}) + t_{prist} \\ t(d) &= \alpha \cdot (d - d_{prist})^2 + \beta \cdot (d - d_{prist}) + t_{prist} \\ t(d) &= t_{prist} \cdot e^{\alpha(d - d_{prist})} \ . \end{aligned} \tag{4.5}$$

They include a linear, a quadratic, and an exponential description, constructed to reproduce the pristine parameters for the pristine distances between atoms. The results for the hopping parameters along with the distance dependence can be found in Figure 4.6. All proposed dependencies show good results except for t_{2nn}^{NN}. Here, we observe a change of sign that makes it impossible for the exponential functions to properly capture it. Due to the constrained condition that the pristine parameter is reproducible, the linear function also variates for higher strains. Although the absolute values are low, the additional

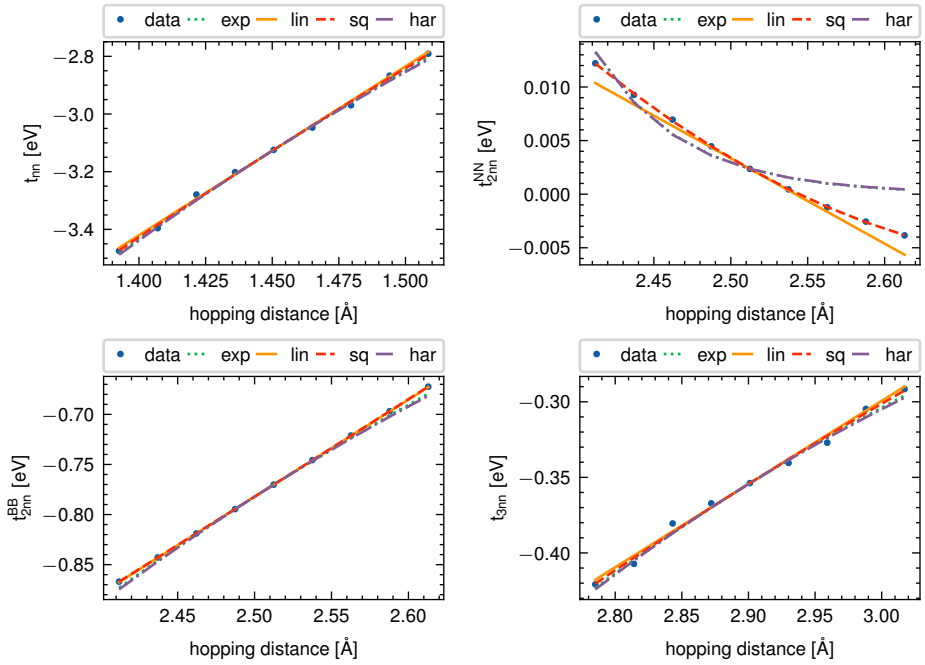

Figure 4.6: Distance dependence of the hopping parameters obtained from strain dependent calculations and the corresponding fits of the exponential (exp), linear (lin), quadratic (sq), and Harrison (har) for different strain. All show similar results, however, the change in sign can not be described with the exponential nor the Harrison model.

computational cost of using two parameters is negligible, and the quadratic fit is used to describe the impact on the geometry (LR).

4.6 Benchmark

To benchmark our methods, we investigate the carbon monomers and the carbon dimer. Both show deep defect states, respectively, and are perfect candidates to further testing of the workflow. The defect states within the band gap are also p_z orbitals and therefore match the p_z model for the pristine crystal. In the following, we will benchmark the workflow and method. The number of projected density of states (PDOS) on which the model to be trained can be easily adapted. Therefore, we will first investigate the impact of the number of PDOS included in the training.

4.6.1 Number of projected density of states as input

To understand the impact, data sets with up to four PDOS have been generated and trained on. To investigate the importance of the two hopping models, we have done this for all possible combinations of the hopping contributions (SR+LR, SR, LR, w/o). The results have been investigated by an evaluation of 10 trainings, respectively.

We evaluated the trainings by how frequent a peak is predicted correctly and if it is how well the bands are described. The first can be understood by checking how often the peak was predicted at the right energy (number of hits). For this we include results where the peak is predicted within one energy step compared to the DFT PDOS, since the discretization of the energy does not always show the maximum at the actual energy. The latter is understood by comparing the cosine similarities of different parts of the PDOS for those results for which the peak was predicted properly. The separation is necessary to see differences in the bands, as the defect peak outweighs these. The similarities were calculated up to the sixth nearest neighbor and include the defect site and three atom sites of nitrogen and boron, respectively.

For C_N, the results in Figure 4.7 indicate that the bands are sufficiently described when two PDOS are used as input for the training. However, we find that the order of the neighbor contributions for the defect peak is only guaranteed when we include three PDOS (defect site, first and second nearest neighbor). Three PDOS is also the optimal

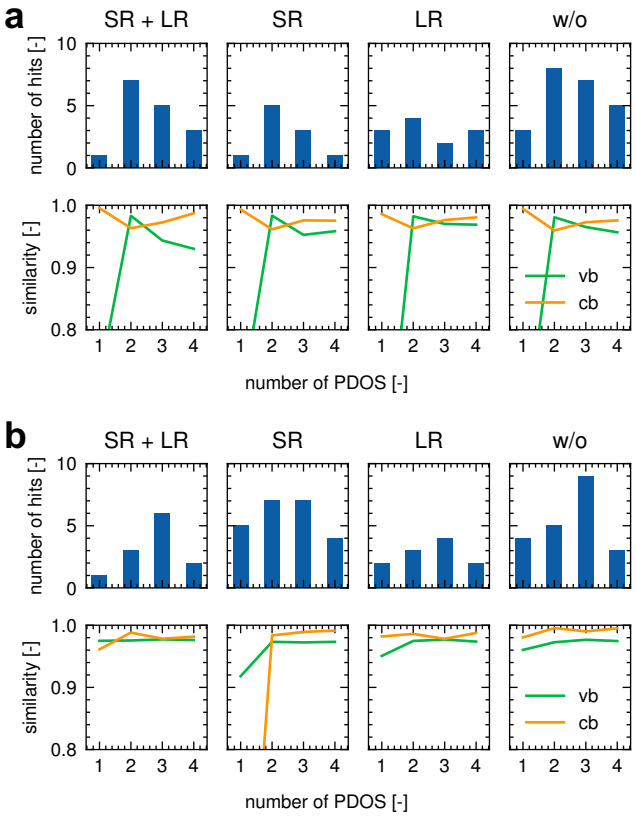

Figure 4.7: Number of hits (upper plots) and cosine similarities (lower plots) for **a** C_N and **b** C_B for different number of PDOS used as input for the training. For all results six PDOS are used for comparison with the *ab-initio* results.

input number number for C_B, for which the bands are better described for all hopping models. We therefore use three PDOS for both defects to further benchmark the method.

This will be done by comparison of the DFT and predicted tight-binding results. Since the parameters are generated by the DFT PDOS as the input for the neural network, we check how well the resulting TB PDOS aligns with the corresponding DFT PDOS. We will investigate, in particular, the importance of including SR and/or LR contributions to the TB model. Additionally, we want to understand whether the PDOS is sufficient to also obtain good agreement between the respective band structures. The model is further assessed by considering the strain and different defect distances.

4.6.2 Hopping models

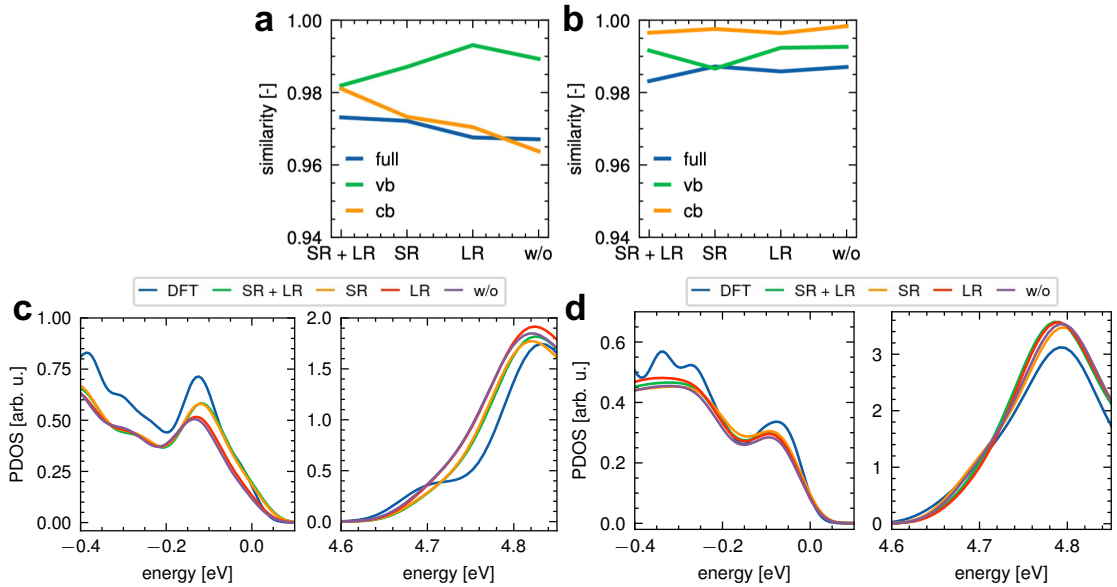

Figure 4.8: Cosine similarity of the different hopping description, including the complete model (SR+LR), either just the short-range (SR) or the long-range (LR) model, and without any adaptation (w/o). The cosine similarity is calculated for the valence (vb), conduction band (cb), and the full PDOS C_N in **a** and C_B in **b**. Valence and conduction band PDOS of the sum of the defect and its first 6 nearest neighbors of the C_N in **c** and C_B in **d** for the different hopping approximations.

For both carbon monomers, we want to answer the question of how important the adjustments of the pristine hopping parameters are. To do so, we analyze the individual contributions of all components of our model. Therefore, we use the complete model (SR + LR), along with models obtained by inclusion of either short- or long-range (SR, LR) model or the exclusion of both (w/o). Based on the previous results, for how many PDOS are optimal as input for the network, we can generate new data sets with a denser sampling of the PDOS to obtain the best parameters (Table 4.4).

These parameters are then used for the evaluation of the hopping contributions. Again, we compare the sum of the PDOS values and their cosine similarities up to the 6th nearest neighbor (defect site and 3 atom sites of each element). Since the defect peak outweighs the differences in the bands, we additionally separate the evaluation into different parts of the PDOS, namely the valence band (-0.4 eV to 0.1 eV) and the conduction band (4.6

C_N	ϵ^{defect}	σ	$t^{CB}_{1^{st}}$	$t^{CN}_{2^{nd}}$	$t^{CB}_{3^{rd}}$	t^{BB}_{SR}
SR+LR	2.611	0.110	-2.772	-0.044	-0.415	-0.830
SR	2.813	0.106	-2.835	-0.015	-0.471	-0.904
LR	3.208	0.094	-2.932	0.028	-0.400	
w/o	3.092	0.088	-2.917	-0.031	-0.380	
C_B	ϵ^{defect}	σ	$t^{CN}_{1^{st}}$	$t^{CB}_{2^{nd}}$	$t^{CN}_{3^{rd}}$	t^{NN}_{SR}
SR+LR	0.142	0.091	-3.164	-0.758	-0.735	-0.090
SR	-0.268	0.092	-3.414	-0.801	-0.698	-0.323
LR	0.053	0.086	-3.106	-0.830	-0.837	
w/o	0.178	0.092	-3.102	-0.761	-0.683	

Table 4.2: The table contains the predicted parameters for the respective hopping models used for the comparison of the monomers and the parameters for the carbon dimer. Superscript of the hopping terms denote the elements of the hopping neighbours. The hopping terms for the dimer are the hopping terms of the respective w/o model. The energies are in eV.

eV to 4.85 eV). The cosine similarities of the respective contributions are shown in Figure 4.8 **a** and **b** and they reveal no major differences with high similarities of above 0.96. For the C_N, the valence band is described slightly worse for models that include the additional hopping parameter (SR contribution). The opposite is observed for the conduction band, where the additional parameter improves the cosine similarity. For all contributions, the cosine similarity of the full PDOS does not change considerably. The C_B on the other hand is described equally well for all different hopping models.

The good agreement of similarities is also reflected in the direct comparison of the summed PDOS of the bands (Figure 4.8 **c** and **d**). Again, we observe a slightly worse description for the valence band when including the SR contribution for the C_N, which also results in a shift of the valence band maximum. In general, we find that the differences for all contributions are negligible.

This indicates that the perturbation of both monomers to the host material is sufficiently accounted for without the additional parameter. Therefore, considering new defect tight-binding parameters and its impact on the pristine onsite energies is enough to describe the carbon monomers. For modeling deformation, on the other hand, one should include

the distant dependence between the hopping neighbors (LR model). Both models include the same number of fitting parameters for the super cell. This enables a tight-binding model with few fitting parameters, but also for more complex defects like the carbon dimer.

4.6.3 Projected density of states

The PDOS of the entire range, including the respective contribution of each site, provides additional insight into the descriptive power of the predicted tight-binding parameters (Figure 4.9). We compare the features of the PDOS for all different hopping contributions. This is done by plotting the tight-binding PDOS, which are calculated with the tight-binding parameters predicted from the DFT PDOS.

Both of the carbon monomers show very good agreement between their respective tight-binding and DFT PDOS. In particular, for C_N, the second nearest neighbor has a larger contribution to the defect peak than the first, whereas the opposite is observed for C_B. Both are properly described with the respective tight-binding Hamiltonians. For the C_N the additional hopping parameter within the SR model increases the split between the first and second nearest neighbor at the defect site. Although this is closer to the DFT results, the difference compared to the models without the additional parameter is small (less than 0.1 eV) and we conclude it is sufficient to use five parameters for the respective defects.

In summary, our results show that the machine learning algorithm is capable of predicting tight-binding models for carbon defects in hBN via the PDOS. Although the tight-binding descriptions are relatively simple, they reproduce the DFT PDOS fairly well. One remaining question is whether the PDOS as a fitting observable is enough to also reproduce the electron band structure of the defective super cell.

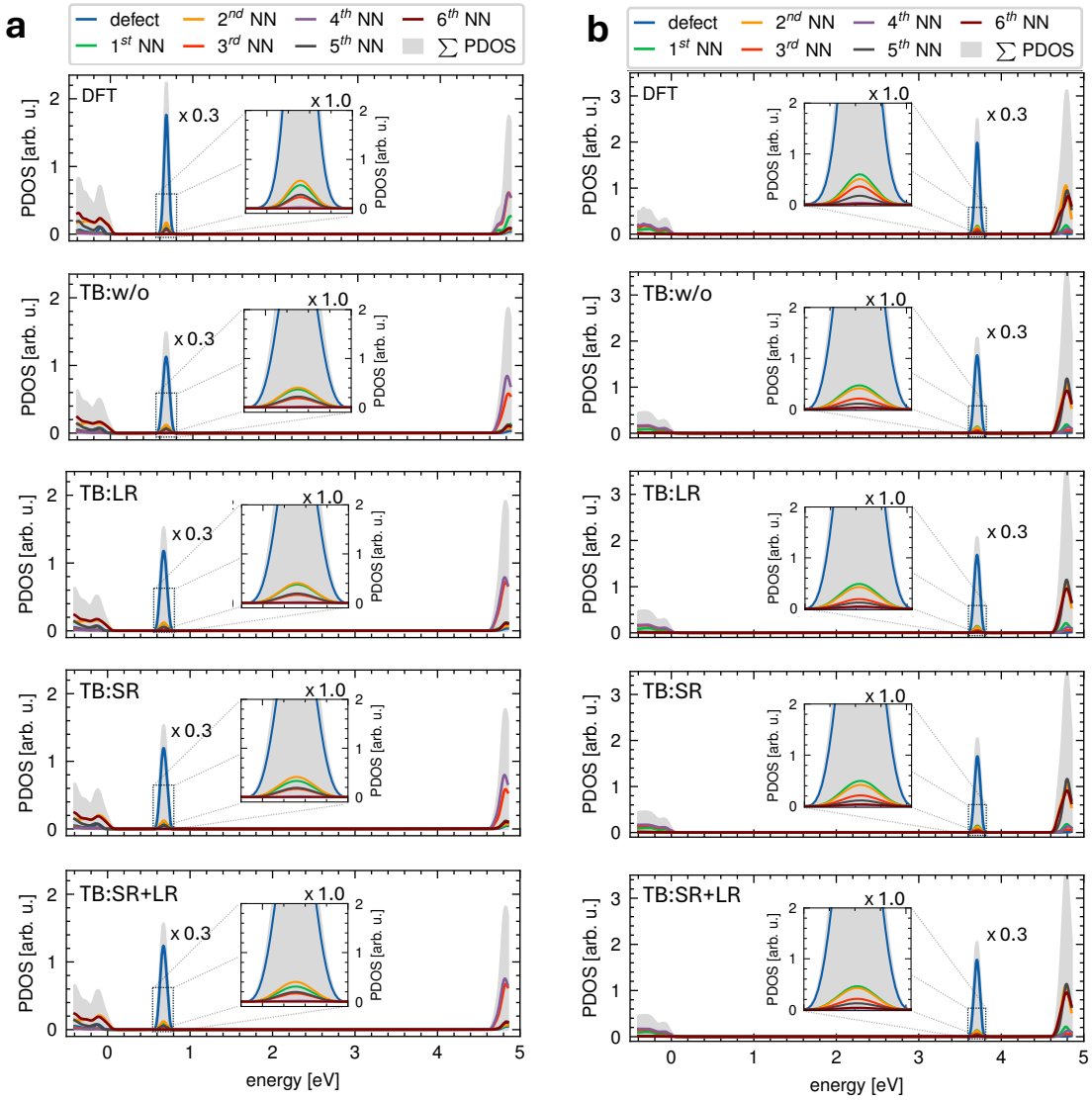

Figure 4.9: Projected densities of states obtained from DFT (upper subplots) and the corresponding TB results (lower subplots) for all of the respective hopping contributions. In **a** these are depicted for C_N and in **b** for C_B. The PDOS include sites for the defect and its first 6 nearest neighbors and the sum of all. The defect peaks are scaled by 0.3 for better visibility in comparison to the band edge contributions to the PDOS. The inset shows the defect peak without rescaling to assess the different contributions of the neighbors.

4.6.4 Band structure

We compare the band structures of the parameters for the different hopping contributions in Figure 4.10. As expected, for both defects and all models, the narrow defect peaks observed in the PDOS result in a correct description of the undispersive defect states in the band gap. The introduction of the defect results in a splitting of the bands from which the defect states emerge. This feature is captured by all tight-binding Hamiltonians as well. However, for the C_N not all hopping adaptations lead to an equally good description of the band structure. Although both w/o and LR contributions properly capture this splitting of the valence band, the additional parameter for the short-range contribution lifts the valence band maximum. This aligns with the previous assessment of the cosine similarity calculations and the sum of PDOS in Figure 4.8. The defect peak emerges from a valence band state and therefore we do not observe an impact on the conduction band which is similar to the pristine crystal (Figure 4.4).

For the C_B the defect state emerges from the conduction band. Aligning with the observation for the C_N, for all different hopping contributions we observe a splitting of the conduction bands. However, this splitting is underestimated, which is enhanced by the difference of the conduction band minimum already observed in the pristine fit. This difference is related to the simple pristine model with six parameters. It is not able to capture simultaneously the shape of the conduction band and the position of its minimum. More hopping parameters might fix this, but the aim of this work is to use a pristine model with its limitations to describe a defect. An improvement of the pristine model should also result in a better description of the defective model.

Overall, we have demonstrated that the use of the PDOS to obtain tight-binding parameters also shows good agreement for the band structures, thus enabling a novel method to obtain a tight-binding model for defective systems. The observed differences are related to the simplicity of tight-binding models (pristine and defective), rather than to the fitting method.

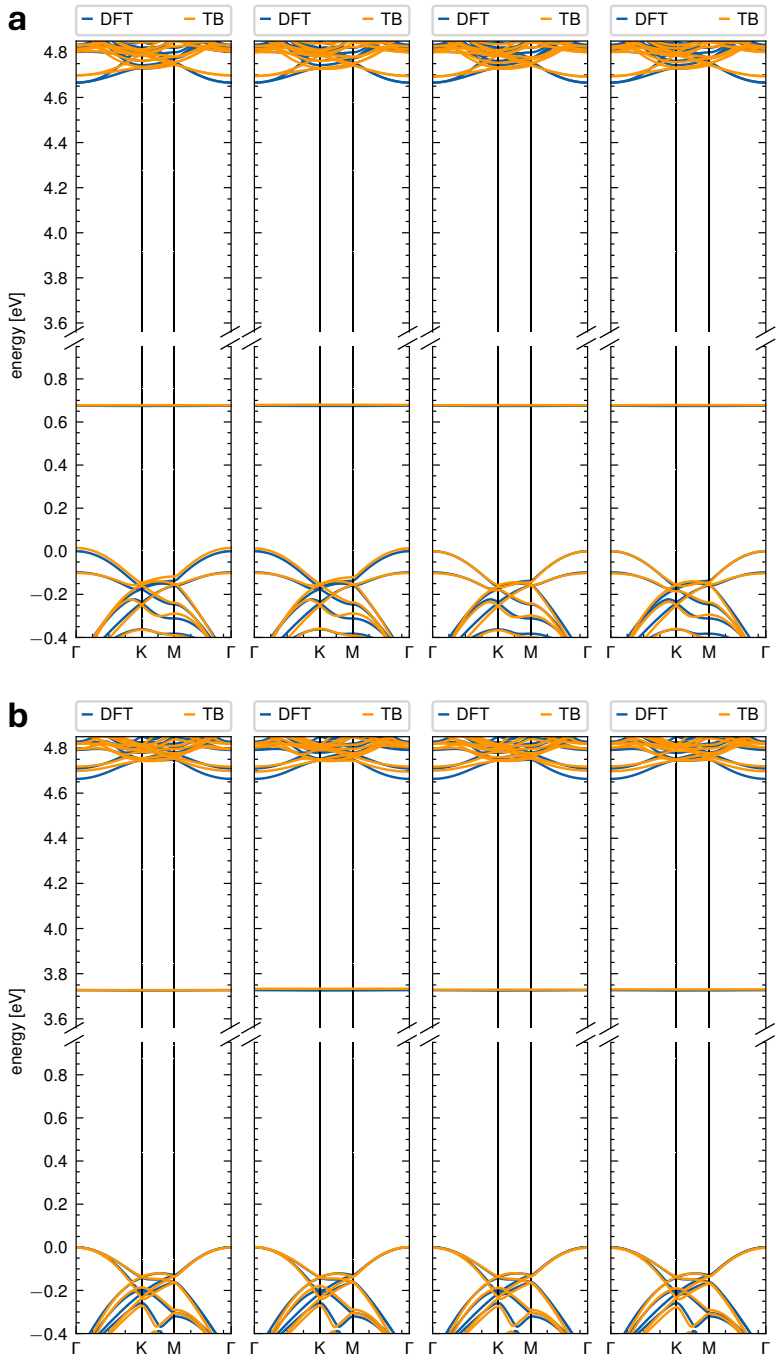

Figure 4.10: DFT (blue) and tight-binding (orange) band structures of the different hopping models for C_N (**a**) and C_B (**b**). From left to right: SR+LR, SR, LR, w/o. The tight-binding band structures are calculated with the parameters obtained from the fit to the PDOS found in table 4.4. The energy axis has been cut. DFT band structures are aligned to be zero for the valence band maximum.

4.6.5 Lattice deformation

In order to run strain-dependent calculations and to be consistent with the tight-binding model, we need to introduce a distance dependence for the hopping terms that include a defect. Given that the valence and conduction bands are expected to be properly described within the LR model, we only need to adjust the distance dependence in a way such that the peak position is described correctly. In order to assess the strain dependence of the position of the defects states, we must carry out two additional DFT calculations (compressive and tensile strain).

Instead of refitting all parameters for the distance dependence of the defect hoppings (α_i, β_i), we find that it is sufficient to introduce a scaling factor of the pristine dependence for each defect:

$$\alpha_i = w_\alpha \alpha_i \tag{4.6}$$

$$\beta_i = w_\beta \beta_i \ . \tag{4.7}$$

We obtain the weights w_α and w_α minimizing the peak difference for the $\pm 3\%$ strain, respectively, and the results can be found in Table 4.3.

Figure 4.11 demonstrates how peak positions vary as a function of strain. Our results show that the tight-binding model accurately captures the overall trend when accounting for the distance dependence of carbon hopping matrix elements. Under strain, the two monomers show different behaviors, with the shift of C_N being less pronounced compared to C_B. The absolute peak difference between DFT and TB is less than 0.01 eV, which is within the range of the unstrained results. The band structures are illustrated in Figure

	C_N	C_B
w_α	1.64	-1.54
w_β	0.47	0.75

Table 4.3: The table contains weights determining the distance dependence for the respective defect hoppings.

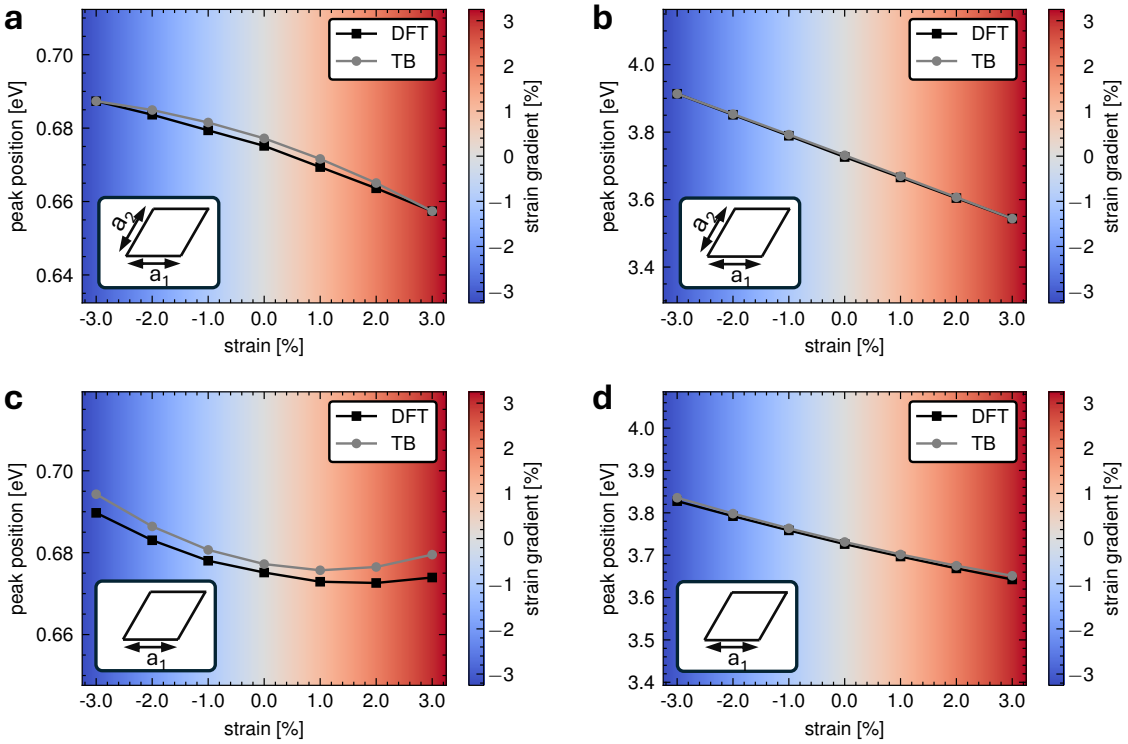

Figure 4.11: Peak position of DFT (black squares) and TB (grey circles) under different strain conditions (gradient). Tight-binding results include distance dependence for the carbon hopping matrix elements. Left: C_N; Right: C_B. Figures **a** and **b** include bi-axial strain, whereas figures **c** and **d** include strain along one axis. The shift of C_N is relatively small compared to C_B for which the state shifts more than 0.1 eV.

4.12. They reproduce the shift of the conduction p_z bands. However, we note that under compressive (negative) strain the sigma band exhibits a shift towards lower energies, resulting in it becoming the conduction band minimum for a compressive strain exceeding -2%. This is also observed for the pristine calculations. Since our tight-binding model only describes the p_z orbitals, it fails to describe these sigma bands and, therefore, the shift of the conduction band minimum.

To evaluate the generalizability of the model, we performed further calculations under strained conditions along a single axis. We find that the tight-binding results agree very well with the DFT. The change in peak position is described correctly and when applied along one axis.

Although these results include two additional DFT calculations, the results demonstrate the capability of the tight-binding model with a minimal set of parameters. By introducing the hopping distance dependence for the new hopping parameters, the model is able to reproduce *ab-initio* band structures. For the followup project, we want to include all strain dependencies in the initial training.

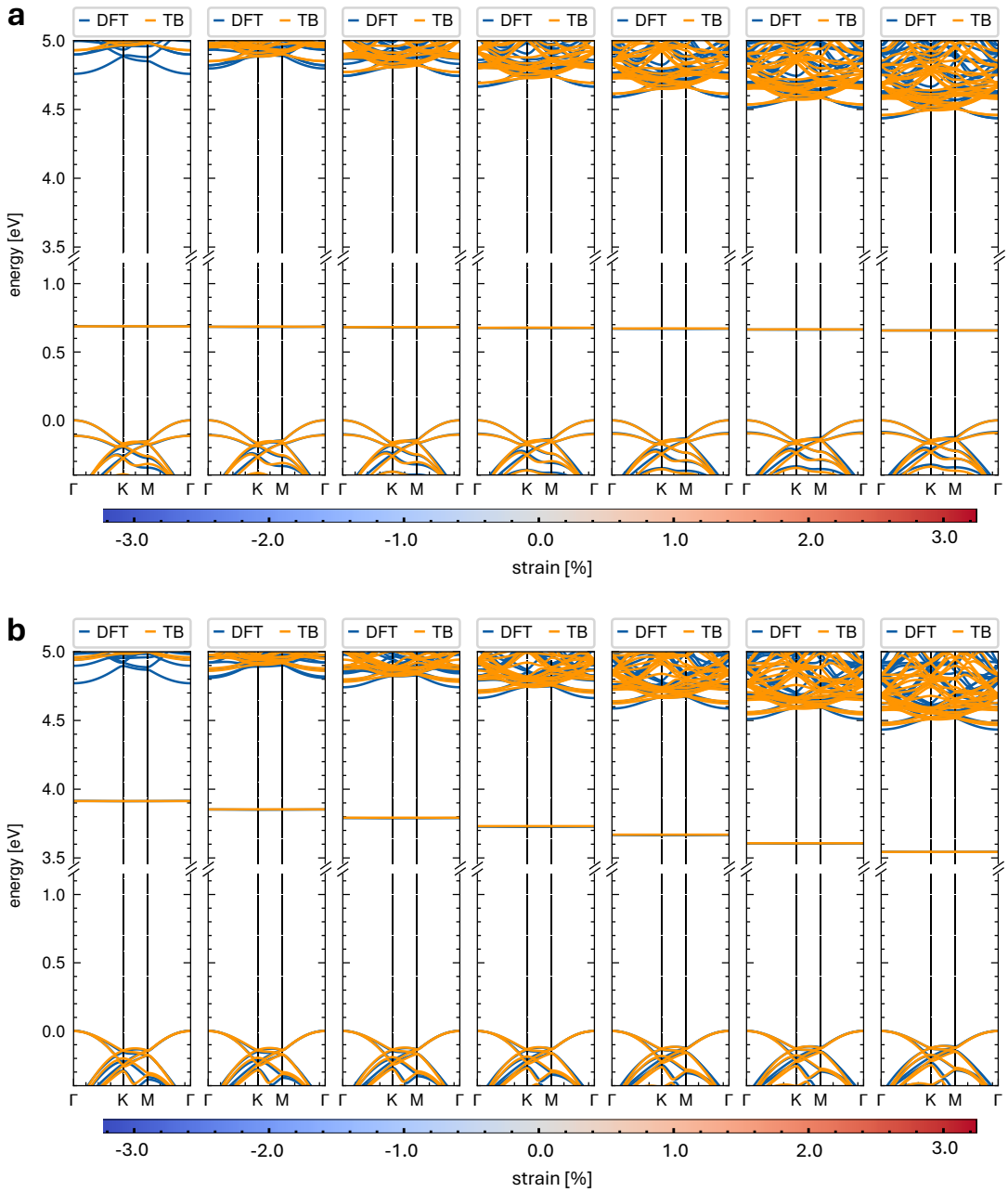

Figure 4.12: DFT (blue) and tight-binding (orange) band structures of the C_N **a** and C_B **b** for varying compressive and tensile strain. The tight-binding results are calculated with the long-range contributions (LR) and include an additional distance dependence for the carbon hoppings (Table 4.3). We observe a minor change in the defect state related to the C_N and a large shift of the defect state related to the C_B. The band gap increases as compressive strain is enhanced. As the p_z bands continue to shift higher for compression beyond -1%, the sigma bands end up lower in energy. The band structures are aligned to be zero at the valence band maximum and the energy axis has been cut for better visibility of the results.

4.6.6 Distance between defects

So far, we have obtained parameters for isolated point defects and analyzed their generalization under strained conditions. Since the supercells of the respective defects have been chosen large enough to maintain the pristine distances at the boundaries, we now want to address how transferable these parameters are for different distances between the defects. To do so, we increased the size of the supercell to $18 \times 9 \times 1$, which allows us to gradually reduce the distance between the carbon atoms. We start with the "dilute" limit, where each defect is at the center of the respective cell. We then sample the distances by moving one of the defects along the connection (Figure 4.13). The respective supercells have been relaxed to their geometric ground state, and we analyze the local density of states for the defect sites and the bands (Figure 4.13).

For both carbon monomers, we find that the characteristics of the bands and defect states are well reproduced up to a distance of 7.5 Å. Below 5 Å, we observe that the TB and DFT results differ, which could be related to the fact that the respective Gaussians begin to overlap (Figure 4.13). A fit with a smaller supercell might therefore be sufficient. However, to ensure smooth integration into larger structures, it is important to maintain neighbor distances to adjacent cells close to pristine values. If one is interested in multiple defects within a material that have different distances and need precise DFT relaxations, one can optimize the geometries of different supercells with multiple defects and use the parameters obtained for these structures.

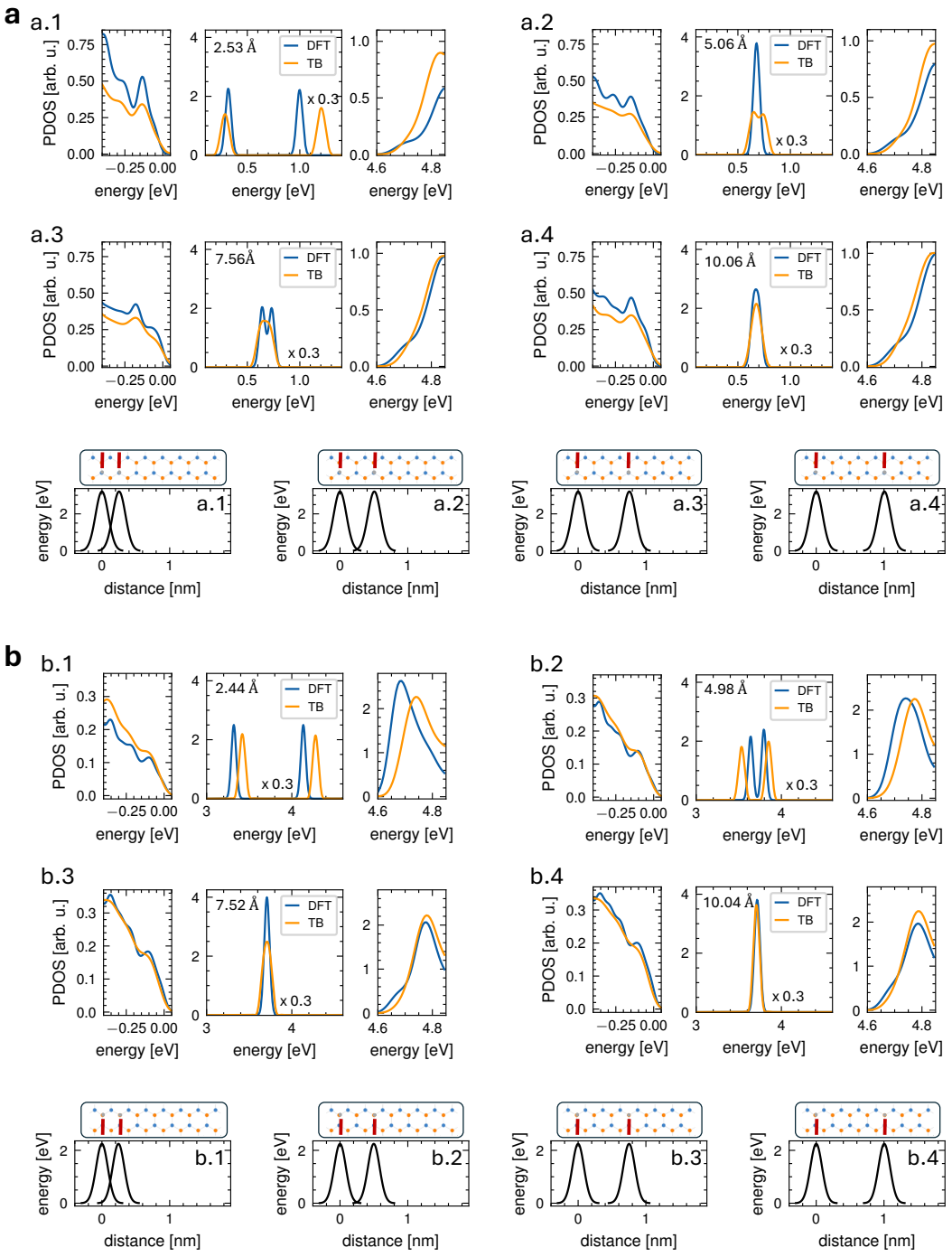

Figure 4.13: PDOS of the valence band (right), defect peaks (middle) and conduction band (right) for two C_N **a** and C_B **b** monomers with increasing distances from **x.1** to **x.4**. DFT PDOS are depicted in blue and TB PDOS in orange. The tight-binding calculations include the long-range contribution. The respective distances are indicated in the lower figures. The model fails to reproduce the characteristics for **x.1** and **x.2**. For **a.3** we observe two peaks, which are visible with reduced broadening.

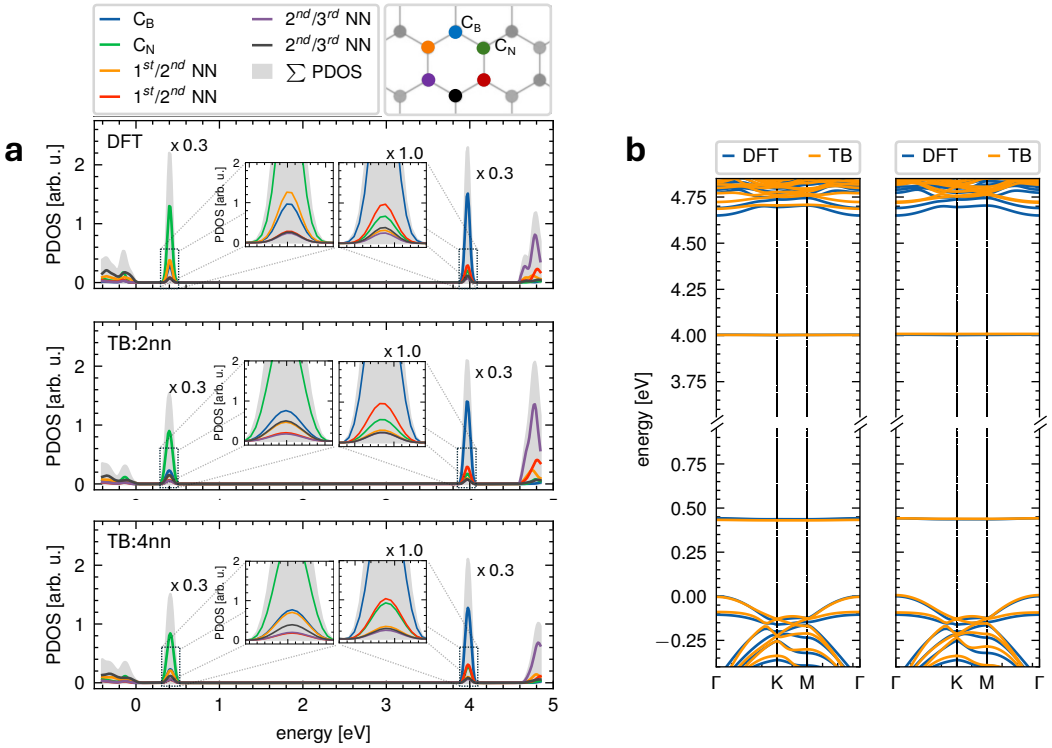

Figure 4.14: **a** Projected densities of states obtained from DFT (upper subplot) and from the fitted TB parameters (lower subplots) for the different number (2 and 4) of PDOS used as input for the training. **b** DFT band structure with corresponding tight-binding results for 2 (left) and 4 (right) PDOS included in the training.

4.6.7 Carbon-dimer

The carbon dimer consists of two carbon atoms that substitute for one pair of neighboring nitrogen and boron atoms. This results in a shift of the two respective monomer states towards the band edges, an effect that is similar to that of two atoms forming a molecule with a bonding state and an antibonding state. This requires the introduction of an additional hopping parameter t_{1st}^{CC}. However, our fitting attempts show that this is not sufficient to obtain a good fit for the positions of both defect peaks at the same time. This aligns with our calculations for varying distances between carbon monomers indicating that for distances where the Gaussian perturbations of the respective defects overlap, the parameters are not transferable (Section 4.6.6). Thus, we have to fit the respective Gaussian dependencies ($\epsilon_N^{defect}, \sigma_N, \epsilon_B^{defect}, \sigma_B$) and the new hopping parameter

C_{BN}	ϵ_N^{defect}	σ_N	ϵ_B^{defect}	σ_B	$t_{1^{st}}^{CC}$
LR 2	4.363	0.077	0.303	0.124	-3.003
LR 4	4.198	0.090	-0.299	0.090	-3.254

Table 4.4: The table contains the predicted parameters for the carbon dimer. Results include two and four PDOS used as input for the training. Subscripts for denote the species of the substituted host atom. The energies are in eV.

($t_{1^{st}}^{CC}$). Additionally, the carbon dimer introduces additional symmetry breaking which leads to different distances between the hopping neighbors in the vicinity of the defect. To properly account for all changes, more hopping parameters than those already used for the respective monomers would be necessary. However, we have obtained a distance dependence for the carbon hopping matrix elements that describe the evolution of the respective peaks under compressive and tensile strain (Section 4.6.5). These can be used to properly account for the different distances of the carbon atoms from its neighbors.

We perform tests using varying numbers of PDOS throughout the training phase and discover that the inclusion of the two defect sites offers the most accurate description of the conduction band (Figure 4.14). The split off band is captured with the tight-binding model, when we include two PDOS. Although the splitting itself is similar, we observe a difference between the conduction band minima, similar to the C_B. Nevertheless, our very simple model reproduces the main features of the band structure, including the valence band, the defect states and the split off conduction band.

4.7 Discussion

In this chapter, we proposed an alternative way to access the electronic properties of defects. As seen in the previous chapter, the necessity of a supercell for defect calculations limited us to total energy evaluations for complex materials like $CuInS_2$ and $CuGaS_2$. Although this allowed us to access some properties of the defects, we have reached the computational limits.

Here, we proposed a new way to construct semi-empirical models, at a tight-binding

level, of technologically relevant defects by means of a machine learning approach. The embedding of the defect perturbs the pristine Hamiltonian and the refit of all the parameter gets out of hand quickly. We suggest using a perturbed version of a pristine tight-binding model along with additional defect parameters to overcome this problem and to enable a description with only a few parameters.

We make use of the atom- and orbital-projected density of states to obtain the parameters, instead of the common way of fitting parameters to the band structure. This overcomes the well known problem of disentanglement of the bands of the supercell. We demonstrate that a neural network is capable of predicting a tight-binding model from the projected density of states (PDOS) for defects with localized peaks within a large band gap. We show that these parameters are also reproducing the band structure, thus confirming the PDOS as a well suited observable for fitting.

We investigate different adjustments of the pristine hopping terms and find that neither a short-range (additional parameter) nor long-range contribution (distance dependent hopping parameters) show a significant improvement of the results. However, we propose to include the long-range contribution, as it allows to describe strain dependencies. Therefore, a very simple tight-binding model is able to deliver an adequate description of carbon defects in hBN.

Our tight-binding approach is able to describe defects under strained conditions. The distance dependence for the hopping parameters obtained for the pristine crystal under strained conditions is transferable, and we only need to introduce a distance dependence for the carbon hopping terms. Both distance dependencies have been obtained for biaxial strain. A comparison of uniaxial strain suggests that they accurately describe the behavior. We used these results to build a tight-binding model for the well known carbon dimer, expected to be responsible for the single photon emission at around 4.1 eV.

Calculations for varying distances between two monomers indicate that the parameters are also transferable to systems where the defects are closer to each other. Comparisons for different distances show that the DFT results are well reproduced up to the point where the respective Gaussian functions for the onsite perturbations overlap.

We emphasize that, while this has been done for relatively simple defects, the use of the projected density of states is expected to be transferable to more complex systems. The projection on the orbitals enables a simple access to the description of multi-orbital states. The machine learning architecture might need to be adapted to describe a more complex tight-binding description, but this is not expected to be a barrier and we suggest a few model improvements in this work.

In conclusion, we show that a tight-binding model consisting of defect parameters and a perturbation of the pristine onsite energies is sufficient to describe the carbon monomers in monolayer hBN. We propose a simple model for the carbon dimer that makes use of the respective monomer parameters, thus enabling a description with few parameters for more complex carbon defects. We demonstrate that the tight-binding Hamiltonian can be fitted to the projected density of states with a neural network. This opens a novel pathway for obtaining tight-binding parameters for defective systems.

The workflow has been successfully applied using hBN as a test scenario. Here, a simple Gaussian function was enough to describe the carbon monomers. However, we believe that, for more complex materials, the key to success might be the correct description of the onsite perturbation. We tested to include multiple Gaussian functions and different dependencies, such as the Laplacian but for the present defects no improvement was found. However, a more complex description might be necessary for other materials.

Another point to address is the distance dependence for the defect hopping parameters that are necessary to describe the strain properly. Currently, they are obtained by correcting the peak position for additional strain-dependent DFT calculations for the defective supercell. Although this aligns with the theory, we believe that it would be beneficial to integrate this into the training phase. In particular, one idea is to include geometry information in the training, as well as parameters for the distance dependence.

Chapter 5

Discussion and perspectives

Summary

In this thesis, we discuss various approaches for describing defects in solids. In the first part, we use *ab-initio* methods to investigate intrinsic point defects in $CuInS_2$ and $CuGaS_2$, thereby revealing some of the limitations. State-of-the-art calculations primarily involve charge transition level calculations, which are sometimes compared with optical spectra without further discussion.

However, they do not describe the same mechanism. Charge transition levels are ionization energies, meaning that the lattice has time to accommodate the additional charge and relax to its ground state. Although this phenomenon is well known and discussed in theory, we demonstrate that the lattice contribution is also significant for defects in chalcopyrites. Our calculations show that the difference is substantial for deep defects and should not be ignored. We hope this serves as an incentive for future studies that compare these results to optical measurements, emphasizing the need to be aware of the limitations. Due to the complexity of chalcopyrites, the calculations are challenging; however, vertical transition levels are feasible whenever charge transition levels are.

We can improve the accuracy and comparability in describing optical transitions using the Δ-SCF approach, sometimes referred to as constrained DFT (cDFT). This was also the initial goal for studying chalcopyrites. However, due to the complexity of these materials

and the lack of well-defined Kohn-Sham states, most calculations remained unconverged. The only two defects for which we were able to obtain Δ-SCF results were the donor defects In_{Cu} and Ga_{Cu}. A comparison with the vertical transition levels shows good agreement for the absorption from the valence band to the defect. However, we observed a difference in the relaxed geometries of the excited state, which is reflected in different transition characteristics. A future study should try to pinpoint the origin of this discrepancy.

The difficulties and constraints observed in the chalcopyrites led us to explore alternative methods to investigate defects. Semi-empirical tight-binding is a practical tool because it can handle supercells containing thousands of atoms, unlike DFT. However, determining the optimal way to derive tight-binding parameters for defects remains a challenging problem.

A defect can be modeled by embedding it within a supercell of the pristine crystal. This embedded defect perturbs the crystal, creating multiple challenges in obtaining accurate parameters. In theory, because of this perturbation, all the parameters of the supercell must be fitted. Another issue is the fitting observable. Tight-binding parameters are commonly fitted to the band structure; however, the supercell's increased lattice constant in real space results in band folding in reciprocal space. Both problems complicate the fitting process and must be addressed.

We propose a new workflow for obtaining tight-binding parameters for defective systems. Considering the defect as a perturbation to the host crystal introduces two distance dependencies for the pristine tight-binding parameters. Rather than refitting all parameters of the supercell, this approach minimizes the number of fitting parameters. We address the cumbersome disentanglement of the folded band structure by using the atom- and orbital-projected density of states (PDOS).

This approach allows us to train a neural network that learns the mapping between the tight-binding PDOS as input and the tight-binding parameters as output. The neural network can then predict the tight-binding parameters from a PDOS obtained with a single DFT calculation. To demonstrate the method's capabilities, we tested the workflow on carbon substitutional defects in hBN. The use of PDOS results in band structures that

are consistent with *ab-initio* results, addressing the challenges faced by previous studies using total DOS [149]. Furthermore, we evaluated the tight-binding model under strained conditions and showed that it accurately describes the behavior of the bands and peak positions.

In conclusion, we have used various approaches to understand defects in crystals. In the first part, we employed state-of-the-art DFT methods to investigate the intrinsic defect landscape in $CuInS_2$ and $CuGaS_2$, accompanied by a valuable discussion of differences related to optical transitions. This enabled us to identify potential candidates for the different transitions observed in photoluminescence measurements. In the second part, our focus shifted toward method development by training a neural network on a few parameter tight-binding model for defects that utilizes the atom- and orbital-projected density of states instead of the band structure.

Outlook

Both parts still present open questions that we partially already discussed throughout the thesis. Regarding chalcopyrites, although we identified candidates for some of the experimentally observed optical transitions, a third acceptor has not yet been found. One possibility is extrinsic defects. Previous works have addressed various extrinsic defects, such as antimony (Sb), phosphorus (P), and others [98, 109–111], but theoretical work on sodium is still lacking. Because chalcopyrites are grown on soda lime glass [11], sodium can migrate to the absorption layer. Another focus could be on defect complexes, for instance, Yang et al. [109] investigated the formation energies of intrinsic complexes, while Xiang et al. [110] examined complexes involving a phosphorus defect. The supercells used in these studies are relatively small (32–64 atoms), so an additional study with larger supercells to address possible defect complexes could be beneficial. Furthermore, the present studies lack a comparison to optical properties.

Further improvements for the work presented in the first part of this thesis will emerge over time, thanks to increased computational resources in the future. However, we should not rely solely on computational power, but also focus on method development. In

this context, machine learning has proven to be a powerful tool for electronic structure calculations [153]. This is also reflected in the field of defects. For example, Mosquera-Lois et al. [154] use machine learning to enhance geometric optimization, while Rahman et al. [155] implement a graph neural network that can predict charge transition levels. Exploring similar ideas to obtain defect properties more cost-efficiently is both interesting and directly linked to the machine learning project we have initiated.

We plan to further test the robustness of the tight-binding parameters for the carbon defects in hBN and refine the discussed workflow. One approach to evaluate robustness is to develop an accurate description of the excitons associated with the defects in hBN, which aligns with our interest in the optical properties of defects. To this end, we will explore the possibility of using these parameters within a Bethe-Salpeter framework, as introduced by Galvani et al. [156].

Building on the promising results obtained for carbon defects in hBN, our next objective is to extend and test our workflow on multi-orbital systems. In order to be able to describe more complex materials, such as chalcopyrites, we have to gradually increase the complexity of the tight-binding model. The next step is therefore to apply the workflow for transition-metal dichalcogenides, which is a promising system to explore the transferability to a multi-orbital system. The distance dependence of the hoppings between the nearest neighbors should be transferable to any system since it is obtained from strain-dependent calculations of the pristine crystal separately. However, one idea is to include the fit of the pristine parameters (including their distance dependence) into the training itself and investigate if the PDOS alone it is enough to reproduce the band structure. The machine learning model also offers room for improvement and it might be beneficial to include the geometry information of the system.

In the long run, we want to further develop the workflow so that users can obtain tight-binding parameters for various defects in different materials. We intend to explore two concepts. First, we plan to design a neural network trained on additional information rather than solely on the projected density of states. This approach would allow us to either utilize pretrained models and learn new materials or incorporate various materials

(and their defects) into a single dataset. It would be interesting to investigate whether this leads to transferability to previously unseen systems. Secondly, we want the code to be accessible to the general user. To achieve this, we aim to create an out-of-the-box workflow that requires only minimal inputs (e.g., the *ab-initio* results).

Final remarks

We have observed both the beauty and the challenges associated with calculating the properties of defects. A good description of defects is necessary to accelerate our understanding, which can, for example, help to improve the efficiency of applications such as solar cells in the future. We have seen that state-of-the-art methods allow us to characterize defects; however, we have also encountered their limitations. We have introduced an initial step toward a novel approach for obtaining tight-binding parameters using machine learning and the projected density of states. This approach offers great potential for macroscopic descriptions of defects, and we hope that further improvements in this direction will eventually enable us to routinely describe defects in complex materials like chalcopyrites.

Appendix A

Carbon dimer: $\triangle$-SCF

The following part is based on [23] to clarify how we can construct the excited state for a spin-degenerate system with single slater determinants.

The carbon dimer has two distinct defect states within the band gap. For spin-degenerate cases where the final state is unoccupied, the singlet excitation state can be expressed as [23]:

$$|S\rangle = \frac{1}{\sqrt{2}}(|v_{\downarrow}d_{\uparrow}\rangle - |v_{\uparrow}d_{\downarrow}\rangle) \ , \tag{A.1}$$

which is the asymmetric solution for two excitations that include either a spin up or a spin down promotion from the valence band v to the defect d and therefore includes two Slater determinants. However, this can be overcome by considering degenerate triplet excitations

$$|T; -1\rangle = (|v_{\downarrow}d_{\downarrow}\rangle) \ , \tag{A.2}$$

$$|T;\ 0\rangle = \frac{1}{\sqrt{2}}(|v_{\downarrow}d_{\uparrow}\rangle + |v_{\uparrow}d_{\downarrow}\rangle) \ , \tag{A.3}$$

$$|T; +1\rangle = (|v_{\uparrow}d_{\uparrow}\rangle) \tag{A.4}$$

and defining a mixed state $|S/T\rangle = |v_{\downarrow}d_{\uparrow}\rangle$ for a neutral spin excitation. We can obtain this single Slater determinant with a superposition ($|S/T\rangle = \frac{1}{\sqrt{2}}(|S\rangle + |T;\ 0\rangle)$) and we can

evaluate $\langle S/T| \mathcal{H} |S/T\rangle$

$$\langle v_\downarrow d_\uparrow| \mathcal{H} |v_\downarrow d_\uparrow\rangle = \frac{1}{2}(\langle S| \mathcal{H} |S\rangle + \langle T;\ 0| \mathcal{H} |T;\ 0\rangle) \tag{A.5}$$

where the parts with both $|S\rangle$ and $|T;\ 0\rangle$ vanish because they are orthogonal eigenstates of the system. We finally rearrange for the singlet excitation

$$\underbrace{\langle S| \mathcal{H} |S\rangle}_{E(S)} = 2\underbrace{\langle S/T| \mathcal{H} |S/T\rangle}_{E(S/T)} - \underbrace{\langle T| \mathcal{H} |T\rangle}_{E(T)} \ . \tag{A.6}$$

Because of the degeneracy of the triplet states both $E(S/T)$ and $E(T)$ are representable by total energies of single Slater determinants that we can construct by defining the occupations for a DFT calculation.

Appendix B

Computational details

B.1 Calculation details chapter 3

The *ab-initio* calculation were performed with the VASP software package [157–160]. The cutoff energy for the plane waves is converged at 400 eV. The forces for the converged of $1 \cdot 10^{-3}$ eV/Å and an electronic sc-criteria of $1 \cdot 10^{-6}$ eV was used. For all caclulations we used the PBE exchange correlation functional and the projector augmented-wave method (PAW).

For the parabolicity calculations, the conventional cell was pre-relaxed with a GGA functional before relaxing the structure for different HSE parameters. Since this includes many calculations, these calculations were performed with a 2×2×1 Monkhorst-Pack grid to pre-screen the band gaps. We increased the mesh density to 4×4×2 to find a range of band gaps on which to perform parabolic calculations.

For atomic relaxation, the defect and its nearest neighbors were displaced from their equilibrium positions. Initially, a prerelaxation step was carried out using a semi-local functional with fixed supercell dimensions and without enforcing symmetry constraints, employing an atomic convergence criterion of $1 \cdot 10^{-3}$ eV/Å and an electronic convergence criterion of $1 \cdot 10^{-5}$ eV. Subsequent relaxations using the HSE functional included symmetry constraints and utilized a convergence criterion for the forces of $1 \cdot 10^{-1}$ eV/Å and for the electronic convergence of $1 \cdot 10^{-4}$ eV .

For the parabolicity calculations we keep the geometry fixed. We identify the trends for $CuInS_2$ for a Monkorst-Pack grid of $4 \times 4 \times 2$ and expect a similar behavior for the $CuGaS_2$.

For all charge difference and density plots an isosurface of $3 \cdot 10^{-4}$ was used.

B.2 Computational details chapter 4

The *ab-initio* calculations have been performed with the VASP software package [157–160]. For all density functional theory calculations, we used the Perdew-Burke-Ernzerhof (PBE) exchange correlation functional and the projector augmented-wave method (PAW). The pristine unit cell has been optimized prior to the defect calculations using a plane wave cutoff at 580 eV and a 9×9×1 Monkhorst-Pack grid for the sampling of the first Brillouin zone. The total energy did not change more than $1 \cdot 10^{-4}$ eV after a vacuum distance of 10 Å.

The defects were embedded in a supercell of 9×9×1 to include all high symmetry points in the folding process. The forces have been converged to be below $1 \cdot 10^{-3}$ eV/Å, whereas the electronic optimization converged for $1 \cdot 10^{-5}$ eV. The projected density of states was post processed from the VASP calculations with the VASPKIT package [161]. The projected density of states converged for a Monkhorst-Pack grid of 9×9×1 in the 1^{st} BZ of the supercell (corresponding to 81×81×1 in the 1^{st} BZ of the primitive unit cell) with a Gaussian smearing of 0.04 eV.

The tight-binding models have been implemented using the pybinding package [162]. The projected density of states has been calculated with the kernel polynomial method and a Gaussian broadening of 0.04 eV as implemented in pybinding.

For the training of the neural network, we used the Adam optimizer with a learning rate of $\eta = 1.0 \cdot 10^{-3}$ and an exponential learning rate decay with $\lambda = 0.99$. Before the training we split the data into train, validation and test data (80/10/10) to monitor the training and possible over-fitting. The training converges for 700 epochs (Figure C.2) and takes around 1.5-3h on a NVIDIA Tesla V100 SXM2 16G/32G.

The first component of the neural network is a 1D convolutional layer. The convolu-

tional layer is followed by a self-attention mechanism and a multilayer perceptron. The dimensions of the neural network can be found in the Table B.1.

Type	dimensions
Conv1D	[4, 32, 8, 0]
self-attention:	
Head num	1
Head dim	128
MLP	[128, 256, 512, 256, 64]

Table B.1: Dimensions used for the layers of the neural network.

Appendix C

Supplemental information chapter 4

C.1 Data generation

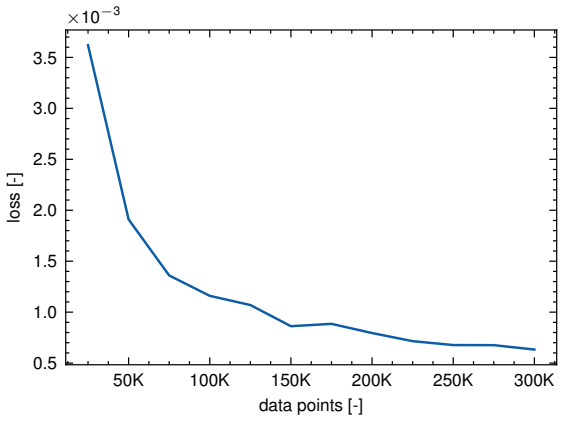

Figure C.1: Average loss of three training runs of the test set versus data points for the carbon on nitrogen defect including SR+LR contributions.

The data set is obtained by a uniform distribution of defect parameters ϵ_i^{defect}, σ_i and $t_{n',m}^{defect}$. The ranges are shown in Table C.1. We used 200K data points, as the average test loss of 3 trainings did not improve much after that (Figure C.1).

For each combination of tight-binding parameters, a PDOS is created on the site of the defect and its nearest neighbors up to the 3rd nearest neighbor. The PDOS is sampled by 400 steps to identify the optimal number of PDOS for the training and is increased to

$C_{monomer}$	ϵ^{defect}	σ	t_{1st}	t^{CN}_{2nd}	t^{CB}_{2nd}	t_{3rd}	t^{BB}_{SR}	t^{NN}_{SR}
lower limit	-1.00	0.00	-3.5	-0.15	-1.25	-0.85	-1	-0.5
upper limit	3.30	0.15	-2.5	0.15	-0.25	0.15	0	0.5

Table C.1: The table contains the energy ranges for the data generation of the data sets used for the number of PDOS evaluation.

800 for the hopping analysis. We also narrow down the defect's onsite energy ranges with the knowledge from the previous training. Since there is no Jahn-Teller distortion for the investigated defects, it is sufficient to take the PDOS at one of the neighbors, respectively. If needed, this can easily be adapted.

C.2 Loss curves

Two examples for training and validation loss curves are displayed in Figure C.2. The two examples are for different samplings of the projected density of states. Independently of the input sampling, the training is done at around 700 epochs.

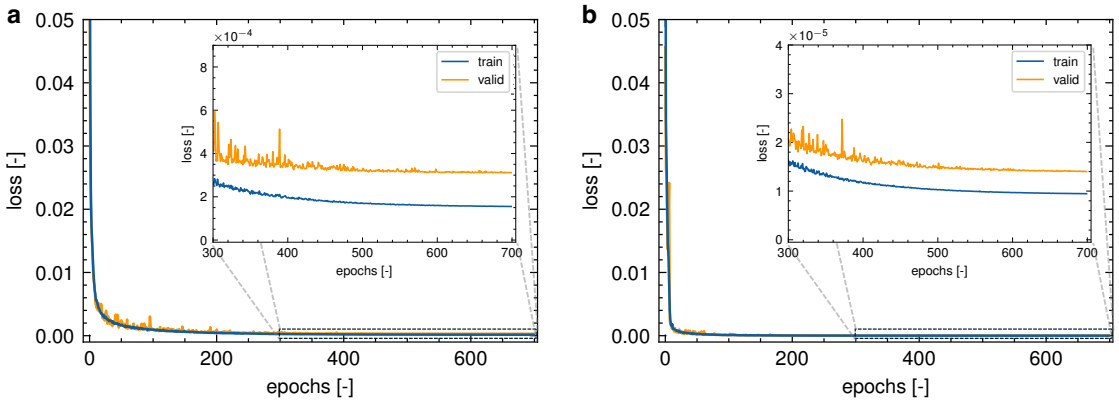

Figure C.2: Example loss curves for trainings (blue) and validation (orange) set of a SR+LR training with an PDOS sampling of **a** 400 and **b** 800. The training included three PDOS and their sum. The training is finished after 700 Epochs.

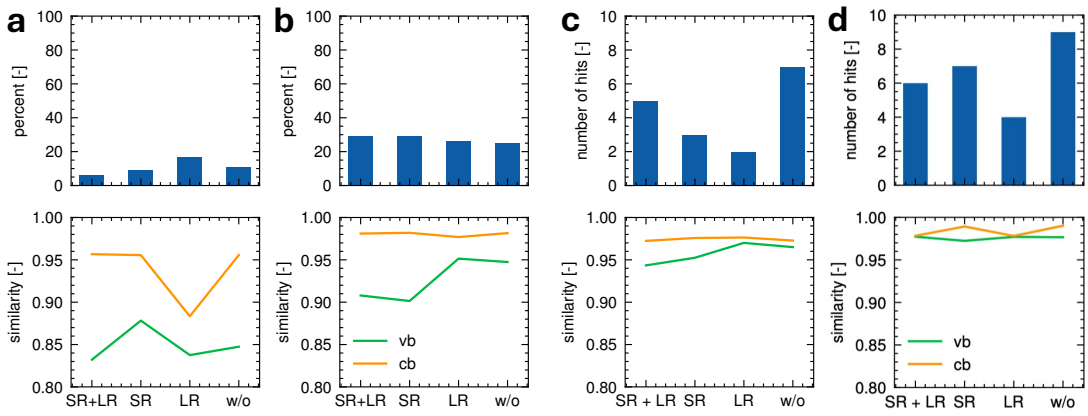

Figure C.3: Standard fitting tools for **a** C_N and **b** C_B. Machine learning results: Number of successful runs for **c** C_N and **d** C_B. Upper plots are the number of hits. Lower plots cosine similarities or the PDOS of the respective bands. The optimization has been done for three PDOS.

C.3 Comparison with the standard fitting procedure

In this section, we give insight into the difficulty of the task for standard fitting procedures. To objectively evaluate the difficulty, we perform an optimization using a standard Nelder-Mead algorithm with 100 random initial guesses. (We use the same range of initial guesses as for the ML data generation and we again use three PDOS.) We find that for the majority of initial guesses (in particular those which do not display an initial defect state), we fail to converge towards the DFT result.

Additionally, we find that for C_N the chance of finding such a peak is around 9% to 17% for all different hopping contributions and around 25% to 29% for C_B (Figure C.3). Of those successful runs, the bands are not always described properly, indicating that the algorithm is stuck in a local minimum. The average cosine similarity for the bands shows an overall worse description of the bands than our machine learning results (Figure C.3). We observe that the machine learning method outperforms not only the number of correct peak predictions but also the description of the bands.

References

[1] Jeff Tollefson. “Earth breaches 1.5 ℃ climate limit for the first time: what does it mean?” In: *Nature* (Jan. 2025). ISSN: 1476-4687. DOI: 10.1038/d41586-025-00010-9.

[2] United Nations. *Paris Agreement*. Accessed: 14.06.2025. 2015. URL: https://unfccc.int/files/essential_background/convention/application/pdf/english_paris_agreement.pdf.

[3] Svante Arrhenius. “XXXI. On the influence of carbonic acid in the air upon the temperature of the ground”. In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 41.251 (1896), pp. 237–276. DOI: 10.1080/14786449608620846.

[4] *Scientific Consensus - NASA Science*. [Online; accessed 17. Feb. 2025]. Oct. 2024. URL: https://science.nasa.gov/climate-change/scientific-consensus.

[5] European Environment Agency. *Renewable Energy Consumption in the EU*. Accessed: 19. Feb. 2025. 2023. URL: https://www.eea.europa.eu/en/analysis/indicators/share-of-energy-consumption-from?utm_source=chatgpt.com.

[6] Daniel Abou-Ras, Thomas Kirchartz, and Uwe Rau. *Advanced Characterization Techniques for Thin Film Solar Cells*. Mar. 2011. ISBN: 978-3-52741003-3. DOI: 10.1002/9783527636280.

[7] Arno Smets et al. *Solar Energy: The Physics and Engineering of Photovoltaic Conversion, Technologies and Systems*. Bloomsbury USA, Feb. 2016. ISBN: 978-1-90686032-5.

[8] Conrad Spindler et al. “Electronic defects in Cu(In,Ga)Se_2: Towards a comprehensive model”. In: *Phys. Rev. Mater.* 3.9 (Sept. 2019), p. 090302. DOI: 10.1103/PhysRevMaterials.3.090302.

[9] Jan Keller et al. “High-concentration silver alloying and steep back-contact gallium grading enabling copper indium gallium selenide solar cell with 23.6% efficiency”. In: *Nat. Energy* 9 (Apr. 2024), pp. 467–478. ISSN: 2058-7546. DOI: 10.1038/s41560-024-01472-3.

[10] Damilola Adeleye et al. “Composition dependence of electronic defects in $CuGaS_2$”. In: *Prog. Photovoltaics Res. Appl.* 32.8 (Aug. 2024), pp. 528–545. ISSN: 1062-7995. DOI: 10.1002/pip.3778.

[11] Sudhanshu Shukla et al. “Over 15% efficient wide-band-gap Cu(In,Ga)S_2 solar cell: Suppressing bulk and interface recombination through composition engineering”. In: *Joule* 5.7 (July 2021), pp. 1816–1831. ISSN: 2542-4351. DOI: 10.1016/j.joule.2021.05.004.

[12] J. R. Botha et al. “Photoluminescence of vacuum-deposited $CuGaS_2$ thin films”. In: *Physica B* 308-310 (Dec. 2001), pp. 1065–1068. ISSN: 0921-4526. DOI: 10.1016/S0921-4526(01)00848-1.

[13] J. Eberhardt et al. “Optical properties of epitaxial $CuGaS_2$ layers on Si(111)”. In: *J. Phys. Chem. Solids* 64.9 (Sept. 2003), pp. 1781–1785. ISSN: 0022-3697. DOI: 10.1016/S0022-3697(03)00199-9.

[14] Mete Atatüre et al. “Material platforms for spin-based photonic quantum technologies”. In: *Nat. Rev. Mater.* 3 (May 2018), pp. 38–51. ISSN: 2058-8437. DOI: 10.1038/s41578-018-0008-9.

[15] Jared Rovny et al. “Nanoscale diamond quantum sensors for many-body physics”. In: *Nat. Rev. Phys.* 6 (Dec. 2024), pp. 753–768. ISSN: 2522-5820. DOI: 10.1038/s42254-024-00775-4.

[16] Kamyar Parto et al. “Defect and strain engineering of monolayer WSe_2 enables site-controlled single-photon emission up to 150 K”. In: *Nat. Commun.* 12.3585 (June 2021), pp. 1–8. ISSN: 2041-1723. DOI: 10.1038/s41467-021-23709-5.

[17] Marianne Etzelmüller Bathen and Lasse Vines. “Manipulating Single-Photon Emission from Point Defects in Diamond and Silicon Carbide”. In: *Adv. Quantum Technol.* 4.7 (July 2021), p. 2100003. ISSN: 2511-9044. DOI: 10.1002/qute.202100003.

[18] Yi Yu et al. “Tunable single-photon emitters in 2D materials”. In: *Nanophotonics* 13.19 (Aug. 2024), pp. 3615–3629. ISSN: 2192-8614. DOI: 10.1515/nanoph-2024-0050.

[19] Toan Trong Tran et al. “Quantum emission from hexagonal boron nitride monolayers”. In: *Nat. Nanotechnol.* 11 (Jan. 2016), pp. 37–41. ISSN: 1748-3395. DOI: 10.1038/nnano.2015.242.

[20] Zai-Quan Xu et al. “Single photon emission from plasma treated 2D hexagonal boron nitride”. In: *Nanoscale* 10.17 (May 2018), pp. 7957–7965. ISSN: 2040-3364. DOI: 10.1039/C7NR08222C.

[21] Brian Shevitski et al. “Blue-light-emitting color centers in high-quality hexagonal boron nitride”. In: *Phys. Rev. B* 100.15 (Oct. 2019), p. 155419. DOI: 10.1103/PhysRevB.100.155419.

[22] Romain Bourrellier et al. “Bright UV Single Photon Emission at Point Defects in h-BN”. In: *Nano Lett.* 16.7 (July 2016), pp. 4317–4321. ISSN: 1530-6984. DOI: 10.1021/acs.nanolett.6b01368.

[23] M. Mackoit-Sinkevičienė et al. “Carbon dimer defect as a source of the 4.1 eV luminescence in hexagonal boron nitride”. In: *Appl. Phys. Lett.* 115.21 (Nov. 2019). ISSN: 0003-6951. DOI: 10.1063/1.5124153.

[24] Michael Winter et al. “Photoluminescent properties of the carbon-dimer defect in hexagonal boron-nitride: A many-body finite-size cluster approach”. In: *Physical Review Materials* 5.9 (2021), pp. 1–29. ISSN: 24759953. DOI: 10.1103/PhysRevMaterials.5.095201. arXiv: 2107.12795.

[25] Cesar Jara et al. “First-Principles Identification of Single Photon Emitters Based on Carbon Clusters in Hexagonal Boron Nitride”. In: *J. Phys. Chem. A* 125.6 (Feb. 2021), pp. 1325–1335. ISSN: 1089-5639. DOI: 10.1021/acs.jpca.0c07339.

[26] J. Plo et al. “Isotope substitution and polytype control for point defects identification: the case of the ultraviolet color center in hexagonal boron nitride”. In: *arXiv* (May 2024). DOI: 10.48550/arXiv.2405.20837. eprint: 2405.20837.

[27] Noah Mendelson et al. “Identifying carbon as the source of visible single-photon emission from hexagonal boron nitride”. In: *Nat. Mater.* 20 (Mar. 2021), pp. 321–328. ISSN: 1476-4660. DOI: 10.1038/s41563-020-00850-y.

[28] Philipp Auburger and Adam Gali. “Towards ab initio identification of paramagnetic substitutional carbon defects in hexagonal boron nitride acting as quantum bits”. In: *Phys. Rev. B* 104.7 (Aug. 2021), p. 075410. DOI: 10.1103/PhysRevB.104.075410.

[29] Ph. Lambin et al. “Long-Range Interactions between Substitutional Nitrogen Dopants in Graphene: Electronic Properties Calculations”. In: *Physical Review B* 86.4 (July 2012), p. 045448. ISSN: 1098-0121, 1550-235X. DOI: 10.1103/PhysRevB.86.045448. (Visited on 01/24/2024).

[30] Christoph Schattauer et al. “Machine Learning Sparse Tight-Binding Parameters for Defects”. In: *npj Computational Materials* 8.1 (May 2022), p. 116. ISSN: 2057-3960. DOI: 10.1038/s41524-022-00791-x.

[31] N. W. Ashcroft and N. D. Mermin. *Solid State Physics*. Holt-Saunders, 1976.

[32] Walter A. Harrison. *Electronic Structure and the Properties of Solids: The Physics of the Chemical Bond*. Mineola, NY, USA: Dover Publications, July 1989. ISBN: 978-0-48666021-9.

[33] Rudolf Gross and Achim Marx. *Festkörperphysik*. München, Germany: Oldenbourg Wissenschaftsverlag, Sept. 2012. ISBN: 978-3-48671294-0.

[34] Peter Y. Yu and Manuel Cardona. *Fundamentals of Semiconductors*. Berlin, Germany: Springer, 2010. ISBN: 978-3-642-00710-1. URL: https://link.springer.com/book/10.1007/978-3-642-00710-1.

[35] M. Born and R. Oppenheimer. "Zur Quantentheorie der Molekeln". In: *Annalen der Physik* 389.20 (Jan. 1927), pp. 457–484. DOI: 10.1002/andp.19273892002.

[36] Wenhui Mi et al. "Orbital-Free Density Functional Theory: An Attractive Electronic Structure Method for Large-Scale First-Principles Simulations". In: *Chem. Rev.* 123.21 (Nov. 2023), pp. 12039–12104. ISSN: 0009-2665. DOI: 10.1021/acs.chemrev.2c00758.

[37] P. Hohenberg and W. Kohn. "Inhomogeneous Electron Gas". In: *Phys. Rev.* 136.3B (Nov. 1964), B864–B871. ISSN: 0031-899X. DOI: 10.1103/PhysRev.136.B864.

[38] W. Kohn and L. J. Sham. "Self-Consistent Equations Including Exchange and Correlation Effects". In: *Phys. Rev.* 140.4A (Nov. 1965), A1133–A1138. ISSN: 0031-899X. DOI: 10.1103/PhysRev.140.A1133.

[39] John P. Perdew, Kieron Burke, and Matthias Ernzerhof. "Generalized Gradient Approximation Made Simple". In: *Phys. Rev. Lett.* 77.18 (Oct. 1996), pp. 3865–3868. DOI: 10.1103/PhysRevLett.77.3865.

[40] Gabriela Herrero Saboya. "Defects in silicon : revisiting theoretical frameworks to guide ab initio characterization". PhD thesis. Toulouse, France: Université Paul Sabatier - Toulouse III, Nov. 2020. URL: https://univ-tlse2.hal.science/tel-03158564.

[41] John P. Perdew et al. "Density-Functional Theory for Fractional Particle Number: Derivative Discontinuities of the Energy". In: *Phys. Rev. Lett.* 49.23 (Dec. 1982), p. 1691. DOI: 10.1103/PhysRevLett.49.1691.

[42] John P. Perdew and Mel Levy. “Physical Content of the Exact Kohn-Sham Orbital Energies: Band Gaps and Derivative Discontinuities”. In: *Phys. Rev. Lett.* 51.20 (Nov. 1983), pp. 1884–1887. DOI: 10.1103/PhysRevLett.51.1884.

[43] J. F. Janak. “Proof that $\frac{\partial E}{\partial n_i} = \epsilon$ in density-functional theory”. In: *Phys. Rev. B* 18.12 (Dec. 1978), pp. 7165–7168. DOI: 10.1103/PhysRevB.18.7165.

[44] Jochen Heyd, Gustavo E. Scuseria, and Matthias Ernzerhof. “Hybrid functionals based on a screened Coulomb potential”. In: *J. Chem. Phys.* 118.18 (May 2003), pp. 8207–8215. ISSN: 0021-9606. DOI: 10.1063/1.1564060.

[45] Michel Lannoo and Jacques Bourgoin. *Point Defects in Semiconductors I*. Berlin, Germany: Springer. ISBN: 978-3-642-81574-4. URL: https://link.springer.com/book/10.1007/978-3-642-81574-4.

[46] Christoph Freysoldt et al. “First-principles calculations for point defects in solids”. In: *Rev. Mod. Phys.* 86.1 (Mar. 2014), pp. 253–305. ISSN: 1539-0756. DOI: 10.1103/RevModPhys.86.253.

[47] Audrius Alkauskas, Matthew D. McCluskey, and Chris G. Van de Walle. “Tutorial: Defects in Semiconductors—Combining Experiment and Theory”. In: *Journal of Applied Physics* 119.18 (May 2016), p. 181101. ISSN: 0021-8979. DOI: 10.1063/1.4948245.

[48] S. B. Zhang and John E. Northrup. “Chemical potential dependence of defect formation energies in GaAs: Application to Ga self-diffusion”. In: *Phys. Rev. Lett.* 67 (17 Oct. 1991), pp. 2339–2342. DOI: 10.1103/PhysRevLett.67.2339. URL: https://link.aps.org/doi/10.1103/PhysRevLett.67.2339.

[49] Chris G. Van De Walle et al. “First-principles calculations of solubilities and doping limits: Li, Na, and N in ZnSe”. In: *Physical Review B* 47.15 (1993), pp. 9425–9434. ISSN: 01631829. DOI: 10.1103/PhysRevB.47.9425.

[50] Sunghyun Kim et al. “Quick-start guide for first-principles modelling of point defects in crystalline materials”. In: *J. Phys.: Energy* 2.3 (July 2020), p. 036001. ISSN: 2515-7655. DOI: 10.1088/2515-7655/aba081.

[51] Christoph Freysoldt, Jörg Neugebauer, and Chris G. Van de Walle. “Fully Ab Initio Finite-Size Corrections for Charged-Defect Supercell Calculations”. In: *Phys. Rev. Lett.* 102.1 (Jan. 2009), p. 016402. DOI: 10.1103/PhysRevLett.102.016402.

[52] Stephan Lany and Alex Zunger. “Accurate prediction of defect properties in density functional supercell calculations”. In: *Model. Simul. Mater. Sci. Eng.* 17.8 (Nov. 2009), p. 084002. ISSN: 0965-0393. DOI: 10.1088/0965-0393/17/8/084002.

[53] Christoph Freysoldt, Jörg Neugebauer, and Chris G. Van de Walle. “Electrostatic interactions between charged defects in supercells”. In: *Phys. Status Solidi B* 248.5 (May 2011), pp. 1067–1076. ISSN: 0370-1972. DOI: 10.1002/pssb.201046289.

[54] Yu Kumagai and Fumiyasu Oba. “Electrostatics-based finite-size corrections for first-principles point defect calculations”. In: *Phys. Rev. B* 89.19 (May 2014), p. 195205. DOI: 10.1103/PhysRevB.89.195205.

[55] Hannu-Pekka Komsa and Alfredo Pasquarello. “Finite-Size Supercell Correction for Charged Defects at Surfaces and Interfaces”. In: *Phys. Rev. Lett.* 110.9 (Feb. 2013), p. 095505. DOI: 10.1103/PhysRevLett.110.095505.

[56] Tomoya Gake et al. “Finite-size corrections for defect-involving vertical transitions in supercell calculations”. In: *Phys. Rev. B* 101.2 (Jan. 2020), p. 020102. DOI: 10.1103/PhysRevB.101.020102.

[57] Stefano Falletta, Julia Wiktor, and Alfredo Pasquarello. “Finite-size corrections of defect energy levels involving ionic polarization”. In: *Phys. Rev. B* 102.4 (July 2020), p. 041115. DOI: 10.1103/PhysRevB.102.041115.

[58] Christoph Freysoldt et al. “First-principles calculations for point defects in solids”. In: *Reviews of Modern Physics* 86.1 (2014), pp. 253–305. ISSN: 15390756. DOI: 10.1103/RevModPhys.86.253.

[59] Christoph Schattauer. “Defects in two-dimensional crystals”. PhD thesis. Wien, Austria: Technische Universität. DOI: 10.34726/hss.2023.106942.

[60] Muhammad Usman Hadi et al. *Large Language Models: A Comprehensive Survey of Its Applications, Challenges, Limitations, and Future Prospects*. Nov. 2023. DOI: 10.36227/techrxiv.23589741. (Visited on 11/05/2024).

[61] Joseph Redmon et al. "You Only Look Once: Unified, Real-Time Object Detection". In: *arXiv* (June 2015). DOI: 10.48550/arXiv.1506.02640. eprint: 1506.02640.

[62] Fung Fung Ting, Yen Jun Tan, and Kok Swee Sim. "Convolutional Neural Network Improvement for Breast Cancer Classification". In: *Expert Systems with Applications* 120 (Apr. 2019), pp. 103–115. ISSN: 09574174. DOI: 10.1016/j.eswa.2018.11.008. (Visited on 11/05/2024).

[63] Arslan Khalid et al. "Breast Cancer Detection and Prevention Using Machine Learning". In: *Diagnostics* 13.19 (Oct. 2023), p. 3113. ISSN: 2075-4418. DOI: 10.3390/diagnostics13193113. (Visited on 11/05/2024).

[64] *Nobel Prize in Physics 2024*. [Online; accessed 18. Feb. 2025]. Feb. 2025. URL: https://www.nobelprize.org/prizes/physics/2024/press-release.

[65] David Dao et al. "daviddao/awful-ai: Awful AI - 2021 Edition". In: *Zenodo* (Jan. 2022). DOI: 10.5281/zenodo.5855972.

[66] Christopher M. Bishop. *Pattern Recognition and Machine Learning*. New York, NY, USA: Springer. ISBN: 978-0-387-31073-2. URL: https://link.springer.com/book/9780387310732.

[67] Simon J.D. Prince. *Understanding Deep Learning*. [Online; accessed 18. Feb. 2025]. June 2024. URL: https://mitpress.mit.edu/9780262048644/understanding-deep-learning.

[68] Kevin P. Murphy. *Probabilistic Machine Learning: An introduction*. MIT Press, 2022. URL: http://probml.github.io/book1.

[69] Michael A. Nielsen. "Neural Networks and Deep Learning". In: *Determination Press* (2015). URL: http://neuralnetworksanddeeplearning.com/chap2.html.

[70] Grant Sanderson. *3Blue1Brown*. Accessed: 2025-02-25. 2025. URL: https://www.3blue1brown.com/.

[71] Raul Rojas and Jerome Feldman. *Neural networks : a systematic introduction*. eng. Berlin, Heidelberg, Germany: Springer, 1996 - 1996. ISBN: 9783642610684.

[72] Warren S. McCulloch and Walter Pitts. "A Logical Calculus of the Ideas Immanent in Nervous Activity". In: *The bulletin of mathematical biophysics* 5.4 (Dec. 1943), pp. 115–133. ISSN: 1522-9602. DOI: 10.1007/BF02478259.

[73] F. Rosenblatt. *The perceptron - A perceiving and recognizing automaton*. Tech. rep. 85-460-1. Ithaca, New York: Cornell Aeronautical Laboratory, Jan. 1957.

[74] Nils J. Nilsson. *The Quest for Artificial Intelligence*. Cambridge University Press, 2009.

[75] Marvin Minsky and Seymour Papert. *Perceptrons : : an introduction to computational geometry /*. eng. Cambridge, Massachusetts ; The MIT Press, 2017 - 1969. ISBN: 9780262343930.

[76] Terrence J. Sejnowski. *The Deep Learning Revolution*. Cambridge, MA, USA: The MIT Press, Oct. 2018. ISBN: 978-0-26234682-5. DOI: 10.7551/mitpress/11474.001.0001.

[77] Andrey Kurenkov. "A Brief History of Neural Nets and Deep Learning". In: *Skynet Today* (2020).

[78] David E Rumelhart, Geoffrey E Hinton, and Ronald J Williams. "Learning representations by back-propagating errors". In: *nature* 323.6088 (1986), pp. 533–536.

[79] Kurt Hornik, Maxwell Stinchcombe, and Halbert White. "Multilayer feedforward networks are universal approximators". In: *Neural Networks* 2.5 (1989), pp. 359–366. ISSN: 0893-6080. DOI: http://dx.doi.org/10.1016/0893-6080(89)90020-8. URL: http://www.sciencedirect.com/science/article/pii/0893608089900208.

[80] Yann LeCun et al. “Handwritten Digit Recognition with a Back-Propagation Network”. In: *Advances in Neural Information Processing Systems* 2 (1989).

[81] Li Deng. “The mnist database of handwritten digit images for machine learning research”. In: *IEEE Signal Processing Magazine* 29.6 (2012), pp. 141–142.

[82] Serkan Kiranyaz et al. “1D Convolutional Neural Networks and Applications: A Survey”. In: *Mechanical Systems and Signal Processing* 151 (Apr. 2021), p. 107398. ISSN: 08883270. DOI: 10.1016/j.ymssp.2020.107398. (Visited on 11/08/2024).

[83] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. “Deep Learning”. In: *Nature* 521.7553 (May 2015), pp. 436–444. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/nature14539. (Visited on 11/08/2024).

[84] Hossein Gholamalinezhad and Hossein Khosravi. “Pooling Methods in Deep Neural Networks, a Review”. In: (2020).

[85] Ashish Vaswani et al. “Attention is all you need”. In: *Advances in neural information processing systems* 30 (2017).

[86] Hafid Khalfoun et al. “B and N codoping effect on electronic transport in carbon nanotubes”. In: *Phys. Rev. B* 81.19 (May 2010), p. 193411. DOI: 10.1103/PhysRevB.81.193411.

[87] H. P. Fried, D. Barragan-Yani, and L. Wirtz. “Electronic and optical properties of native point defects in $CuInS_2$ and $CuGaS_2$”. Submitted to *Phys. Rev. Mater.*, under review. 2025.

[88] S. M. Sze and Kwok K. Ng. *Physics of Semiconductor Devices*. Apr. 2006. ISBN: 978-0-47114323-9. DOI: 10.1002/0470068329.

[89] McCluskey McCluskey. *Dopants and Defects in Semiconductors*. Andover, England, UK: Taylor & Francis, Feb. 2018. ISBN: 978-1-31526908-5. DOI: 10.1201/b21986.

[90] Hans J. Queisser and Eugene E. Haller. “Defects in semiconductors: Some fatal, some vital”. In: *Science* 281.5379 (1998), pp. 945–950. ISSN: 00368075. DOI: 10.1126/science.281.5379.945.

[91] Gang Zhang et al. “Material platforms for defect qubits and single-photon emitters”. In: *Appl. Phys. Rev.* 7.3 (Sept. 2020). DOI: 10.1063/5.0006075.

[92] A Smets et al. *Solar Energy: The Physics and Engineering of Photovoltaic Conversion, Technologies and Systems*. UIT Cambridge, 2016. ISBN: 9781906860325. URL: https://books.google.lu/books?id=vTkdjgEACAAJ.

[93] R Kaigawa et al. “Tandem solar cells with Cu(In,Ga)S_2 top cells on ZnO coated substrates”. In: *Solar Energy Materials and Solar Cells* 94.11 (2010), pp. 1880–1883.

[94] Damilola Adeleye et al. “Wide-Bandgap Cu(In,Ga)S_2 Solar Cell: Mitigation of Composition Segregation in High Ga Films for Better Efficiency”. In: *Small* 21.8 (Feb. 2025), p. 2405221. ISSN: 1613-6810. DOI: 10.1002/smll.202405221.

[95] Filip Tuomisto. *Characterisation and Control of Defects in Semiconductors*. Edison, NJ, USA: Institution of engineering and technology, 2019. ISBN: 978-1-78561-655-6.

[96] J. R. Botha et al. “Defect chemistry in $CuGaS_2$ thin films: A photoluminescence study”. In: *Thin Solid Films* 515.15 (May 2007), pp. 6246–6251. ISSN: 0040-6090. DOI: 10.1016/j.tsf.2006.12.083.

[97] Julien Vidal et al. “Strong interplay between structure and electronic properties in $CuIn(S,Se)_2$: A first-principles study”. In: *Physical Review Letters* 104.5 (2010), pp. 1–4. ISSN: 10797114. DOI: 10.1103/PhysRevLett.104.056401.

[98] Christine L. Bailey et al. “Defect physics of $CuGaS_2$”. In: *Phys. Rev. B* 81.20 (May 2010), p. 205214. DOI: 10.1103/PhysRevB.81.205214.

[99] Johan Pohl and Karsten Albe. “Thermodynamics and kinetics of the copper vacancy in $CuInSe_2$, $CuGaSe_2$, $CuInS_2$, and $CuGaS_2$ from screened-exchange hybrid density functional theory”. In: *J. Appl. Phys.* 108.2 (July 2010), p. 023509. ISSN: 0021-8979. DOI: 10.1063/1.3456161.

[100] Johan Pohl and Karsten Albe. “Intrinsic point defects in $CuInSe_2$ and $CuGaSe_2$ as seen via screened-exchange hybrid density functional theory”. In: *Phys. Rev. B* 87.24 (June 2013), p. 245203. DOI: 10.1103/PhysRevB.87.245203.

[101] L. E. Oikkonen et al. “Vacancies in $CuInSe_2$: new insights from hybrid-functional calculations”. In: *J. Phys.: Condens. Matter* 23.42 (Oct. 2011), p. 422202. ISSN: 0953-8984. DOI: 10.1088/0953-8984/23/42/422202.

[102] L. E. Oikkonen et al. “Redirecting focus in $CuInSe_2$ research towards selenium-related defects”. In: *Phys. Rev. B* 86.16 (Oct. 2012), p. 165115. DOI: 10.1103/PhysRevB.86.165115.

[103] Hui Chen et al. “First-principles study of point defects in solar cell semiconductor $CuInS_2$”. In: *Journal of Applied Physics* 112.8 (2012). ISSN: 00218979. DOI: 10.1063/1.4762001.

[104] J. Bekaert et al. “Native point defects in $CuIn_{1-x}Ga_xSe_2$: hybrid density functional calculations predict the origin of p- and n-type conductivity”. In: *Phys. Chem. Chem. Phys.* 16.40 (Sept. 2014), pp. 22299–22308. ISSN: 1463-9076. DOI: 10.1039/C4CP02870H.

[105] Bing Huang et al. “Origin of Reduced Efficiency in Cu(In,Ga)Se_2”. In: *IEEE J. Photovoltaics* 4.1 (Oct. 2013), pp. 477–482. DOI: 10.1109/JPHOTOV.2013.2285617.

[106] Ye Sheng Yee et al. “Deep recombination centers in $Cu_2ZnSnSe_4$ revealed by screened-exchange hybrid density functional theory”. In: *Phys. Rev. B* 92.19 (Nov. 2015), p. 195201. DOI: 10.1103/PhysRevB.92.195201.

[107] Maria Malitckaya et al. “First-Principles Modeling of Point Defects and Complexes in Thin-Film Solar-Cell Absorber $CuInSe_2$”. In: *Adv. Electron. Mater.* 3.6 (June 2017), p. 1600353. ISSN: 2199-160X. DOI: 10.1002/aelm.201600353.

[108] Alice D.P. Leach et al. “Defect Luminescence from Wurtzite $CuInS_2$ Nanocrystals: Combined Experimental and Theoretical Analysis”. In: *Journal of Physical Chemistry C* 120.9 (2016), pp. 5207–5212. ISSN: 19327455. DOI: 10.1021/acs.jpcc.6b00156.

[109] Pei Yang et al. “Tuning to the Band Gap by Complex Defects Engineering: Insights from Hybrid Functionals Calculations in $CuInS_2$”. In: *J. Phys. D: Appl. Phys.* 51.2 (Dec. 2017). ISSN: 1361-6463. DOI: 10.1088/1361-6463/aa9c17.

[110] Huiwen Xiang et al. “Defect-Complex Engineering to Improve the Optoelectronic Properties of $CuInS_2$ by Phosphorus Incorporation”. In: *Phys. Rev. Appl.* 19.6 (June 2023), p. 064036. DOI: 10.1103/PhysRevApplied.19.064036.

[111] Miaomiao Han et al. “Defect physics in intermediate-band materials: Insights from an optimized hybrid functional”. In: *Phys. Rev. B* 96.16 (Oct. 2017), p. 165204. DOI: 10.1103/PhysRevB.96.165204.

[112] M. M. Han, X. L. Zhang, and Z. Zeng. “Sn doping induced intermediate band in $CuGaS_2$”. In: *RSC Adv.* 6.112 (Nov. 2016), pp. 110511–110516. ISSN: 2046-2069. DOI: 10.1039/C6RA16855H.

[113] Peter Deák et al. “Choosing the correct hybrid for defect calculations: A case study on intrinsic carrier trapping in $\beta - \mathrm{Ga_2O_3}$”. In: *Phys. Rev. B* 95.7 (Feb. 2017), p. 075208. DOI: 10.1103/PhysRevB.95.075208.

[114] B. Tell, J. L. Shay, and H. M. Kasper. “Electrical Properties, Optical Properties, and Band Structure of $CuGaS_2$ and $CuInS_2$”. In: *Physical Review B* 4.8 (1971), pp. 2463–2471.

[115] J L Shay, J H Wernick, and B R Pamplin. *Ternary Chalcopyrite Semiconductors: Growth, Electronic Properties, and Applications: International Series of Monographs in The Science of The Solid State*. International series of monographs in the science of the solid state. Elsevier Science, 2013. ISBN: 9781483157924. URL: https://books.google.lu/books?id=PqU3BQAAQBAJ.

[116] Johan Pohl. “Structure and Properties of Defects in Photovoltaic Absorber Materials: Atomic Scale Computer Simulations of Si and Cu(In, Ga)Se_2”. Available at https://tuprints.ulb.tu-darmstadt.de/3497/. PhD thesis. Darmstadt: University of Darmstadt, Jan. 2013.

[117] Bernard Gil, Didier Felbacq, and Shigefusa F. Chichibu. “Exciton binding energies in chalcopyrite semiconductors”. In: *Phys. Rev. B* 85.7 (Feb. 2012), p. 075205. DOI: 10.1103/PhysRevB.85.075205.

[118] H. P. Fried, D. Barragan-Yani, F. Libisch, and L. Wirtz. “A machine learning approach to predict tight-binding parameters for point defects via the projected density of states”. In: *npj Comput. Mater.* (2025). Accepted for publication.

[119] Audrius Alkauskas et al. *Advanced Calculations for Defects in Materials*. Apr. 2011. ISBN: 978-3-52741024-8. DOI: 10.1002/9783527638529.

[120] Audrius Alkauskas et al. “First-principles theory of the luminescence lineshape for the triplet transition in diamond NV centres”. In: *New Journal of Physics* 16.7 (July 2014), p. 073026. DOI: 10.1088/1367-2630/16/7/073026. URL: https://dx.doi.org/10.1088/1367-2630/16/7/073026.

[121] Audrius Alkauskas et al. “First-Principles Calculations of Luminescence Spectrum Line Shapes for Defects in Semiconductors: The Example of GaN and ZnO”. In: *Phys. Rev. Lett.* 109 (26 Dec. 2012), p. 267401. DOI: 10.1103/PhysRevLett.109.267401. URL: https://link.aps.org/doi/10.1103/PhysRevLett.109.267401.

[122] Lukas Razinkovas et al. “Vibrational and vibronic structure of isolated point defects: The nitrogen-vacancy center in diamond”. In: *Phys. Rev. B* 104 (4 July 2021), p. 045303. DOI: 10.1103/PhysRevB.104.045303. URL: https://link.aps.org/doi/10.1103/PhysRevB.104.045303.

[123] Patrick Rinke et al. “Defect Formation Energies without the Band-Gap Problem: Combining Density-Functional Theory and the GW Approach for the Silicon Self-Interstitial”. In: *Phys. Rev. Lett.* 102 (2 Jan. 2009), p. 026402. DOI: 10.1103/PhysRevLett.102.026402. URL: https://link.aps.org/doi/10.1103/PhysRevLett.102.026402.

[124] Wei Chen and Alfredo Pasquarello. “Accuracy of GW for calculating defect energy levels in solids”. In: *Phys. Rev. B* 96 (2 July 2017), p. 020101. DOI: 10.1103/

PhysRevB.96.020101. URL: https://link.aps.org/doi/10.1103/PhysRevB.96.020101.

[125] Feng Wu et al. “First-principles engineering of charged defects for two-dimensional quantum technologies”. In: *Phys. Rev. Mater.* 1 (7 Dec. 2017), p. 071001. DOI: 10.1103/PhysRevMaterials.1.071001. URL: https://link.aps.org/doi/10.1103/PhysRevMaterials.1.071001.

[126] Tyler J. Smart et al. “Fundamental principles for calculating charged defect ionization energies in ultrathin two-dimensional materials”. In: *Phys. Rev. Mater.* 2 (12 Dec. 2018), p. 124002. DOI: 10.1103/PhysRevMaterials.2.124002. URL: https://link.aps.org/doi/10.1103/PhysRevMaterials.2.124002.

[127] Du Li, Zhen-Fei Liu, and Li Yang. “Accelerating GW Calculations of Point Defects with the Defect-Patched Screening Approximation”. In: *J. Chem. Theory Comput.* 19.24 (Dec. 2023), pp. 9435–9444. ISSN: 1549-9618. DOI: 10.1021/acs.jctc.3c01032.

[128] Ferenc Karsai et al. “F center in lithium fluoride revisited: Comparison of solid-state physics and quantum-chemistry approaches”. In: *Phys. Rev. B* 89.12 (Mar. 2014), p. 125429. ISSN: 2469-9969. DOI: 10.1103/PhysRevB.89.125429.

[129] Paul Tiwald et al. “Ab initio perspective on the Mollwo-Ivey relation for F centers in alkali halides”. In: *Phys. Rev. B* 92.14 (Oct. 2015), p. 144107. ISSN: 2469-9969. DOI: 10.1103/PhysRevB.92.144107.

[130] C. Attaccalite et al. “Coupling of excitons and defect states in boron-nitride nanostructures”. In: *Phys. Rev. B* 83.14 (Apr. 2011), p. 144115. ISSN: 2469-9969. DOI: 10.1103/PhysRevB.83.144115.

[131] Alexander Kirchhoff et al. “Electronic and optical properties of a hexagonal boron nitride monolayer in its pristine form and with point defects from first principles”. In: *Phys. Rev. B* 106.4 (July 2022), p. 045118. ISSN: 2469-9969. DOI: 10.1103/PhysRevB.106.045118.

[132] Zifeng Wang et al. “Graph representation-based machine learning framework for predicting electronic band structures of quantum-confined nanostructures”. In: *Sci. China Mater.* 65.11 (Nov. 2022), pp. 3157–3170. ISSN: 2199-4501. DOI: 10.1007/s40843-022-2103-9.

[133] Xiaoxun Gong et al. “General framework for E(3)-equivariant neural network representation of density functional theory Hamiltonian”. In: *Nat. Commun.* 14.2848 (May 2023), pp. 1–10. ISSN: 2041-1723. DOI: 10.1038/s41467-023-38468-8.

[134] Yang Zhong et al. “Transferable equivariant graph neural networks for the Hamiltonians of molecules and solids”. In: *npj Comput. Mater.* 9.182 (Oct. 2023), pp. 1–13. ISSN: 2057-3960. DOI: 10.1038/s41524-023-01130-4.

[135] Zifeng Wang et al. “Machine Learning Method for Tight-Binding Hamiltonian Parameterization from Ab-Initio Band Structure”. In: *npj Computational Materials* 7.1 (Jan. 2021), p. 11. ISSN: 2057-3960. DOI: 10.1038/s41524-020-00490-5.

[136] M. Nakhaee, S. A. Ketabi, and F. M. Peeters. “Machine Learning Approach to Constructing Tight Binding Models for Solids with Application to BiTeCl”. In: *Journal of Applied Physics* 128.21 (Dec. 2020), p. 215107. ISSN: 0021-8979. DOI: 10.1063/5.0023980.

[137] Daniele Soccodato et al. “Machine learned environment-dependent corrections for a $spds^*$ empirical tight-binding basis”. In: *Mach. Learn.: Sci. Technol.* 5.2 (May 2024), p. 025034. ISSN: 2632-2153. DOI: 10.1088/2632-2153/ad4510.

[138] Qiangqiang Gu et al. “Deep learning tight-binding approach for large-scale electronic simulations at finite temperatures with ab initio accuracy”. In: *Nat. Commun.* 15.6772 (Aug. 2024), pp. 1–12. ISSN: 2041-1723. DOI: 10.1038/s41467-024-51006-4.

[139] Haichen Li et al. “A Density Functional Tight Binding Layer for Deep Learning of Chemical Hamiltonians”. In: *J. Chem. Theory Comput.* 14.11 (Nov. 2018), pp. 5764–5776. ISSN: 1549-9618. DOI: 10.1021/acs.jctc.8b00873.

[140] Chang Liu et al. “Efficient Parameterization of Density Functional Tight-Binding for 5f-Elements: A Th–O Case Study”. In: *J. Chem. Theory Comput.* 20.14 (July 2024), pp. 5923–5936. ISSN: 1549-9618. DOI: 10.1021/acs.jctc.4c00145.

[141] Nicola Marzari et al. “Maximally localized Wannier functions: Theory and applications”. In: *Rev. Mod. Phys.* 84.4 (Oct. 2012), pp. 1419–1475. ISSN: 1539-0756. DOI: 10.1103/RevModPhys.84.1419.

[142] Pietro F. Fontana et al. “Spread-balanced Wannier functions: Robust and automatable orbital localization”. In: *Phys. Rev. B* 104.12 (Sept. 2021), p. 125140. ISSN: 2469-9969. DOI: 10.1103/PhysRevB.104.125140.

[143] Yu. V. Skrypnyk and V. M. Loktev. “Impurity Effects in a Two-Dimensional System with the Dirac Spectrum”. In: *Physical Review B* 73.24 (June 2006), p. 241402. ISSN: 1098-0121, 1550-235X. DOI: 10.1103/PhysRevB.73.241402. (Visited on 04/10/2024).

[144] T. O. Wehling et al. “Local Electronic Signatures of Impurity States in Graphene”. In: *Physical Review B* 75.12 (Mar. 2007), p. 125425. ISSN: 1098-0121, 1550-235X. DOI: 10.1103/PhysRevB.75.125425. (Visited on 04/17/2024).

[145] Aurélien Lherbier et al. “Charge Transport in Chemically Doped 2D Graphene”. In: *Physical Review Letters* 101.3 (July 2008), p. 036808. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.101.036808. (Visited on 04/17/2024).

[146] Shangduan Wu et al. “Average Density of States in Disordered Graphene Systems”. In: *Physical Review B* 77.19 (May 2008), p. 195411. ISSN: 1098-0121, 1550-235X. DOI: 10.1103/PhysRevB.77.195411. (Visited on 04/17/2024).

[147] Vitor M. Pereira, J. M. B. Lopes dos Santos, and A. H. Castro Neto. “Modeling disorder in graphene”. In: *Phys. Rev. B* 77.11 (Mar. 2008), p. 115109. DOI: 10.1103/PhysRevB.77.115109.

[148] A. McSloy et al. “TBMaLT, a flexible toolkit for combining tight-binding and machine learning”. In: *J. Chem. Phys.* 158.3 (Jan. 2023). ISSN: 0021-9606. DOI: 10.1063/5.0132892.

[149] Wenbo Sun et al. “Machine Learning Enhanced DFTB Method for Periodic Systems: Learning from Electronic Density of States”. In: *J. Chem. Theory Comput.* 19.13 (July 2023), pp. 3877–3888. ISSN: 1549-9618. DOI: 10.1021/acs.jctc.3c00152.

[150] L. Brey and C. Tejedor. “New optical transitions in Si-Ge strained superlattices”. In: *Phys. Rev. Lett.* 59.9 (Aug. 1987), pp. 1022–1025. DOI: 10.1103/PhysRevLett.59.1022.

[151] G. Grosso and C. Piermarocchi. “Tight-binding model and interactions scaling laws for silicon and germanium”. In: *Phys. Rev. B* 51.23 (June 1995), pp. 16772–16777. DOI: 10.1103/PhysRevB.51.16772.

[152] Shang Yuan Ren, Xia Chen, and John D. Dow. “Tight-binding sp3d5 Hamiltonian for Si”. In: *J. Phys. Chem. Solids* 59.3 (Mar. 1998), pp. 403–410. ISSN: 0022-3697. DOI: 10.1016/S0022-3697(97)00190-X.

[153] H. J. Kulik et al. “Roadmap on Machine learning in electronic structure”. In: *Electron. Struct.* 4.2 (Aug. 2022), p. 023004. ISSN: 2516-1075. DOI: 10.1088/2516-1075/ac572f.

[154] Irea Mosquera-Lois et al. “Machine-learning structural reconstructions for accelerated point defect calculations”. In: *npj Comput. Mater.* 10.121 (June 2024), pp. 1–9. ISSN: 2057-3960. DOI: 10.1038/s41524-024-01303-9.

[155] Md Habibur Rahman et al. “Accelerating defect predictions in semiconductors using graph neural networks”. In: *APL Mach. Learn.* 2.1 (Mar. 2024). DOI: 10.1063/5.0176333.

[156] Thomas Galvani et al. “Excitons in boron nitride single layer”. In: *Phys. Rev. B* 94.12 (Sept. 2016), p. 125303. DOI: 10.1103/PhysRevB.94.125303.

[157] G. Kresse and J. Hafner. “Ab initio molecular dynamics for liquid metals”. In: *Phys. Rev. B* 47 (1 Jan. 1993), pp. 558–561. DOI: 10.1103/PhysRevB.47.558. URL: https://link.aps.org/doi/10.1103/PhysRevB.47.558.

[158] G. Kresse and J. Furthmüller. “Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set”. In: *Computational Materials Science* 6.1 (1996), pp. 15–50. ISSN: 0927-0256. DOI: https://doi.org/10.1016/0927-0256(96)00008-0. URL: https://www.sciencedirect.com/science/article/pii/0927025696000080.

[159] G. Kresse and J. Furthmüller. “Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set”. In: *Phys. Rev. B* 54 (16 Oct. 1996), pp. 11169–11186. DOI: 10.1103/PhysRevB.54.11169. URL: https://link.aps.org/doi/10.1103/PhysRevB.54.11169.

[160] G. Kresse and D. Joubert. “From ultrasoft pseudopotentials to the projector augmented-wave method”. In: *Phys. Rev. B* 59 (3 Jan. 1999), pp. 1758–1775. DOI: 10.1103/PhysRevB.59.1758. URL: https://link.aps.org/doi/10.1103/PhysRevB.59.1758.

[161] Vei Wang et al. “VASPKIT: A user-friendly interface facilitating high-throughput computing and analysis using VASP code”. In: *Computer Physics Communications* 267 (2021), p. 108033. DOI: https://doi.org/10.1016/j.cpc.2021.108033.

[162] Dean Moldova, Miša Anđelković, and Francois Peeters. “pybinding v0.9.5: a Python package for tight-binding calculations (v0.9.5)”. In: (2020). DOI: https://doi.org/10.5281/zenodo.4010216.