DTU

Technical University of Denmark

Effects of lattice imperfections on the optical and electronic properties of two-dimensional materials

Simone Manti
smanti@fysik.dtu.dk

Supervisor:
Kristian Sommer Thygesen

Center for Atomic-scale Materials Design

Department of Physics

September 2021

Summary

After the first exfoliation of graphene in 2004, the field of two-dimensional (2D) materials has received a lot of research attention due to their unique properties. 2D materials are just one layer thick and they exhibit novel features compare to their bulk counterpart and they have become a platform for discovering exotic phenomena and developing innovative applications.

At the same time, the last few years have seen a renewed interest in studying defects in materials. This happened with a paradigm shift: from an unintended and ubiquitous presence in materials to a source for developing new applications. Therefore, defects are no longer considered to limit the material's performance, but more as a tool to design specific target properties.

In this thesis, from the title *Effects of lattice imperfections on the optical and electronic properties of two-dimensional materials*, the methods to investigate the effects of defects in 2D materials from first principles were developed with the electronic structure code GPAW.

The electronic properties were characterized by describing the symmetry of new defect states that appear in the material's band gap. A symmetry analysis of such states was performed for thousands of defects states, for vacancy and antisite defects in 2D materials. Then, it was also considered the impact of the defects on the optical properties. This was done by studying the optical transitions between defect states and looking at the interaction of the defects with vibrations, an important aspect for application of defects. The focus was also on the nonradiative recombination of photo-excited charge carriers to the defect states, which is the main loss mechanism in opto-electronic devices. The modern state of the art methods to describe such transitions from first principles were implemented with GPAW and within the ASR framework.

A part of the work was also spent in predicting new 2D materials with a high throughput approach. This method was applied to a new class of 2D material, the Janus monolayers. These materials exhibit an intrinsic finite dipole moment that comes from the difference in electronegativity in the chemical elements on the two sides of the layer. A set of new possible monolayers of this class was proposed. Finally, the structural instabilities of 2D materials were also considered. A simple method was used and validated to assess the stability of a monolayer without more expensive calculations, that are not suitable for high throughput computational studies. The effect of the lattice distortion on the electronic properties was studied for a set of distorted monolayers.

Resume

Efter den første eksfoliering af grafen i 2004 har forskningsfeltet to-dimensionale (2D) materialer modtaget megen opmærksomhed på grund af deres unikke egenskaber. 2D-materialer er kun et atomart lag tykke og udviser unikke egenskaber i forhold til deres bulk-modstykke, og de er blevet en platform for opdagelsen af eksotiske fænomener og udvikling af innovative applikationer.

Samtidig har de sidste par år oplevet en fornyet interesse for at studere defekter i materialer. Dette skete med et paradigmeskifte: fra en utilsigtet og allestedsnærværende tilstedeværelse i materialer til en kilde til udvikling af nye applikationer. Derfor anses defekter ikke længere for at begrænse materialets evne, men mere som et værktøj til at designe specifikke egenskaber.

I denne afhandling med titlen *Effects of lattice imperfections on the optical and electronic properties of two-dimensional materials* blev metoder til at undersøge effekterne af defekter i 2D-materialer fra første principper udviklet med elektronisk struktur-koden GPAW.

De elektroniske egenskaber blev karakteriseret ved at beskrive symmetrien i nye defekttilstande, der optræder i materialets båndgab. En symmetri-analyse af sådanne tilstande blev udført for tusinder af defekttilstande, for vakance- og antisite-defekter i 2D -materialer. Derefter blev defekternes indvirkning på de optiske egenskaber også betragtet. Dette blev gjort ved at studere de optiske overgange mellem defekttilstande og ved at se på defekternes interaktion med vibrationer, hvilket er et vigtigt aspekt for anvendelse af defekter. Fokus var også på den ikke-radiative rekombination af foto-eksiterede ladningsbærere til defekttilstandene, som er den vigtigste tabsmekanisme i opto-elektroniske enheder. De moderne metoder til at beskrive sådanne overgange fra de første principper blev implementeret med GPAW inden for ASR-rammen.

En del af arbejdet blev også brugt på at forudsige nye 2D-materialer med en storskala DFT-tilgang. Denne metode blev anvendt på en ny klasse af 2D -materiale; Janus-monolagene. Disse materialer udviser et intrinsisk finit dipolmoment, der kommer fra forskellen i elektronegativitet i de kemiske elementer på de to sider af laget. Et sæt af nye mulige monolag i denne klasse blev foreslået. Endelig blev 2D-materialers strukturelle ustabilitet også overvejet. En simpel metode til at vurdere stabiliteten af et enkeltlag uden tungere beregninger, der ikke er egnede til storskala-undersøgelser, blev brugt og valideret. Effekten af gitterforvrængninger på de elektroniske egenskaber blev undersøgt for et sæt forvrængede monolag.

Preface

This thesis is submitted for the candidacy of PhD degree in Physics from the Technical University of Denmark. The work contained in the thesis was carried out at the Center for Atomic-scale Materials Design (CAMD), Department of Physics in the period from May 2018 to September 2021 under the supervision of Prof. Kristian S. Thygesen.

Kongens Lyngby, September 2021

Simone Manti

Acknowledgements

First of all, I would like to express my deep and sincere gratitude to my supervisor Kristian for allowing me to start the PhD at CAMD and for giving me the possibility to come to Denmark. I would also like to thank him for his patience during the PhD, when I encountered many problems, and also for teaching me to believe more in myself and my capabilities.

Then, I would like to thank Audrius Alkauskas for helping me during and after the external stay, for the interesting discussions, but also for the pleasant experience in Lithuania.

Starting from CAMD, I would like to thank, first, the people I worked more closely: Fabian and Sajid. We shared interesting weekly meetings and they also helped me with the proofreading of this thesis.

Then I would like to thank Bettina, for helping me, even though I have always been late with deadlines. Next, also Jens Jørgen for solving all my problems with GPAW and Ole for the support with Niflheim and the discussions on phonons and defects.

Then chronologically, I would like to thank Thorsten for helping at the beginning of the PhD and for the interesting discussions in the office.

Next, all the people at CAMD, that can be divided in before and after the pandemic, or maybe it is better, before and after moving from 307 to 309. Starting from Sten, Daniele, Thorbjørn, Anders, Estefania, Morten, Asbjørn, Luca, Matthew, Mark (also for the proofreading) and Hadeel. The PhD itself has been really challenging for me, even more during the lockdown period. Then, thanks to Alireza, Nikolaj (for the danish translation), Stéfano, Mads, Sami, Urko, Cuauhtémoc, Ask, Jiban, Sahar, and Joachim.

Outside CAMD, I would like to thank Mohnish for his social support during the lockdown and his "supervision" during the last part of thesis writing. Finally, also Rohitesh, for reintroducing me to the social life during and after the lockdown (Ciao!).

List of Figures

Contents

1 | Introduction and Outline

After Galileo Galilei said that Nature, *the book that lies before our eyes is written in the mathematical language* [1], scientists realized that this language assumes different dialects that depend on what scale it is observed. For example, the description of quantum mechanics required a new mathematical formalism that made it possible to understand the behavior of microscopic particles.

A quote that well represents this complexity that comes from changing the scale is: "More is different" from Philip Warren Anderson [2], which summarizes the fact that a complex system can not be explained by the sum of its parts, but the observed phenomena is an interplay of individual components.

More recently, also within the same scale, different characteristics seem to emerge. In condensed matter, reducing the dimensionality of structures generates materials with totally new properties. The best example in two dimensions is graphene [3], that unlike to its three-dimensional version, i.e., Graphite, exhibits new exotic behavior. Decreasing the dimensionality even more, results to one-dimensional carbon nanotubes [4] and zero-dimensional with quantum dots [5]. Therefore, it seems that by fixing the scale in condensed matter and reducing the dimensionality, one can also say that "less is different".

This observation describes two-dimensional (2D) materials, which are atomically thin materials with novel properties, with reduced screening and enhanced quantum confinement effects. Some examples of this class of materials are transition metal dichalcogenides (TMDs). A well-known TMD, MoS2 [6] has a direct bandgap compared to the indirect in its bulk form. Using state-of-the-art experimental techniques and exploiting the large surface-to-volume ratio, species on the surface can be manipulated to discover new 2D monolayers with interesting characteristics. For instance, if in MoS2 the sulphur atoms on one side of the layer are replaced by selenium atoms, it leads to MoSSe structure [7]. Further in this thesis, these materials are explored in great detail.

Additionally, these 2D materials are not only interesting in their pristine form, their defective counterparts could also have promising applications. Recently, the field of quantum defect design [8] shed light on the positive prospects of defects, which were otherwise considered detrimental for material's performance. For example in bulk material, the negatively charged nitrogen impurity and carbon vacancy in diamond (NV center) [9,10] can be used for application in quantum computations [11] and nanoscale sensing [12]. Another interesting host material for defects for qubits applications is silicon carbide SiC in its different polymorphs [13,14]. In 2D materials, promising color centers [15,16] are found in

1

hexagonal boron nitride [17], that could be used as single-photon emitters. In the family of TMDs, the role of defects in single-photon emission in MoS2 [18,19], MoSe2 [20], WS2 [21],WSe2 [22], is still unclear and open for research.

In this thesis, we have developed methods to characterize the effect of defects in 2D materials from first principles. In particular we explored their impact on the electronic and optical properties for many materials.

The thesis has the following outline:

- **Chapter 2** introduces the different approximations and methods needed to approach the many-body problem. Starting from the Born-Oppenheimer approximation, where the dynamics of electrons and nuclei are decoupled, to address each problem separately. Then density functional theory is introduced, together with all the schemes and methods to practically solve the electronic problem. At the same time, the dynamic of the nuclei is described within the harmonic approximation, where the idea of phonons is naturally introduced. Finally, the methods to study defects from first principles are also presented in the last part of the chapter.

- **Chapter 3** begins by stressing the importance of studying the symmetry of the electronic states, even in the presence of defects, despite the symmetry of the system is usually decreased. Consequently, the states introduced inside the band gap will be analyzed under the symmetry transformations of the point group. During the chapter, we will present the tools and methods for carrying out this symmetry analysis and we will show two practical applications at the end of the chapter.

- **Chapter 4** analyzes the impact of defects on lattice vibrations of the crystal. We will introduce how the phononic modes are affected and how to quantitatively describe the new vibrations in the presence of defects. Consequently, two observable effects of the interaction between defects and phonons will be introduced. The first, in which defects induce nonradiative transitions from the bands of the material to midgap states, without the emission of light. The second, in which the emission of light between two defect states in the gap is modulated by the interaction of defects with the phonons.

- **Chapter 5** introduces the concept of the high-throughput approach in material discovery with first-principle methods. Examples will be presented in relation to the published articles. In particular, how this approach

can propose new interesting materials (paper I), or systemically define the structural instabilities of 2D materials (paper III), or study the impact of defects for many materials (paper IV).

- **Chapter 6** briefly summarizes the results obtained.

2 | The many-body problem

The methods to describe electrons and atoms at the quantum level are addressed in this chapter. This is done by presenting the approximations needed to practically solve the many body problem, to then make the comparison between experiments and first principles calculations possible. The main results and formalism are also introduced, which will be used in the following chapters of this thesis.

2.1 | Born-Oppenheimer approximation

The description of the many-body problem is, at least in principle, obtained from the solution of the Schrödinger equation, once a total Hamiltonian is defined. For a system with electrons (e) and nuclei (N) the most general is:

$$\hat{H}_{tot} = \hat{T}_N + \hat{T}_e + \hat{V}_{N-e} + \hat{V}_{N-N} + \hat{V}_{e-e} \tag{2.1}$$

Which is a sum of the electron and nuclei kinetic energies terms T and all possible interaction terms between: electrons $\hat{V}_{e-e}$, nuclei $\hat{V}_{N-N}$ and electron-nuclei $\hat{V}_{N-e}$.

 This general Hamiltonian can be expressed in terms of spatial coordinates for electrons $\mathbf{r}$ and nuclei $\mathbf{R}$:

$$\hat{H}_{tot} = -\sum_i \frac{\hbar^2 \nabla_i^2}{2m_e} - \sum_I \frac{\hbar^2 \nabla_I^2}{2M_I} - \sum_{i,I} \frac{Z_I e^2}{|\mathbf{r}_i - \mathbf{R}_I|} + \sum_{I<J} \frac{Z_I Z_J e^2}{|\mathbf{R}_I - \mathbf{R}_J|} - \sum_{i<j} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|} \tag{2.2}$$

And solving the many-body problem is equivalent to find the solution of the time-independent Schrödinger equation:

$$\hat{H}_{tot}(\mathbf{r}, \mathbf{R}) \, \Psi(\mathbf{r}, \mathbf{R}) = E \Psi(\mathbf{r}, \mathbf{R}) \tag{2.3}$$

With $\Psi(\mathbf{r}, \mathbf{R})$ the wavefunction of both electrons and nuclei and E their combined energy. This is an intractable problem even for simple molecules. The main idea of the Born-Oppenheimer (BO) approximation [23] is to neglect the $T_N(\mathbf{R})$ term for electrons and then factorize (2.1) in electronic and nuclear parts:

$$\hat{H}_e(\mathbf{r}, \mathbf{R}) = T_e(\mathbf{r}) + V_{N-e}(\mathbf{r}, \mathbf{R}) + V_{e-e}(\mathbf{r})$$
$$\hat{H}_{tot}(\mathbf{r}, \mathbf{R}) = T_N(\mathbf{R}) + V_{N-N}(\mathbf{R}) + \hat{H}_e(\mathbf{r}, \mathbf{R}) \tag{2.4}$$

And solve the electronic problem:

$$\hat{H}_e(\mathbf{r}, \mathbf{R})\, \Psi_n(\mathbf{r}, \mathbf{R}) = E_n(\mathbf{R})\, \Psi_n(\mathbf{r}, \mathbf{R}) \tag{2.5}$$

Where the electronic Hamiltonian $\hat{H}_e(\mathbf{r}, \mathbf{R})$ depends parametrically on the positions $\{\mathbf{R}\}$ of the nuclei. Then the equation (2.3) is approached with an ansatz for the wavefunction that separates the dynamics of electrons and nuclei:

$$\Psi(\mathbf{r}, \mathbf{R}) = \chi(\mathbf{R})\, \Psi_0(\mathbf{r}, \mathbf{R}) \tag{2.6}$$

$\Psi_0(\mathbf{r}, \mathbf{R})$ is the ground state wavefunction of the electronic Hamiltonian, while $\chi(\mathbf{R})$ is the vibronic wavefunction for the nuclei. Then the problem is treated variationally [24], leading to:

$$T_N(\mathbf{R})\,\chi(\mathbf{R}) + E_0(\mathbf{R})\,\chi(\mathbf{R}) + \Lambda(\mathbf{R})\,\chi(\mathbf{R}) = W\chi(\mathbf{R}) \tag{2.7}$$

Where $E_0(\mathbf{R})$ is the lowest adiabatic energy surface and $\Lambda(\mathbf{R})$ is the so called non-adiabatic operator (NAO) [24]:

$$\Lambda(\mathbf{R}) = -\frac{\hbar^2}{2M}\left(\int d\mathbf{r}\; \psi_0^* \frac{\partial \psi_0}{\partial \mathbf{R}} \right) \frac{\partial}{\partial \mathbf{R}} - \frac{\hbar^2}{2M}\left(\int d\mathbf{r}\; \psi_0^* \frac{\partial^2 \psi_0}{\partial^2 \mathbf{R}} \right) \tag{2.8}$$

The NAO is a gauge-dependent term, and in most applications, one can choose a gauge with real wavefunctions and consequently, the first term is zero. Also, the second term can be neglected because it is proportional to the ratio of electron and nucleus masses and therefore, negligible. Hence, the problem for the nuclei reduces to:

$$T_N(\mathbf{R})\,\chi(\mathbf{R}) + E_0(\mathbf{R})\,\chi(\mathbf{R}) = W\chi(\mathbf{R}) \tag{2.9}$$

Where the nuclei feel the potential obtained by parametrically solved for every $\mathbf{R}$ the electronic problem and the equilibrium geometry is given by:

$$\frac{\partial E_0}{\partial \mathbf{R}} = 0 \tag{2.10}$$

A case where the BO approximation is failing is the Jahn-Teller (JT) effect, which is an important effect especially for defects in solids [25, 26].

The Jahn-Teller effect [27] is a spontaneous distortion that lowers the symmetry of the system. It is due to the presence of a degenerate electronic configuration and the effect of the distortion is to remove this degeneracy. Mathematically, this

means that the ansatz (2.6) for the wavefunction should include contributions from n different degenerate energy minima:

$$\Psi(\mathbf{r}, \mathbf{R}) = \sum_k^n \chi_k(\mathbf{R})\, \psi_k(\mathbf{r}, \mathbf{R}) \tag{2.11}$$

With different contributions due to the mixing of degenerate states with the distortion. The new ansatz will lead to a Jahn-Teller Hamiltonian:

$$\hat{H}_{JT} = -\frac{\hbar^2}{2M}\frac{\partial^2}{\partial^2 \mathbf{R}} + \begin{pmatrix} V_{11} & \cdots & V_{1n} \\ \vdots & \ddots & \vdots \\ V_{n1} & \cdots & V_{nn} \end{pmatrix} \tag{2.12}$$

where the mixing terms of the degenerate states with the distortion are:

$$V_{mk} = \int d\mathbf{r}\, \psi_m^*(\mathbf{r}, \mathbf{R})\, \hat{H}_e(\mathbf{r}, \mathbf{R})\, \psi_k(\mathbf{r}, \mathbf{R}) \tag{2.13}$$

2.2 | Density functional theory

Once the many-body problem is decoupled, the electronic problem needs to be solved. Density Functional Theory (DFT) has become the standard method to do this. Within DFT, the main quantity needed is not anymore the wavefunction but the electronic density defined for a system with N electrons as:

$$n(\mathbf{r}) = N \int d\mathbf{r}_2...d\mathbf{r}_N\, \Psi^*(\mathbf{r}, \mathbf{r}_2, ..., \mathbf{r}_N)\, \Psi(\mathbf{r}, \mathbf{r}_2, ..., \mathbf{r}_N) \tag{2.14}$$

This change of perspective from the wavefunction to the density is advantageous because the density $n(\mathbf{r})$ depends only on one variable $\mathbf{r}$ and it is a real quantity unlike the wavefunction, which is complex and depends on the coordinates of all electrons. This is possible because the energy of the system can be written as functional of the density. However, this is not a trivial aspect because the energy is a functional of the wavefunction:

$$E[\Psi] = \frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} \tag{2.15}$$

And the density is a functional of the wavefunction, as it is clear from (2.14). However, writing the energy directly as a functional of the density is not trivial. One of the main results of DFT is that this relation exists between the ground

state density and the energy. This is possible because the functional relation between the ground state wavefunction and the ground state density can be inverted and then energy can be written directly as functional of the ground state density:

$$E[\Psi^{\mathrm{GS}}] = E[\Psi^{\mathrm{GS}}[n^{\mathrm{GS}}]] = E[n^{\mathrm{GS}}] \tag{2.16}$$

This can be proved rigorously [28] by considering two different mappings, studying their uniqueness and hence their invertibility. The ground state wavefunction

$$C : \{v_{ext}\} \longrightarrow \{\Psi^{\mathrm{GS}}\}$$

$$D : \{\Psi^{\mathrm{GS}}\} \longrightarrow \{n^{\mathrm{GS}}\}$$
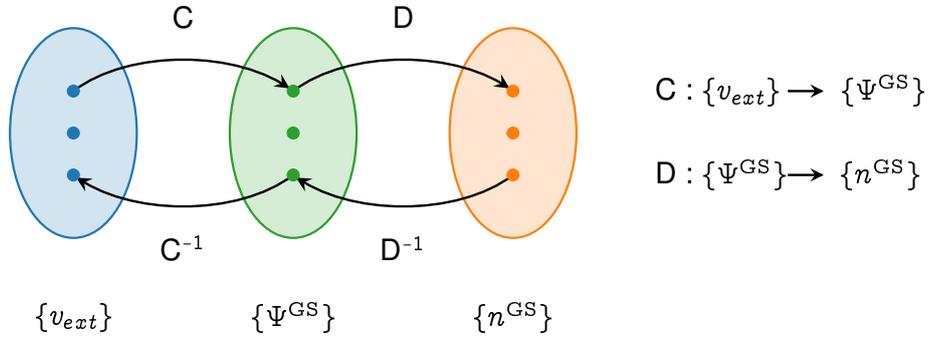
Figure 2.1: The mappings among the external potential, the wavefunction and the density are unique and therefore invertible.

for a particular physical system is uniquely determined by an external potential v_{ext}. Consequently, there are two mappings to consider: one from the external potential to the ground state wavefunction and the other one from the ground state wavefunction to the ground state density (Fig. 2.1). These mappings can be proven to be unique and therefore invertible. This makes writing the wavefunction as functional of the density and the energy as functional of the ground state density possible.

As a result, the energy as a functional of the ground state density exists and another important result is that the energy has a variational property with respect to the ground state density of the system:

$$E[n^{\mathrm{GS}}] \leq E[n] \tag{2.17}$$

This allows the variational access of the theory, starting from a trial density and then calculating the energy extremum.

The other important result of DFT is how the energy as functional of the ground state density $E[n]$ can be written explicitly. The different contributions to the energy functional are:

$$E[n^{\mathrm{GS}}] = T[n^{\mathrm{GS}}] + V^{e-e}[n^{\mathrm{GS}}] + \int d\mathbf{r} \; v^{ext}(\mathbf{r}) \, n^{\mathrm{GS}}(\mathbf{r}) \tag{2.18}$$

With the kinetic energy term and the potential divided in the external and electron-electron parts.

The idea is to introduce a fictitious non-interacting system, called the Kohn-Sham system [29], which has the same ground state density as the interacting system (Fig. 2.2). The density of the two systems are:

$$n_{\mathrm{I}}(\mathbf{r}) = N \int d\mathbf{r}_2 ... d\mathbf{r}_N \; \Psi^*(\mathbf{r}, \mathbf{r}_2, ..., \mathbf{r}_N) \, \Psi(\mathbf{r}, \mathbf{r}_2, ..., \mathbf{r}_N)$$

$$n_{\mathrm{NI}}^{\mathrm{KS}}(\mathbf{r}) = \sum_i^{occ} |\varphi_i^{\mathrm{KS}}(\mathbf{r})|^2 \tag{2.19}$$

And the functions $\{\varphi_i^{\mathrm{KS}}(\mathbf{r})\}$ come from the self-consistent solution of the Kohm-Sham equations:

$$\left[-\frac{\hbar^2 \nabla^2}{2m} + v^{\mathrm{KS}}(\mathbf{r}) \right] \varphi_i^{\mathrm{KS}}(\mathbf{r}) = \varepsilon_i^{\mathrm{KS}} \varphi_i^{\mathrm{KS}}(\mathbf{r}) \tag{2.20}$$

Where $v^{\mathrm{KS}}(\mathbf{r})$ is an effective potential for the fictitious system:

$$v^{\mathrm{KS}}(\mathbf{r}) = v^{\mathrm{H}}(\mathbf{r}) + v^{\mathrm{ext}}(\mathbf{r}) + v^{\mathrm{xc}}(\mathbf{r}) \tag{2.21}$$

And $v^{\mathrm{H}}(\mathbf{r})$ is the functional derivative of the Hartree energy:

$$\mathrm{V}_H[n] = \frac{e^2}{2} \int \frac{n(\mathbf{r}) \, n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' \tag{2.22}$$

Similar functional derivatives hold for the external potential and also for the $v^{\mathrm{xc}}(\mathbf{r})$ that comes from the exchange-correlation functional $E^{\mathrm{xc}}[n]$ which incorporates all the unknown quantities inside the energy functional of the real system:

$$E^{\mathrm{xc}}[n] = T^{\mathrm{I}}[n] - T^{\mathrm{NI}}[n] - V^{\mathrm{H}}[n] + V^{\mathrm{el-el}}[n] \tag{2.23}$$

And the energy functionals of the two systems are:

$$E_{\mathrm{I}}[n] = T^{\mathrm{NI}}[n] + V^{\mathrm{H}}[n] + V^{\mathrm{ext}}[n] + E^{\mathrm{xc}}[n]$$
$$E_{\mathrm{NI}}[n] = T^{\mathrm{NI}}[n] + V^{\mathrm{KS}}[n]$$
$$(2.24)$$

An approximation to $E^{\mathrm{xc}}[n]$ makes it possible to solve the self-consistent Kohm-

Real   Kohn-Sham

$n_{\mathrm{I}}(\mathbf{r})$ $n_{\mathrm{NI}}^{\mathrm{KS}}(\mathbf{r})$
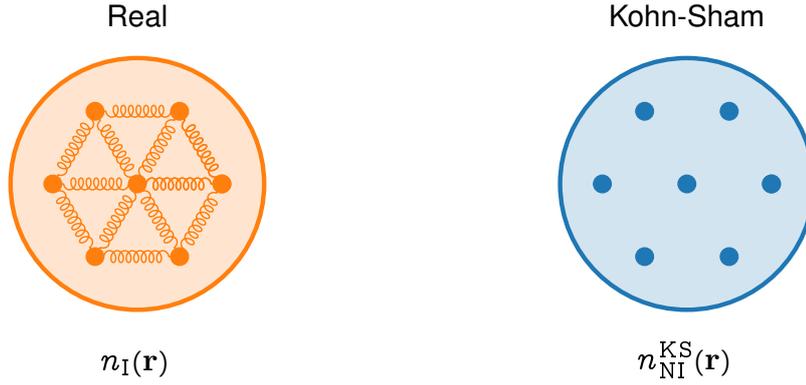
Figure 2.2: The non-interacting Kohn-Sham system is constructed from the real interacting system by ensuring that the electronic density is the same.

Sham equations and to describe the electronic properties of a system, always having in mind to give the right interpretation to the eigenvalues and eigenvectors of (2.20).

The basic approximation to the exchange and correlation functional is the local density approximation (LDA), where the $E^{\mathrm{xc}}[n]$ is split into two contributions. The exchange part is given from the results of the homogeneous electron gas:

$$E^{\mathrm{x}}[n] = -\int \frac{3}{4}\frac{e^2}{\pi}\left(3\pi^2 n(\mathbf{r})\right)^{\frac{4}{3}}d\mathbf{r}$$
$$(2.25)$$

The correlation part is obtained from Quantum Monte Carlo calculations [30]. In spite of the simple form of LDA, it gives good results also for systems different from the homogeneous electron gas. However, it poorly describes systems with an inhomogeneous charge density. A way to improve this is with the generalized gradient approximation (GGA), which adds a gradient dependence in the energy functional. One of the most important functionals within the GGA is the PBE [31].

It is important to mention hybrid functionals, as it will be crucial their use in the case of defects [32]. They are called hybrid because the PBE exchange is mixed with the Hartree-Fock exchange. In the HSE [33,34] functional the degree of mixing between the two different exchanges is determined by the α parameter. This mixing is affecting the short-range (SR) behavior of the two terms, while the long-range (LR) is taken from PBE. The range is defined with an error function, with the parameter ω:

$$E_{xc}^{HSE} = \alpha E_x^{HF,SR}(\omega) + (1-\alpha) E_x^{PBE,SR}(\omega) + E_x^{PBE,LR}(\omega) + E_c^{PBE} \tag{2.26}$$

Typical values for the fraction α is 0.25, while for the screening length ω is 0.11 in the unit of the Bohr radius.

One of the effects of the mixing parameter α for a semiconductor is to enlarge the gap of the material. This can be understood roughly from the fact that Hartree-Fock overestimates the band gap, while with PBE, it is underestimated. Consequently, the gap increases as α is increasing.

All the DFT calculations performed in this thesis are done with electronic structure code GPAW [35] with the PBE functional. The benchmark in section (4.1) was done with the HSE functional.

2.3 | Projector augmented wave formalism

We have seen that DFT allows to systematically solved the electronic problem once an exchange and correlation functional is chosen. Another important choice to practically solve the electronic problem is how to represent the wavefunctions.

In GPAW, three basis sets are available: plane waves (PW), real space grid and localized atomic orbitals (LCAO). Independently from the basis chosen, it is challenging to capture the variation of the wavefunctions near the atoms. The problem is that it requires a very large set of basis functions to expand the states due to the large oscillations near the nuclei.

The approach of the projector augmented wave (PAW) [36] maps this problem in finding a new set of wavefunctions $\tilde{\psi}_n$. These pseudo-wavefunctions are smooth near the nuclei, while they represent the true all-electron wavefunctions ψ_n in the bonding region. In the PAW method therefore, exists a transformation of the type:

$$|\psi_n\rangle = \hat{\mathcal{T}} |\tilde{\psi}_n\rangle \tag{2.27}$$

Which defines a new set of transformed Kohn-Sham equations:

$$\hat{\mathcal{T}}^\dagger \hat{H} \hat{\mathcal{T}} |\tilde{\psi}_n\rangle = \varepsilon_n \hat{\mathcal{T}} |\tilde{\psi}_n\rangle \tag{2.28}$$
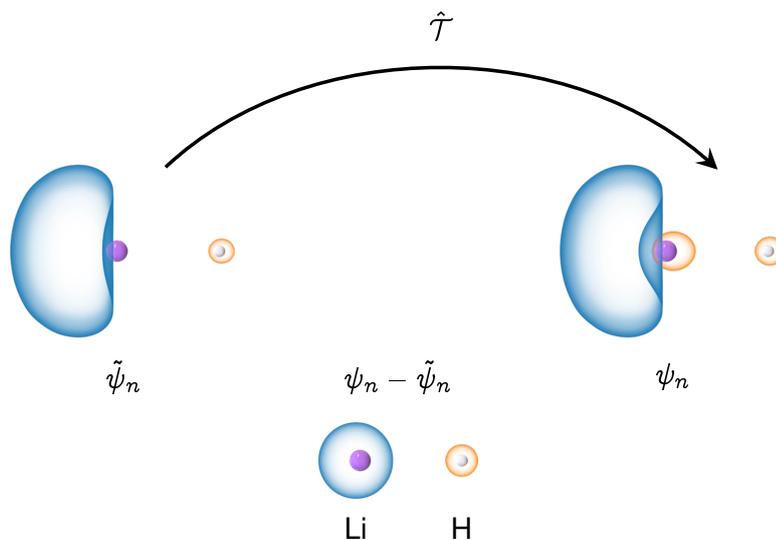
Figure 2.3: The difference from the pseudo (left) and all-electron (right) wavefuc-
ntion are only closer to the nuclei as it is clear from their difference
(below). The PAW transformation $\hat{\mathcal{T}}$ allow to obtain the true electron
wavefunction from the pseudo one, here for a wavefunction for the
HLi molecule.

Consequently, the true all-electron wavefunctions can be obtained from the
pseudo ones after solving the new transformed problem. The pseudo-wavefunctions
are the same as the real all-electron wavefunctions outside augmentation spheres
defined around each nucleus, the only differences are close to the nuclei (Fig.
2.3)[1].
    Together with the PAW method, the last approximation needed is the frozen
core approximation, where the core electrons are considered localized on the
atoms and thus Kohm-Sham problem is defined only for the valence electrons
of the system.

---

[1]The positive (negative) part of the wavefunctions are plotted in orange (blue). This convention
   will be used for all isosurfaces in the thesis.

2.4 | Harmonic approximation and phonons

The solution of the electronic problem gives access to the equilibrium geometry of the system with (2.10). Perturbations from the equilibrium positions can be investigated to study the vibrational properties of the material.

The starting point is the harmonic approximation, which is shortly introduced here and naturally, leads to the concept of phonons. We will follow the derivation and the notation of Ref [37].

Let us consider an infinite crystal made by repetitions of the unit cell with lattice vectors $\mathbf{a}_i$ over the three directions (Fig. 2.4). The position vector for a
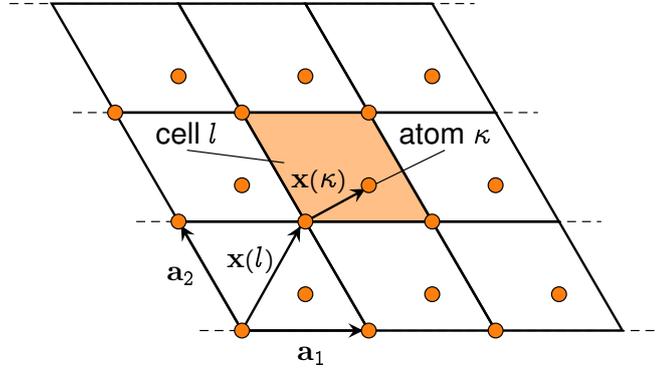
Figure 2.4: The position with respect to the origin of the reference cell l is given by the vector $\mathbf{x}(l)$ and the position of the atom κ in the cell is $\mathbf{x}(\kappa)$ or $\mathbf{x}(l\kappa)$ with respect to the origin.

reference cell l relative to the origin is:

$$\mathbf{x}(l) = l_1\mathbf{a}_1 + l_2\mathbf{a}_2 + l_3\mathbf{a}_3 \tag{2.29}$$

Where the l_i are integer numbers. For a crystal made consisting of N_a atoms in the unit cell, the position of the κ-th atom in the l-th unit cell is:

$$\mathbf{x}(l\kappa) = \mathbf{x}(l) + \mathbf{x}(\kappa) \tag{2.30}$$

If $u_\alpha(l\kappa)$ is the cartesian component α of the displacement from the equilibrium position $\mathbf{x}^{(0)}(lk)$.

$$u_\alpha(l\kappa) = x_\alpha(l\kappa) - x_\alpha^{(0)}(l\kappa) \tag{2.31}$$

The potential energy Φ of the crystal can be expanded in terms of these displacements around the equilibrium configuration and it reads:

$$\Phi = \Phi_0 + \sum_{lk\alpha} \Phi_\alpha(lk)\, u_\alpha(lk) + \frac{1}{2} \sum_{\substack{lk\alpha \\ l'k'\beta}} \Phi_{\alpha\beta}(lk; l'k')\, u_\alpha(lk)\, u_\beta(l'k') + \cdots \quad (2.32)$$

Keeping only terms up to second order in the quadratic term in the Taylor expansion is the Harmonic approximation. Then, the coefficients inside the linear and quadratic terms are the derivatives of the potential evaluated at the equilibrium geometry:

$$\begin{aligned}
\Phi_\alpha(lk) &= \left.\frac{\partial \Phi}{\partial u_\alpha(l\kappa)}\right|_0 \\
\Phi_{\alpha\beta}(lk; l'k') &= \left.\frac{\partial \Phi}{\partial^2 u_\alpha(l\kappa)\, \partial u_\beta(l'\kappa')}\right|_0
\end{aligned} \qquad (2.33)$$

Consequently, $\Phi_\alpha(lk)$ is zero at the equilibrium positions, while the coefficients $\Phi_{\alpha\beta}(l,\kappa; l,\kappa')$ are called the *atomic force constants* and satisfying the following properties:

- $\Phi_{\alpha\beta}(l, \kappa; l', \kappa') = \Phi_{\beta\alpha}(l', \kappa'; l, \kappa)$

- $\displaystyle\sum_{l'\kappa'} \Phi_{\alpha\beta}(l, \kappa; l', \kappa') = 0$

- $\Phi_{\alpha\beta}(l, \kappa; l', \kappa) = \Phi_{\alpha\beta}(l - l', \kappa; 0, \kappa) = \Phi_{\alpha\beta}(0, \kappa; l' - l, \kappa)$

- $\Phi_{\alpha\beta}(L, K; L', K') = \displaystyle\sum_{\mu\nu} D_{\alpha\mu}(R)\, D_{\beta\nu}(R)\, \Phi_{\mu\nu}(l, \kappa; l', \kappa')$

The first property comes from their definition as second-order derivatives, the second from the invariance of the potential under translations, the third from the fact $\Phi_{\alpha\beta}(l, \kappa; l', \kappa)$ depends only from the difference $l - l'$. The last property comes from the invariance of the potential under an operation R of the group of symmetry of the crystal that sends the lattice site $\{l, \kappa; l', \kappa'\}$ to $\{L, K; L', K'\}$ and the matrix elements of the operation are $D_{\alpha\beta}(R)$.

Therefore within the Harmonic approximation, the Hamiltonian for the nuclei can be written with the quadratic terms in (2.32) as a potential term and the equation of motion for the displacement is:

$$M_k \ddot{u}_\alpha(lk) = -\frac{\partial \Phi}{\partial u_\alpha(lk)} = -\sum_{l'k'\beta'} \Phi_{\alpha\beta}(lk; l'k')\, u_\beta(l'k') \qquad (2.34)$$

13

Introducing an ansatz for the displacement:

$$u_\alpha(l\kappa) = \frac{1}{\sqrt{M_\kappa}} u_\alpha(\kappa) e^{-i\omega t + i\mathbf{q}\cdot\mathbf{x}(l)} \tag{2.35}$$

The equation of motion can be rewritten as:

$$\omega^2 u_\alpha(\kappa) = \sum_{\kappa'\beta} \mathcal{D}_{\alpha\beta}(\kappa\kappa'|\mathbf{q})\ u_\beta(\kappa') \tag{2.36}$$

Where the $\mathcal{D}_{\alpha\beta}(\kappa\kappa'|\mathbf{q})$ is the Dynamical matrix:

$$\boxed{\mathcal{D}_{\alpha\beta}(\kappa\kappa'|\mathbf{q}) = \frac{1}{\sqrt{M_\kappa M_{\kappa'}}} \sum_{l'} \Phi_{\alpha\beta}(l,\kappa;l',\kappa') e^{-i\mathbf{q}\cdot(\mathbf{x}(l)-\mathbf{x}(l'))}} \tag{2.37}$$

Which is defined with respect to the reference cell l introduced previously. The $3N_a$ eigenvectors of the Dynamical matrix $e_\alpha(\kappa|\mathbf{q}\lambda)$ satisfy the equation (2.36) with the normal mode or phonon frequencies $\omega_\lambda^2(\mathbf{q})$ of the crystal:

$$\omega_\lambda^2(\mathbf{q}) e_\alpha(\kappa|\mathbf{q}\lambda) = \sum_{\kappa'\beta} \mathcal{D}_{\alpha\beta}(\kappa\kappa'|\mathbf{q}) e_\beta(\kappa'|\mathbf{q}\lambda) \tag{2.38}$$

And they satisfy the following normalization properties:

$$\begin{aligned}
\sum_{\kappa\alpha} e_\alpha^*(\kappa|\mathbf{q}\lambda) e_\alpha(\kappa|\mathbf{q}\lambda') &= \delta_{\lambda\lambda'} \\
\sum_{\lambda} e_\alpha^*(\kappa|\mathbf{q}\lambda) e_\beta(\kappa'|\mathbf{q}\lambda) &= \delta_{\alpha\beta}\delta_{\kappa\kappa'}
\end{aligned} \tag{2.39}$$

With DFT, there are two main approaches to calculate the phonon frequencies of a material. The first is treating the displacement of atoms as a perturbation, with Density Functional Perturbation Theory (DFPT) [38] and look at the variation of the density to build the dynamical matrix. The second approach is the small displacement method [39, 40], where a supercell of the original primitive cell is constructed and the force constant matrix elements are obtained by displacing the atoms in the supercell.

This is done using the fact that the derivative of the potential with respect to the displacement of the atom κ in cell l is equal to the force on that atom:

$$F_\alpha(l\kappa) = -\frac{\partial\Phi}{\partial u_\alpha(l\kappa)} \tag{2.40}$$

Consequently, the force constant matrix elements can be obtained from the derivative of the forces:

$$\Phi_{\alpha\beta}(lk; l'k') = \frac{\partial\Phi}{\partial^2 u_\alpha(l\kappa)\,\partial u_\beta(l'\kappa')} = -\frac{\partial F_\alpha(l\kappa)}{\partial u_\beta(l'\kappa')} \tag{2.41}$$

And they are evaluated with a finite difference, displacing in the positive $(+)$ and negative $(-)$ direction:

$$\frac{\partial F_\alpha(l\kappa)}{\partial u_\beta(l'\kappa')} \approx \frac{F_\alpha^+(l\kappa) - F_\alpha^-(l\kappa)}{2\delta} \tag{2.42}$$

Where δ is the amount of the displacement.

Therefore the $\Phi_{\alpha\beta}(lk; l'k')$ are obtained by displacing the atom κ' in the cell l' in the direction β and calculating the force in the direction α on the atom κ in cell l.

Not all the atoms in the supercell need to be displaced since the force constant matrix in (2.37) is needed only with respect to a reference cell, i.e. the original primitive cell.

For all the phonon calculations in this thesis, the frozen phonon method is used. It was implemented in the in Atomic Simulation Recipes (ASR) [41] in the `asr.phonopy` recipe and using the code PHONOPY [42] to get the displacement of the force constant matrix and to extract the phonon frequencies.

Finally, the interaction between electron and phonon is needed when nonradiative results will be considered in chapter 4. The calculation of the electron-phonon matrix elements is available in GPAW [43], between electronic Bloch states and over all phonon modes, with wavevector $\mathbf{q}$ in the whole phonon Brillouin zone. In this thesis, we need a simplified version which is described here, where we introduce the formula for the matrix element following the approach of [44].

The electron-phonon matrix element between the single-particle states ψ_i and ψ_f, with Hamiltonian $\hat{h}$, and the mode described by the coordinate Q^λ is:

$$W_{if}^\lambda = \langle \psi_i | \frac{\partial\hat{h}}{\partial Q^\lambda} | \psi_f \rangle \tag{2.43}$$

We can drop the phonon branch index λ since in chapter 4 we will only need one particular phonon mode. We now consider that the derivative with respect to Q of the matrix element $\langle\psi_i|\hat{h}|\psi_f\rangle$ is zero, therefore evaluating all the derivatives similarly to the derivation of the Hellman-Feynman theorem [24]:

$$\langle \frac{\partial\psi_i}{\partial Q} | \hat{h} | \psi_f \rangle + \langle \psi_i | \frac{\partial\hat{h}}{\partial Q} | \psi_f \rangle + \langle \psi_i | \hat{h} | \frac{\partial\psi_f}{\partial Q} \rangle = 0 \tag{2.44}$$

15

They can be written as:

$$\langle\psi_i|\frac{\partial\hat{h}}{\partial Q}|\psi_f\rangle = -\langle\psi_i|\hat{h}|\frac{\partial\psi_f}{\partial Q}\rangle - \langle\frac{\partial\psi_i}{\partial Q}|\hat{h}|\psi_f\rangle$$

$$\langle\psi_i|\frac{\partial\hat{h}}{\partial Q}|\psi_f\rangle = -\varepsilon_i\langle\psi_i|\frac{\partial\psi_f}{\partial Q}\rangle - \varepsilon_f\langle\frac{\partial\psi_i}{\partial Q}|\psi_f\rangle$$

(2.45)

Using the same idea again, that derivative of $\langle\psi_i|\psi_f\rangle$ with respect to Q is zero:

$$\langle\frac{\partial\psi_i}{\partial Q}|\psi_f\rangle = -\langle\psi_i|\frac{\partial\psi_f}{\partial Q}\rangle$$

(2.46)

Therefore the electron-phonon matrix element is:

$$\boxed{W_{if} = (\varepsilon_f - \varepsilon_i)\langle\psi_i|\frac{\partial\psi_f}{\partial Q}\rangle}$$

(2.47)

This is the formula that will be used in chapter 4.

2.5 | Defects with density functional theory

Defects with DFT can be modeled by creating a supercell of the material's unit cell and introducing the defect. The starting property to calculate is the formation energy that gives the information on how energetically favorable the formation of a given defect is. The formation energy is usually calculated with the so-called Zhang-Northrup formula [45]:

$$\boxed{E^f[D^q] = E[D^q] - E_0[host] - \sum_i \mu_i n_i + q(\varepsilon_v + E_F) + \Delta_q}$$

(2.48)

The first two terms are the total energies of the supercell with the defect and the pristine supercell of the same size. The third is the energy required to remove ($n_i < 0$) or add ($n_i > 0$) n atoms of species i to form the defect, times its chemical potential μ_i. Another contribution is the energy needed to form the defect in a charge state q, which depends on the position of the Fermi level with respect to the valence band maximum of the pristine material.

The last term Δ_q is a charge correction needed to make the formation energy independent from the supercell size due to the charge q.

In chapter 4, the correction scheme used will be the Freysoldt, Neugebauer and Van de Walle [46], where there are two terms in the correction Δ_q:

$$\Delta_q = -E_l + q\Delta V$$

(2.49)

The first contribution, called the lattice term, removes the electrostatic interaction among the repeated images of the charged distribution density and the interaction with itself, and it is written as:

$$E_{\mathrm{l}} = \frac{2\pi}{\varepsilon \Omega} \sum_{\vec{G} \neq 0} \frac{|n(\vec{G})|^2}{G^2} - \frac{1}{\pi \varepsilon} \int_0^\infty \mathrm{d}g \, |n(g)|^2 \tag{2.50}$$

The second, called the alignment term, takes into account that the charge introduced changes the electrostatic potential, and therefore there is a correction needed in order to align the total energy of the system.

The idea behind the correction scheme is that the difference of the density between the charged defect and the pristine material is localized around the defect (Fig. 2.5). Then the contribution of this repeated localized distribution of charge can be removed and make that formation energy independent of the supercell size.

$$n^{(q)}(\mathbf{r}) - n^{(0)}(\mathbf{r})$$

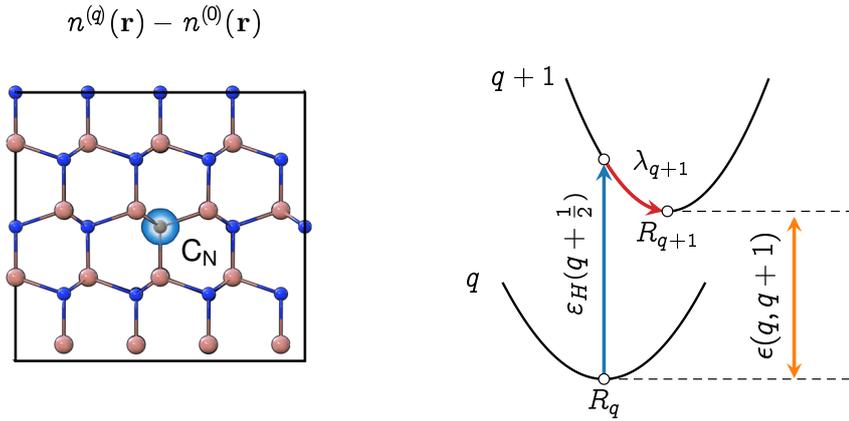Figure 2.5: The density difference (left) between the charged defect and the pristine material is localized on the defect, i.e. the C_N defect in GaN in the center of the supercell. With the SJ approach (right) the CTL can be obtained by the energy of the HOMO level of the $q + \frac{1}{2}$ system and the reorganization energy in the $q + 1$ state.

A standard approximation has become to approximate the density with some

model charge distribution $n_m(r)$ for which the lattice term has an analytical expression. For example, for a Gaussian charge distribution density:

$$n_m^{(q)}(\mathbf{r}) = \frac{q}{[\sqrt{2\pi}\sigma]^3} e^{-\mathbf{r}^2/(2\sigma^2)} \tag{2.51}$$

Which integrates to q and it has a full-width at half maximum (FWHM) of $2\sigma\sqrt{2\ln 2}$. Then the lattice term becomes:

$$E_l = \frac{2\pi}{\varepsilon\Omega} \sum_{\mathbf{G}\neq 0} \frac{q^2 e^{-\mathbf{G}^2\sigma^2}}{\mathbf{G}^2} - \frac{q^2}{2\sqrt{\pi}\varepsilon\sigma} \tag{2.52}$$

With these corrections, the formation energy is independent from the supercell size as long as the density difference is localized around the defect and it can be well described by a gaussian distribution.

For a given material, the formation energies of many defects give the idea of which defects are more likely to be formed. This also depends on the experimental conditions. They can be varied by changing the limit of the chemical potential of the elements to form the defect (rich and poor conditions of the elements).

While formation energies give an idea of how favorable is the formation of defects, the quantity to get closer to the experiment is the charge transition level (CTL). The CTL $\epsilon(q, q')$ gives the Fermi level position needed to change the charge of a defect from a charge state q to q', and it is directly related to their formation energy:

$$\boxed{\epsilon(q, q') = \frac{E_f[D^q; E_F = 0] - E_f[D^{q'}; E_F = 0]}{q' - q}} \tag{2.53}$$

The FNV charge correction scheme is available in GPAW and it was used in chapter 4.

Charge transition levels and formation energies are also calculated in Paper IV for defects in 2D materials, but a different approach was used. Indeed, the FNV scheme should be modified in order to deal with defects in 2D [47] and getting the convergence as a function of the supercell is more challenging. For this reason, since in Paper IV the CTLs are obtained in a high-throughput manner for thousands of defects, the Slater-Janak (SJ) approach was used, which does not rely on total energy considerations and yields reasonable results [48]. The SJ method is based on a basic result of DFT, the Janak's theorem [49]:

$$\frac{\partial E}{\partial n_i} = \varepsilon_i(n_i) \tag{2.54}$$

A relation between the derivative of the total energy with respect to the occupation number n_i and the Kohn-Sham eigenvalue ε_i. It can be used to obtain total energy difference between different numbers of electrons and therefore between different charged states of the defect (Fig. 2.5).

For example, the charge transition level for a defect from the charge state q to the $q + 1$ is given by:

$$\epsilon(q, q+1) = \varepsilon_{\mathrm{H}}(q + \tfrac{1}{2}, R_q) + \lambda_{q+1} \qquad (2.55)$$

Where the $\varepsilon_{\mathrm{H}}(q + \tfrac{1}{2}, R_q)$ is the HOMO level of a system with a $q + \tfrac{1}{2}$ charge at the geometry R_q and λ_{q+1} is the reorganization energy from the R_q geometry with $q + 1$ charge.

The HOMO level of the $q + \tfrac{1}{2}$ system still has some dependence on charge when it is compared to the gap of the pristine material. To avoid this, both HOMO level and the pristine band edges are referenced with respect to the average electrostatic potential for an atom far away from the defect.

3 | Defects and symmetry

The presence of point defects in a material has the effect of lowering its crystal's space group. It is clear that defects break the translational invariance of the lattice, but a further reduction of the symmetry may come from the change of the positions of the atoms surrounding the defect. For this reason, it may seem that characterizing the symmetry in the presence of defects is less crucial compared to the pristine material. Despite this, symmetry still plays an essential role for defects. This happens because defects in a semiconductor may lead to the presence of electronic states with energy that lies in the forbidden gap of the pristine material. The associated wavefunctions of these defect states are spatially localized and a change of their occupation can lead to a significant modification of the electronic density on the defect [50]. This has the effect, for example, of a large ionic relaxation during an electronic excitation and thus an enhancement of the electron-lattice interaction compared to the pristine material.

Therefore, it is crucial to characterize these states under symmetry operations of the point group of the crystal, because changing the occupation of a single state can change the symmetry of the entire material. The reason to do that is twofold: first, the symmetry of the defect states is a fingerprint for a given defect; secondly, all the properties and hence all the non-vanishing matrix elements can be predicted only by symmetry consideration.

This chapter aims to describe the ideas and methods behind the symmetry analysis developed for thousands of defect states in Paper IV and then present some applications.

3.1 | Introduction to group theory

Here we will show some of the basics results of group theory [51] that are needed to classify the symmetry of the Kohn-Sham states in the next section.

Characterizing a function under the symmetry operations of a group is equivalent to decompose the function into different components, where each component transforms in a distinct way. We will clarify this aspect with a well-known and simple example that will help introduce the relevant formalism.

It is an obvious result that starting from a function $f(x)$, it is possible to generate even $f^{(+)}(x)$ or odd $f^{(-)}(x)$ components of the function under reflection along the y axis in the following way:

$$f^{(\pm)}(x) = f(x) \pm f(-x) \tag{3.1}$$

This is trivial, but this is nothing more than an application of a more general method that comes from group theory. To show this, we consider that here in this example, we have a point group with operations $\{R\}$ that are the identity E and the reflection along the y axis, namely σ with the effect $\sigma f(x) = f(-x)$.

If we introduce some unknown quantities for the moment, for every operations and for the even $(+)$ and odd $(-)$ components: $\chi^{(\pm)}(E) = 1$ for the identity and $\chi^{(\pm)}(\sigma) = \pm 1$ for the reflection along the y axis, we can rewrite the even and odd components of the functions as:

$$f^{(\pm)}(x) = \left[\chi^{(\pm)}(E)^* E + \chi^{(\pm)}(\sigma)^* \sigma\right] f(x) = \sum_R \chi^{(\pm)}(R)^* R f(x) \tag{3.2}$$

Where we see that once the $\chi^{(\pm)}(R)$ are known, we have a systematic way of generating even and odd components of the function. The equation above can be seen as the action of a projector operator on the function:

$$f^{(\pm)}(x) = P^{(\pm)} f(x) = \sum_R \chi^{(\pm)}(R)^* R f(x) \tag{3.3}$$

This results can be generalized with group theory to any point group G, where instead of even and odd components, the projector operator projects into the so-called irreducible representations (irreps) [52] of the group, where every irrep transforms in its distinct way, like the even and odd component in the previous example. The numbers $\chi^{(\pm)}(R)$, are called the characters of the group and they are tabulated for every point group [53].

Therefore a projector operator that projects into the irrep α is:

$$\boxed{P^{(\alpha)} = \sum_R \chi^{(\alpha)}(R)^* R} \tag{3.4}$$

To complete the section, we also mention that the characters satisfy an orthogonality relation with the *Little Orthogonality Theorem* (LOT) [54]:

$$\sum_R \chi^{(\alpha)}(R)^* \chi^{(\beta)}(R) = \delta_{\alpha\beta} \tag{3.5}$$

Furthermore, a generalization of the LOT exists in terms of the matrix elements of the representation matrices $D_{kl}^{(\alpha)}(R)$ of the irreps of the group with the *Great Orthogonality Theorem* (GOT):

$$\sum_R D_{ij}^{(\alpha)}(R)^* D_{kl}^{(\beta)}(R) = \frac{N_G}{m_\alpha} \delta_{\alpha\beta} \delta_{ik} \delta_{jl} \tag{3.6}$$

21

Where N_G is the number of the operations in the group and m_α is the degeneracy of the irrep α. The characters $\chi^{(\alpha)}(R)$ introduced before are the trace of the representation matrices:

$$\chi^{(\alpha)}(R) = \text{Tr}(D^{(\alpha)}(R)) = \sum_i D_{ii}^{(\alpha)}(R) \tag{3.7}$$

Finally, from the GOT a generalized projector operator can be also defined, that projects from the component j of the basis of the irrep α to the component i (also called the transfer operator [51]):

$$P_{ij}^{(\alpha)} = \sum_R D_{ij}^{(\alpha)}(R)^* R \tag{3.8}$$

For example, if we want to project inside the basis component i of the irrep α we need:

$$P_{ii}^{(\alpha)} = \sum_R D_{ii}^{(\alpha)}(R)^* R \tag{3.9}$$

And if this is summed over all components of the basis:

$$\sum_i P_{ii}^{(\alpha)} = \sum_R \sum_i D_{ii}^{(\alpha)}(R)^* R = \sum_R \text{Tr}(D^{(\alpha)}(R))^* R = \sum_R \chi^{(\alpha)}(R)^* R \tag{3.10}$$

We obtain again the (3.4) as expected [55]. These are the main results of group theory used in the next sections.

3.2 | Irreducible representations for Kohn-Sham states

After recalling some of the fundamental results of group theory and introducing the projector operator, we can now describe the method to find the irreps of the Kohn-Sham states. This is obtained from the action of the projector operator (3.4) on the Kohm-Sham wavefunction $\Psi(\mathbf{r})$:

$$P^{(\alpha)} \Psi(\mathbf{r}) = \sum_R \chi^{(\alpha)}(R)^* R \Psi(\mathbf{r}) \tag{3.11}$$

When dealing with wavefunctions, it is convenient to work with the overlap $\Gamma(R)$ of the transformed wavefunction under symmetry operation R and itself:

$$\Gamma(R) = \int d\mathbf{r}\ \Psi(\mathbf{r})^* R \Psi(\mathbf{r}) \tag{3.12}$$

And the relation (3.11) becomes an average c_α of the projector operator with the wavefunction:

$$c_\alpha = \langle \Psi | \mathrm{P}^{(\alpha)} | \Psi \rangle = \sum_\mathrm{R} \chi^{(\alpha)}(\mathrm{R})^* \Gamma(\mathrm{R}) \tag{3.13}$$

This can be seen as the fraction of the overlap that transforms like the irrep α. These fractions coefficients c_α are obtained by first inverting the relation (3.13) by multiplying by the character $\chi^{(\beta)}(\mathrm{R})$ and summing over β:

$$\sum_\beta \chi^{(\beta)}(\mathrm{R}) \Gamma(\mathrm{R}) = \sum_\beta \sum_\mathrm{R} \chi^{(\beta)}(\mathrm{R}) \chi^{(\alpha)}(\mathrm{R})^* \Gamma(\mathrm{R}) \tag{3.14}$$

Using the LOT (3.5) and renaming the dummy index we have that the overlap $\Gamma(\mathrm{R})$ is:

$$\Gamma(\mathrm{R}) = \sum_\alpha c_\alpha \chi^{(\alpha)}(\mathrm{R}) \tag{3.15}$$

Where we see that in general the overlap $\Gamma(\mathrm{R})$ can be expanded in terms of the characters of the point group. Therefore for all the N_G operations of the group the overlap are calculated and a linear system of the form $\mathbf{Ax} = \mathbf{B}$ can be constructed:

$$\begin{bmatrix} \Gamma(\mathrm{R}_1) \\ \Gamma(\mathrm{R}_2) \\ \vdots \\ \Gamma(\mathrm{R}_{N_G}) \end{bmatrix} = \begin{bmatrix} \chi^{(\alpha)}(\mathrm{R}_1) & \chi^{(\beta)}(\mathrm{R}_1) & \cdots & \chi^{(\omega)}(\mathrm{R}_1) \\ \chi^{(\alpha)}(\mathrm{R}_2) & \chi^{(\beta)}(\mathrm{R}_2) & \ldots & \chi^{(\omega)}(\mathrm{R}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi^{(\alpha)}(\mathrm{R}_{N_G}) & \chi^{(\beta)}(\mathrm{R}_{N_G}) & \ldots & \chi^{(\omega)}(\mathrm{R}_{N_G}) \end{bmatrix} \begin{bmatrix} c_\alpha \\ c_\beta \\ \vdots \\ c_\omega \end{bmatrix} \tag{3.16}$$

And solved to get the vector $\mathbf{c}$ made from the fraction coefficients c_α.

For example, for the first two wavefunctions of a molecule with $\mathbf{C}_2$ symmetry, the linear system and the wavefunctions are shown in Fig. (3.1) where the operations are just the identity E and the π rotation C_2 around the main axis (perpendicular to the figure).

In this simple example, the irreps can also be found just by looking at the isosurfaces of the wavefunction (Fig. 3.1): Ψ_1 is invariant under all operations and therefore transforms like the total symmetric irrep A; Ψ_2 gets a -1 under the C_2 rotation and consequently, transforms like the irrep B.

A Kohn-Sham state is an eigenvector of the Hamiltonian, hence a base for the point group of the system. This means that if the state Ψ transforms like the component i of the irrep β, then the fraction coefficients c_α are a $\delta_{\alpha\beta}$. This

$$
\begin{bmatrix} \Gamma(E) \\ \Gamma(C_2) \end{bmatrix} = \begin{bmatrix} \chi^A(E) & \chi^B(E) \\ \chi^A(C_2) & \chi^B(C_2) \end{bmatrix} \begin{bmatrix} c_A \\ c_B \end{bmatrix}
$$

$\Psi_2 \sim B$

$$
\begin{bmatrix} \Gamma(E) \\ \Gamma(C_2) \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} c_A \\ c_B \end{bmatrix}
$$

$\Psi_1 \sim A$
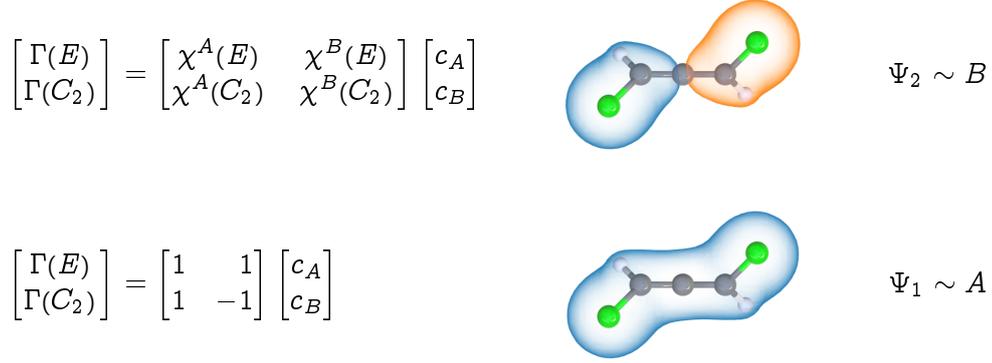
Figure 3.1: The linear system (3.16) and the wavefunctions for a simple molecule (trans-1,2-Dichloroethene), where the C_2 rotation is a π rotation along the axis perpendicular to the figure.

result can be easily proved because if the state transforms like the component i of the irrep β, under a generic symmetry operation we have:

$$
\mathrm{R}\Psi_i^{(\beta)} = \sum_j \mathrm{D}_{ji}^{(\beta)}(R)\,\Psi_j^{(\beta)} \tag{3.17}
$$

This reflects that the state $\Psi_i^{(\beta)}$ is mixed only within the components of the irrep β. If now we act with the projector operator on this state:

$$
\mathrm{P}^{(\alpha)}\,\Psi_i^{(\beta)} = \sum_R \sum_j \chi^{(\alpha)}(\mathrm{R})\,^*\mathrm{D}_{ji}^{(\beta)}(R)\,\Psi_j^{(\beta)} \tag{3.18}
$$

And using the fact the character is the trace $\chi^{(\alpha)}(\mathrm{R}) = \sum_k \mathrm{D}_{kk}^{(\alpha)}(R)$ of the representation matrix, then:

$$
\mathrm{P}^{(\alpha)}\,\Psi_i^{(\beta)} = \sum_R \sum_{jk} \mathrm{D}_{kk}^{(\alpha)}(\mathrm{R})\,^*\mathrm{D}_{ji}^{(\beta)}(R)\,\Psi_j^{(\beta)} = \delta_{ik}\delta_{jk}\delta_{\alpha\beta}\frac{N_G}{m_\alpha}\Psi_i^{(\beta)} \tag{3.19}
$$

Where we have used the GOT (3.6) to obtain deltas on the r.h.s. . Hence the coefficients c_α for a Kohm-Sham state form a vector with all zeros and one in the position of the irrep of the state. Deviations from one are mainly due to small symmetry breaking due to the relaxation of the system. The above method was already available in GPAW [56] for molecules and it was extended and then applied for the symmetry analysis performed in Paper IV.
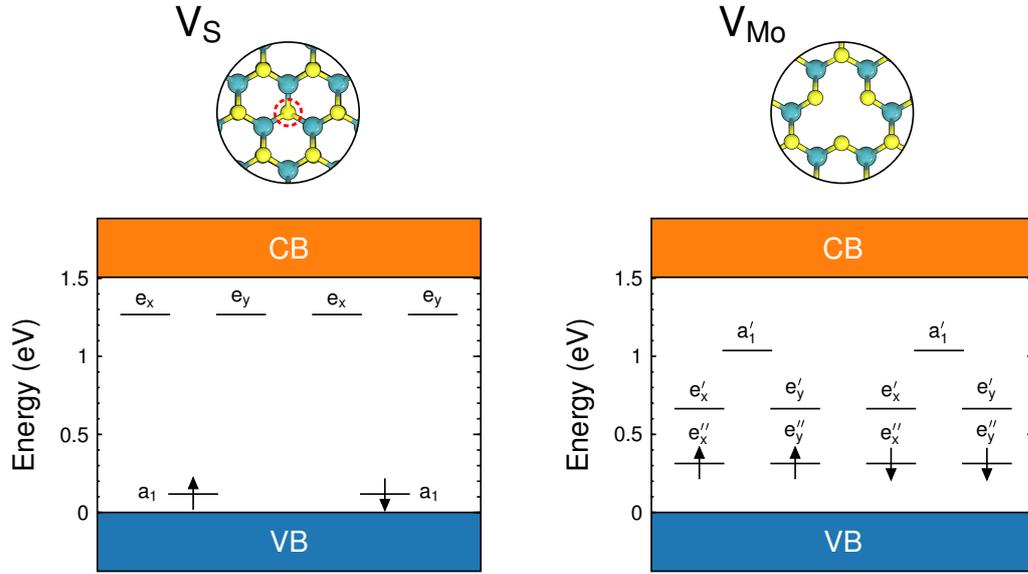
Figure 3.2: Defect states in MoS_2 for the sulphur (yellow) vacancy V_S (left) indicated by the dashed circle and molybdenum (cyan) vacancy V_{Mo} (right) in MoS_2.

As a first application, we will apply this symmetry analysis to the vacancy defects in monolayer MoS_2. A Sulphur vacancy V_S in MoS_2 introduces three defect states in the gap [57,58]: one occupied close to the valence band and two degenerate unoccupied states close to the conduction band. For the Molybdenum vacancy V_{Mo} there are five defect states: two occupied degenerate states, two unoccupied degenerate states and one unoccupied single degenerate state (Fig. 3.2). The point group of V_S is $\mathbf{C}_{3v}$ and the defect states are labeled accordingly to the irreps of the group, while the point group of V_{Mo} is $\mathbf{D}_{3h}$. Qualitatively aspect of the symmetry state can be also seen by looking directly at the wavefunctions. From Fig. (3.3) it is clear how both the total symmetric states a_1 and a_1' of the vacancies are invariant under all operations of the point groups. In contrast, for the V_S the degenerate e states are not, but they transform like in-plane vectors x and y. For V_{Mo} the e' states are invariant respect the in-plane reflection, while the e'' gets a minus one.
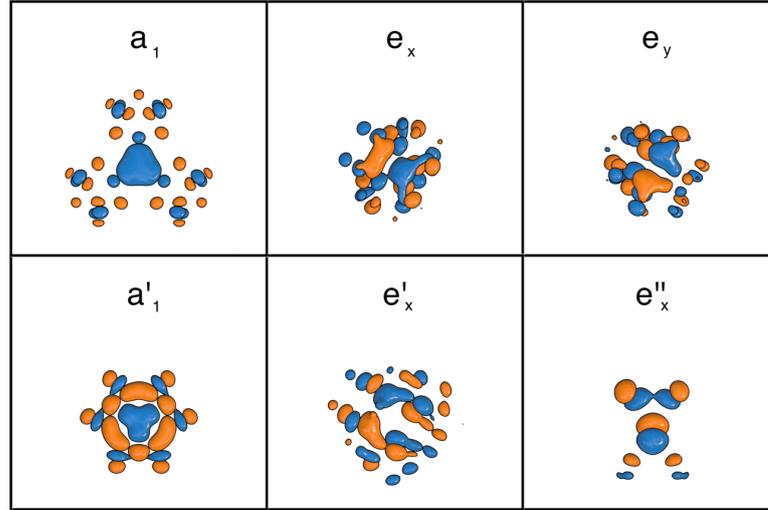
Figure 3.3: Wavefunctions of the defect states for sulphur vacancy V_S (upper panel) and for molybdenum vacancy V_{Mo} (lower panel).

3.3 | Symmetry labels for multiplet states

The importance of performing the symmetry analysis of the defect states is not only essential to predict which matrix elements are different from zero. As introduced at beginning, it is also meaningful to assign the irrep to the multiplet wavefunction made from the occupied defect states.

Obtaining the symmetry properties of the multiplet state is similar to what happens in atomic physics [59], when starting from an electronic configuration, all the allowed different atomic terms are obtained. In that case it is done by summing the angular momenta and spins of the single electrons and finding the possible terms that respect the Pauli exclusion principle, with obvious notation ^{2S+1}L, with L the total angular momentum and S the total spin.

In the same way, the transformation properties of the multiplet wavefunction can be obtained for defects by combining the irrep of the occupied defect states. The notation is somewhat similar to the atomic case $^{2S+1}\Gamma_\alpha$, where S is the total spin of the defect and Γ_α is the irrep of the orbital part of the multiplet.

A well known example is the NV center in diamond [60], with $\mathbf{C}_{3v}$ symmetry and a triplet ground state with $^3A_2$ label and with the first triplet excited state with ^3E. For two-dimensional materials for example in monolayer BN there is the V_N-N_B [61] with a $\mathbf{C}_{2v}$ and a $^2B_2$ ground state.

Here we will show how to obtain these global labels for defects in MoS_2 and for simplicity, we will only consider defects with C_{3v} symmetry. One is the V_S introduced before and the other it is the antisite Mo_S with a triplet ground state with only two degenerate e orbital occupied (Fig. 3.4).

The case of the V_S is trivial since only the single degenerate state is double occupied and hence a spin singlet. The direct product of two a_1 is:

$$a_1 \otimes a_1 = a_1 \tag{3.20}$$

Thus the global label is $^1A_1$ singlet.

The case of the antisite Mo_S with only two degenerate states occupied is more interesting. In this case, the direct product of the irreps of the occupied defect states is:

$$e \otimes e = a_1 \oplus a_2 \oplus e \tag{3.21}$$

And the direct product is not closed on a single irrep of the group. To find which particular irrep corresponds the triplet's orbital part, the projector technique introduced in the first section is needed. This is done by building the product of the occupied states and applying the projector operator. We will indicate the first electron in the e_x as x_1 and the second in e_y as y_2. For example, a rotation of $\pi/3$ has the effect:

$$C_3 x_1 = -\frac{1}{2}x_1 + \frac{\sqrt{3}}{2}y_1 \tag{3.22}$$

For the axes orientation in Fig. (3.4). We need to project this direct product in the irreps of C_{3v}. To do that, we need the projector operator:

$$P^{(\alpha)} = \sum_R \chi^{(\alpha)}(R)^* R \tag{3.23}$$

If now we project the product $x_1 y_2$ over the single degenerate irreps A_1 and A_2 of C_{3v} we have:

$$
\begin{aligned}
P^{(A_1)} x_1 y_2 &= 0 \\
P^{(A_2)} x_1 y_2 &= 3x_1 y_2 - 3x_2 y_1
\end{aligned}
\tag{3.24}
$$

Where we see that the product $x_1 y_2$ is only contained in the A_2 irrep.

To project in the component of the degenerate irrep E, we need a more general projector operator:

$$P_{ii}^{(\alpha)} = \sum_R D_{ii}^{(\alpha)}(R)^* R \tag{3.25}$$

27
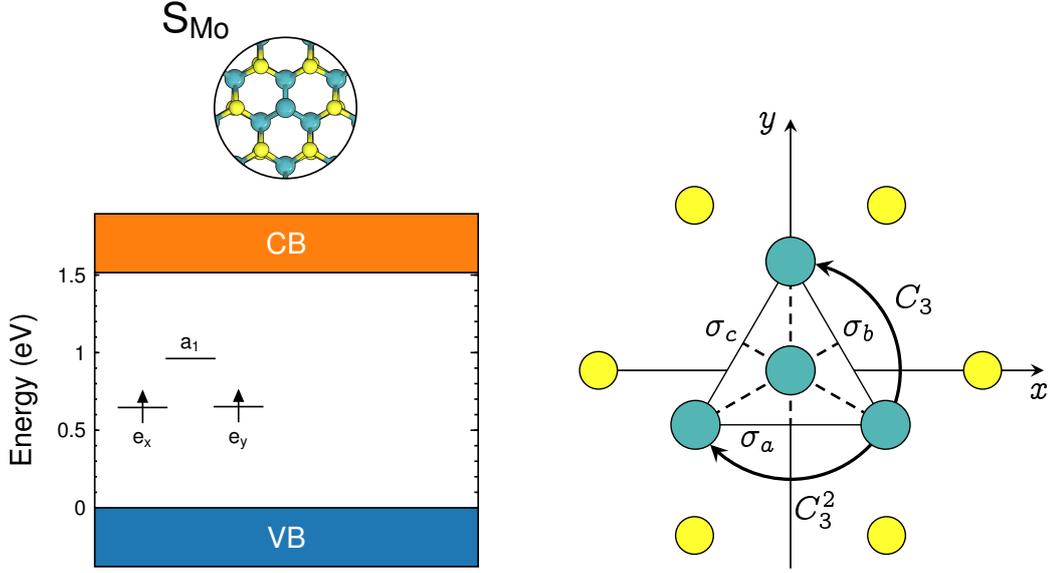
Figure 3.4: The defect states for the triplet ground state of Mo$_S$ (left). The geometry for defining the representation matrices for the antisite defect in $\mathbf{C}_{3v}$ symmetry.

That also projects inside the component of the irrep E, namely x or y in this case. The representation matrices $D(\mathrm{R})$ for the identity, rotations and reflections in $\mathbf{C}_{3v}$, with the orientation of the axes in Fig. (3.4), are:

$$D(E) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad D(C_3) = \begin{bmatrix} -1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & -1/2 \end{bmatrix} \quad D(C_3^2) = \begin{bmatrix} -1/2 & \sqrt{3}/2 \\ -\sqrt{3}/2 & -1/2 \end{bmatrix}$$

(3.26)

$$D(\sigma_a) = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \quad D(\sigma_b) = \begin{bmatrix} 1/2 & \sqrt{3}/2 \\ \sqrt{3}/2 & -1/2 \end{bmatrix} \quad D(\sigma_c) = \begin{bmatrix} 1/2 & -\sqrt{3}/2 \\ -\sqrt{3}/2 & -1/2 \end{bmatrix}$$

And projecting in the x and y component we get:

$$\begin{aligned} \mathrm{P}_{xx}^{(E)} \, x_1 y_2 &= (3x_1 y_2 + 3x_2 y_1)/2 \\ \mathrm{P}_{yy}^{(E)} \, x_1 y_2 &= 0 \end{aligned}$$

(3.27)

Consequently as soon the total wavefunction should be anti-symmetric and the spin part is a triplet and therefore symmetric, the only possible choice is the projection in the irrep A_2. Thus the ground state label for the triplet ground state of Mo$_S$ is $^3A_2$.

3.4 | Spin-orbit and double groups

As a final application of the symmetry analysis, we will discuss the effects of spin-orbit coupling on the defect states of V_S and then we will describe how to characterize the symmetry of the new spin-orbital states.

Spin-orbit coupling is important for the ground-state properties because it can remove the degeneracy of the defect states in the gap. However, it may also affect the dynamics of the defect because it is one of the mechanisms that can flip the spins. For example, spin-orbit coupling is responsible for the transition between multiplet states with different spin components. These transitions are only possible with the interaction of phonons of the system and thus they happen without the emission of light. This effect is called intersystem crossing, and well-known examples are the nonradiative transitions between the triplet to singlet of the NV center in diamond [62].

Consequently, spin-orbit is one of the effects to take into account if the defects states would like to be used for applications like qubit [63, 64].

For example, for the transition metal dichalcogenides (TMDs) this is an even more crucial property that could compromise their possible use as a color centers due to the large spin-orbit coupling. Indeed the heavy transition metal can induce splittings that are of the order of hundreds of meV on the defect states [65].

Here we will consider the case of V_S in MoS_2 as an example to understand how the spin-orbit coupling affects the defect states. The spin-orbit Hamiltonian is [66]:

$$\hat{H}_{SO} = \sum_k \frac{1}{2} \frac{1}{c^2 m_e^2} (\nabla_k V \times \mathbf{p}_k) \cdot \mathbf{s}_k \tag{3.28}$$

In the point group C_{3v} of the V_S it can be written as [60]:

$$\hat{H}_{SO} = \sum_k \lambda_{xy} \left(l_k^x s_k^x + l_k^y s_k^y \right) + \lambda_z \, l_k^z s_k^z \tag{3.29}$$

The $\hat{H}_{SO}$ is evaluated in the Kohm-Sham basis and then diagonalized to extract the new eigenvalues and the new spin-orbital states. The s_k^α are the spin components of the electron k in the direction (θ, ϕ) which defines the polarization axis of the spin. For a polarization along the z-direction, these matrices are just the standard Pauli matrices. To write the $\hat{H}_{SO}$, the Kohm-Sham states are expanded in the different orbital angular momentum projections $l = 0, 1, 2$ and the l_k^α are

the orbital angular momentum matrices that act on every components of the states. For example, for $l = 1$ the matrices are [66]:

$$l_x^{(1)} = \begin{bmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad l_y^{(1)} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{bmatrix} \qquad l_z^{(1)} = \begin{bmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{bmatrix} \qquad (3.30)$$

From the form of the spin and angular momentum matrices, it is possible to understand the effect of the different spin-orbit components in the $\mathbf{C}_{3v}$ symmetry [60]:

- λ_z also called the Axial part of the spin–orbit interaction links states with $m_s = 0$ spin projections among states of the same electronic configuration.

- λ_{xy} called the Non-axial part links states with $m_s \neq 0$ spin projections among different electronic configurations.

For simplicity in the following, we will consider only the Axial part and discuss the splitting of the defect levels due to this perturbation.

When the spin-orbit Hamiltonian is evaluated only considering the λ_z part on the defect state of V_S, the $\hat{H}_{SO}$ matrix is:

$$\hat{H}_{SO}^{(z)} = \begin{array}{c} \\ \langle a_1| \\ \langle \bar{a}_1| \\ \langle e_x| \\ \langle \bar{e}_x| \\ \langle e_y| \\ \langle \bar{e}_y| \end{array} \begin{array}{cccccc} |a_1\rangle & |\bar{a}_1\rangle & |e_x\rangle & |\bar{e}_x\rangle & |e_y\rangle & |\bar{e}_y\rangle \\ \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & i\lambda_z & 0 \\ 0 & 0 & 0 & 0 & 0 & -i\lambda_z \\ 0 & 0 & -i\lambda_z & 0 & 0 & 0 \\ 0 & 0 & 0 & i\lambda_z & 0 & 0 \end{bmatrix} \end{array} \qquad (3.31)$$

Where the bar over the irrep label corresponds to an electron with spin-down component.

We can explain the zeros in this matrix with the symmetry analysis performed previously. Indeed it can be shown [54] that the orbital part of the Non-axial component of $\hat{H}_{SO}$ transforms like the E irrep, while the Axial part transforms as A_2 in the $\mathbf{C}_{3v}$ symmetry.

With group theory the matrix element $\langle \psi^{(\alpha)} | O^{(\gamma)} | \psi^{(\beta)} \rangle$ between two states that transform as the irreps α and β with an operator the transforms as the irrep γ is zero if the irrep α is not contained in the direct product of the irreps $\Gamma_\beta \otimes \Gamma_\gamma$. Alternatively, also if the total symmetric irrep A_1 is not contained in the direct

product of $\Gamma_\alpha \otimes \Gamma_\beta \otimes \Gamma_\gamma$, since the matrix element is a number and therefore invariant under the operations of the group.

This consideration and after looking at the direct product table of $\mathbf{C}_{3v}$ in the appendix, we have that the following matrix elements are zero:

$$\langle a_1 | H_{\mathrm{SO}}^{(A_2)} | a_1 \rangle = \langle a_1 | H_{\mathrm{SO}}^{(A_2)} | e_{x/y} \rangle = 0 \tag{3.32}$$

Because A_1 is not contained in the direct products $A_1 \otimes A_2 \otimes A_1 = A_2$ and $A_1 \otimes A_2 \otimes E = E$. Consequently the only non-vanishing matrix elements are between the degenerate e states. When $H_{\mathrm{SO}}^{(z)}$ is diagonalized there is no effect on the a_1 states, while the double degenerate e states are split by $2\lambda_z$. For $\mathrm{V_S}$ the splitting obtained with GPAW is 50 meV in agreement with other results [65].

The eigenvectors are now spin-orbitals, but it is not possible to label them with the normal irreps of the $\mathbf{C}_{3v}$. The spin part has an additional property, indeed it gets a minus one for 2π rotation. Taking into account this extra symmetry operation leads to the idea of double group [55,67,68], in the following indicated with a bar over the point group, e.g. $\bar{\mathbf{C}}_{3v}$.

Here we will show how to deal with the double group for $\mathrm{V_S}$ and then assign the label of the new irreps to the spin-orbital states. This is important since the as stated earlier for example for TMDs the splitting is so large that the excitation directly happen over the new spin-orbitals states [65].
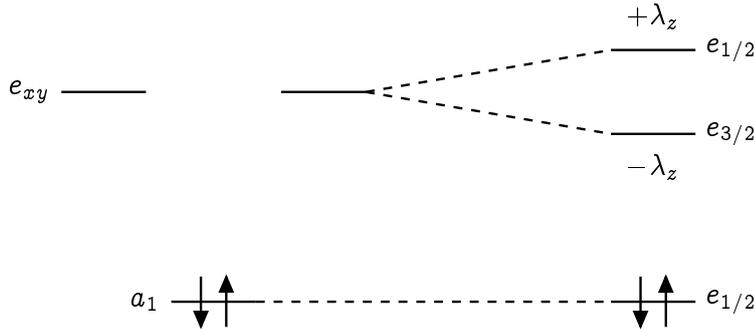

Figure 3.5: For the $\mathrm{V_S}$ in MoS$_2$, the a_1 is not affected by the spin-orbit coupling, while the double degenerate e states are splitted by $2\lambda_z = 50$ meV.

31

The inclusion of the spin adds a new symmetry operation in the group a 2π rotation, indicated with $\bar{E}$, which changes the sing of a spin but leaves the atomic crystal positions and the orbitals unchanged. If there are more symmetry operations, there are more classes and hence more irreps in the group. One of the new extra irreps is the spinorial one, usually called $E_{1/2}$ that describes how spin-1/2 transforms and it is present in every double group. The other irreps can be found just imposing the characters's orthogonality and they are tabulated in the literature for every double group. (The table for $\bar{C}_{3v}$ is in the Appendix). For the C_{3v} case, there are three additional irreps: the spinorial one $E_{1/2}$ and the other two are $E_{3/2}^1$ and $E_{3/2}^2$ that are degenerate when there is time-reversal symmetry [69].

The label for the total symmetric state after the spin-orbit coupling can be obtained by the direct product of the orbital a_1 and the spin $e_{1/2}$:

$$a_1 \otimes e_{1/2} = e_{1/2} \tag{3.33}$$

While for the degenerate e states reads:

$$e \otimes e_{1/2} = e_{1/2} \oplus e_{3/2}^1 \oplus e_{3/2}^2 \tag{3.34}$$

To assign the irrep of the double group we need the projector operator technique again. This time the projector operator is projecting into the irreps of the double group and it requires the characters of C_{3v} and the symmetry operations that act on both the orbital and on the spin part:

$$P^{(\alpha)} = \sum_R \chi^{(\alpha)}(R)^* \, R \, u(R) \tag{3.35}$$

Where we have indicated with $u(R)$ the operation of the point group in the spin representation. The operation R is acting only on the orbital part while $u(R)$ on the spin part.

All the symmetry operations in the spin representation can be obtained starting from the general rotation of an angle ω around the direction $\vec{n} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$

$$u(C_\omega) = \cos(\frac{\omega}{2}) - i(\vec{\sigma}\cdot\vec{n})\sin(\frac{\omega}{2}) = \begin{bmatrix} \cos(\frac{\omega}{2}) - in_z\sin(\frac{\omega}{2}) & -i(n_x - in_y)\sin(\frac{\omega}{2}) \\ -i(n_x + in_y)\sin(\frac{\omega}{2}) & \cos(\frac{\omega}{2}) + in_z\sin(\frac{\omega}{2}) \end{bmatrix} \tag{3.36}$$

32

The reflections can be obtained by taking into account that a reflection can be seen as a rotation of π and then an inversion [55]. For example, if we want to rotate by $\pi/3$ a spin-up electron in the e_x orbital:

$$C_3 u(C_3) \begin{bmatrix} e_x \\ 0 \end{bmatrix} = \begin{bmatrix} (1 - \sqrt{3}i)/2(e_x - \sqrt{3}e_y)/2 \\ 0 \end{bmatrix} \tag{3.37}$$

Having all the symmetry operations, we can construct the projector operator and applying it to the eigenvectors of:

$$\hat{H}_{SO}^{(z)} = \begin{array}{c} \\ \langle e_x| \\ \langle \bar{e}_x| \\ \langle e_y| \\ \langle \bar{e}_y| \end{array} \begin{array}{cccc} |e_x\rangle & |\bar{e}_x\rangle & |e_y\rangle & |\bar{e}_y\rangle \\ \begin{bmatrix} 0 & 0 & i\lambda_z & 0 \\ 0 & 0 & 0 & -i\lambda_z \\ -i\lambda_z & 0 & 0 & 0 \\ 0 & i\lambda_z & 0 & 0 \end{bmatrix} \end{array} \tag{3.38}$$

The eigenvalues are $\pm\lambda_z$ and the eigenvectors in spinorial form are:

$$\begin{aligned} -\lambda_z: && \psi_1 &= \begin{bmatrix} e_y - ie_x \\ 0 \end{bmatrix} & \psi_2 &= \begin{bmatrix} 0 \\ e_y + ie_x \end{bmatrix} \\ +\lambda_z: && \psi_3 &= \begin{bmatrix} e_y + ie_x \\ 0 \end{bmatrix} & \psi_4 &= \begin{bmatrix} 0 \\ e_y - ie_x \end{bmatrix} \end{aligned} \tag{3.39}$$

If we act with (3.35) only on the ψ_1 and ψ_3 because the other states are degenerate and thus they belong to the same irrep, we have for the $E_{1/2}$ irrep that:

$$P^{(E_{1/2})}\psi_1 = 0 \qquad P^{(E_{1/2})}\psi_3 = \begin{bmatrix} e_y + ie_x \\ 0 \end{bmatrix} \tag{3.40}$$

Hence the ψ_3 state with eigenvalue $+\lambda_z$ transforms as the spinorial irrep $E_{1/2}$. While:

$$\begin{aligned} P^{(E_{3/2}^1)}\psi_1 &= \frac{1}{2}\begin{bmatrix} e_y - ie_x \\ e_x - ie_y \end{bmatrix} & P^{(E_{3/2}^1)}\psi_3 &= 0 \\ P^{(E_{3/2}^2)}\psi_1 &= \frac{1}{2}\begin{bmatrix} e_y - ie_x \\ ie_y - e_x \end{bmatrix} & P^{(E_{3/2}^2)}\psi_3 &= 0 \end{aligned} \tag{3.41}$$

And obviously, the other state ψ_1 belongs to the $E_{1/2}^1$ and $E_{1/2}^2$ that are degenerate. We can now label the new spin-orbital states with the irrep of the double group (Fig. 3.5).

33

4 | Defects and phonons

This chapter aims to discuss the combined effects of defects and phonons on the optical properties of the material. As mentioned in the previous chapter, defects, besides of changing the electronic properties of the material, also modify how the electrons interact with the vibrations of the lattice. This has the effect that the emission or absorption from the defect states, is not peaked on the energy difference between the ground and the excited states, but it can be broadened due to the interaction with vibrations of the system [50]. This is an undesirable effect if defects would like to be used as single-photon emitters [70].

The chapter starts by describing from a basic point of view the impact of defects on phonons and generally how the phonon frequencies of the pristine material are perturbed by defects. It is a well-known fact that localized vibrations around the defect can appear, but the situation for a real material is far more complicated [71].

Then the first effect of defects on the optical properties of the material is presented starting from section 4.2, where nonradiative recombinations of photo-excited charge carriers to the defect states are studied. This has been the main project of the PhD, with the intention of studying this mechanism for 2D materials, also in prospective with defects calculated in Paper IV. For this thesis, it is only limited to an implementation in ASR and a benchmark with GPAW for the recombination rate for a bulk defect.

Finally, also the radiative effects of the defects on the emission and absorption properties are studied in section 4.5, and the methods were applied for the most interesting defects in Paper IV.

4.1 | Localized and resonance modes

Here we introduce some general ideas behind the consequences of defects on the phonon frequencies of the material. We will also present a model to treat defects as a perturbation, which allows us to quantitatively define what it is a localized mode and introduce the concept of a resonance mode. We will follow the derivation and the notation in reference [37]. Finally, we will apply the method to the antisite defects in MoS_2 to obtain the perturbed phonon density of states $\delta\rho(\omega)$ from the pristine phonons calculated from first principles, like in Ref [72] for graphene.

In general, defects have two main effects on the dynamics of the pristine material. If we consider a monoatomic crystal with atoms with mass M and

only nearest-neighbor interaction with a force constant C, the effects of defects are: a mass change δM on the masses of the original chemical elements and a change in the atomic force constant δC. A positive (negative) mass change δM has the effect of increasing (decreasing) the phonon frequencies of the material, while a change in force constant C has the opposite effect, increasing (decreasing) of the frequencies for a negative (positive) change of C. These are simple considerations, but this is a starting point to understand the impact of defects on the dynamics of the system. A less trivial fact is that defects can leave all the frequencies unchanged, or they can change them, but by no more than the distance to the next unperturbed frequency. These considerations (Fig. 4.1) and also other more advanced aspects are known as the Rayleigh theorems [37] which can help to understand the physics of phonons with defects as we will show for the case of MoS$_2$.

lighter mass $\delta M < 0$ heavier mass $\delta M > 0$

stronger force $\delta C > 0$ weaker force $\delta C < 0$
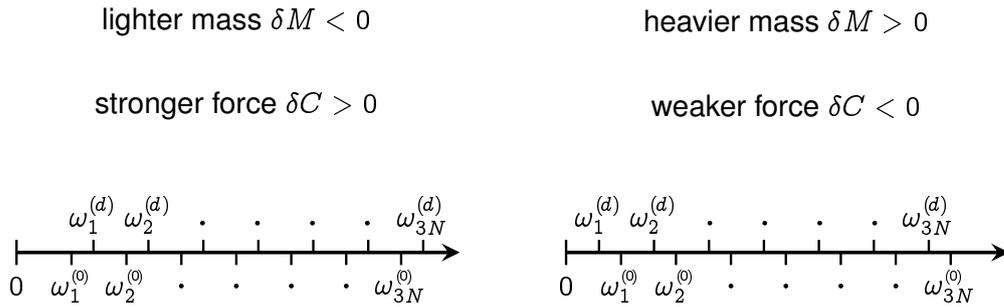
Figure 4.1: A defect with a lighter mass $\delta M < 0$ or a stronger force constant $\delta C > 0$ pushes to frequencies to higher energy. A defect with a heavier mass $\delta M > 0$ or a weaker force constant $\delta C < 0$ pushes to frequencies to lower energy.

Now we will derive the definitions of localized modes and introduce the idea of a resonance modes. The starting point is the equation of motion (EOM) for the displacements (2.34), which is similar to the pristine case, but now we need to take into account the loss of translational symmetry due to defects. If we now take a lattice with a defect, the force constants matrix $\Phi(lk; l'k')$ changes with

35

the effect that it does not depend anymore only on the difference $l - l'$ of the cells. In addition, the masses $M_{l\kappa}$ also depend on the unit cell index l, because there is a defect somewhere in the lattice. The EOM (2.34) now becomes:

$$M_{l\kappa}\ddot{u}_\alpha(lk) = -\sum_{l'k'\beta'} \Phi_{\alpha\beta}(lk; l'k')\, u_\beta(l'k') \tag{4.1}$$

The loss of spatial periodicity due to the defect changes the ansatz for the EOM as:

$$u_\alpha(l\kappa) = u_\alpha(l\kappa)\, e^{-i\omega t} \tag{4.2}$$

Because now there is only the periodicity over time. Substituting the ansatz in the EOM, we have:

$$\sum_{l'k'\beta'} \Big(M_{l\kappa}\omega^2\delta_{ll'}\delta_{kk'}\delta_{\alpha\beta} - \Phi_{\alpha\beta}(l\kappa; l'\kappa') \Big) u_\beta(l'\kappa') = 0 \tag{4.3}$$

If now we define a matrix L with matrix elements:

$$L_{\alpha\beta}(l\kappa; l'\kappa'; \omega^2) = M_\kappa\omega^2\delta_{ll'}\delta_{kk'}\delta_{\alpha\beta} - \Phi_{\alpha\beta}^{(0)}(l\kappa; l'\kappa') \tag{4.4}$$

And define the perturbation due to the defect as:

$$\delta L_{\alpha\beta}(lk; l'k'; \omega^2) = -\omega^2\left(M_{lk} - M_k\right)\delta_{ll'}\delta_{kk'}\delta_{\alpha\beta} + \Phi_{\alpha\beta}(lk; l'k') - \Phi_{\alpha\beta}^{(0)}(lk; l'k') \tag{4.5}$$

Or written in terms of the mass variation $\epsilon = 1 - M_{l\kappa}/M_\kappa$ and the variation of the force constants $\Delta\Phi_{\alpha\beta}(lk; l'k')$ as:

$$\delta L_{\alpha\beta}(lk; l'k'; \omega^2) = \omega^2 M_k\epsilon\,\delta_{ll'}\delta_{kk'}\delta_{\alpha\beta} + \Delta\Phi_{\alpha\beta}(lk; l'k') \tag{4.6}$$

We can rewrite the EOM for the defect (4.3) in matrix form as:

$$(\mathbf{L} - \delta\mathbf{L})\,\mathbf{u} = 0 \tag{4.7}$$

This can be solved if we introduce the inverse of the matrix $\mathbf{L}$ or usually called the Green's function $\mathbf{G}$ which can be shown to have the matrix elements [37]:

$$G_{\alpha\beta}(l, \kappa; l', \kappa'; \omega^2) = \frac{1}{(M_\kappa M_{\kappa'})^{\frac{1}{2}}} \sum_{\mathbf{q}j} \frac{e_\alpha(\kappa|\mathbf{q}j)\, e_\beta^*(\kappa'|\mathbf{q}j)}{\omega^2 - \omega_j(\mathbf{q})^2}\, e^{i\mathbf{q}\cdot(\mathbf{x}(l) - \mathbf{x}(l'))} \tag{4.8}$$

Where the $e_\alpha(\kappa|\mathbf{q}j)$ are the eigenvectors of the dynamical matrix. Multiplying (4.7) from left by $\mathbf{G}$ and using the definition of the Green's function $\mathbf{G} = \mathbf{L}^{-1}$:

$$\mathbf{G}(\mathbf{L} - \delta\mathbf{L})\,\mathbf{u} = \mathbf{u} - \mathbf{G}\delta\mathbf{L}\mathbf{u} = 0 \tag{4.9}$$

We get:

$$\mathbf{u} = \mathbf{G}\delta\mathbf{L}\,\mathbf{u} \tag{4.10}$$

The matrix $\delta\mathbf{L}$ is a sparse matrix as it can be seen from its definition (4.6), it has nonzero elements only for indices (l, κ) close to the defect site. Therefore we can partitionate the matrix as:

$$\delta\mathbf{L} = \left[\begin{array}{c:c} \delta\mathbf{l} & 0 \\ \hdashline 0 & 0 \end{array} \right] \tag{4.11}$$

The same holds for the Green's function $\mathbf{G}$ and the displacement vector $\mathbf{u}$:

$$\mathbf{G} = \left[\begin{array}{c:c} \mathbf{g} & \mathbf{G}_{12} \\ \hdashline \mathbf{G}_{21} & \mathbf{G}_{22} \end{array} \right] \qquad \mathbf{u} = \left[\begin{array}{c} \mathbf{u}_1 \\ \mathbf{u}_2 \end{array} \right] \tag{4.12}$$

Where the global index 1 refers to indices close to the defect and 2 the remaining indices. Therefore (4.10) becomes:

$$\begin{aligned} \mathbf{u}_1 &= \mathbf{g}\,\delta\mathbf{l}\,\mathbf{u}_1 \\ \mathbf{u}_2 &= \mathbf{G}_{21}\,\delta\mathbf{l}\,\mathbf{u}_1 \end{aligned} \tag{4.13}$$

The first is a closed equation for the displacements that are directly affected by the presence of the defect, and the displacements of the remaining atoms are obtained from the second equation.

Solving the first equation in the defect subspace is equivalent to set the following determinant equals to zero:

$$\Delta(\omega^2) = |\mathbb{1} - \mathbf{g}\delta\mathbf{l}| = 0 \tag{4.14}$$

Usually, $\Delta(\omega^2)$ is called the *defect determinant* [73], and its zeros give the frequencies of the modes that are perturbed by the presence of the defect.

In an equivalent way, solving the equation for the displacement in the defect subspace means finding the eigenvalues of the matrix $\delta\mathbf{l}\mathbf{g}$:

$$\mathbf{g}\delta\mathbf{l}\,\psi_n = \lambda_n\psi_n \tag{4.15}$$

And setting the eigenvalues $\lambda_n = 1$. The real and imaginary parts of the eigenvalues λ_n give the properties of the new modes introduced by defects:

• Localized mode: $\operatorname{Re}\lambda_n = 1$ $\operatorname{Im}\lambda_n = 0$

- Resonance mode: $\text{Re}\,\lambda_n = 1 \quad \text{Im}\,\lambda_n \neq 0$

This gives a precise definition of localized modes and they are characterized by the fact the vibration amplitude is localized on the defect. Another fact is that the $\text{Im}\,\lambda_n = 0$ only when the imaginary part of the Green's function is zero and therefore when the pristine phonon density of states is zero $\rho_0(\omega) = 0$ [37]. Consequently, localized modes can appear only above the maximum frequency or in the gap of the phonon spectrum of the pristine material. Resonance modes, instead that have a finite imaginary part and they appear in the region of the spectrum where the pristine density of states is different from zero.

Now from the eigenvalues λ_n, we can obtain first the variation of the number of modes between the defect and the pristine system, with the formula [74]:

$$\delta n(\omega) = \frac{2\omega}{3r\pi} \sum_n \tan^{-1} \frac{-\text{Im}\lambda_n(\omega^2)}{1 - \text{Re}\lambda_n(\omega^2)} \tag{4.16}$$

Its derivative is the variation of the phonon density of states:

$$\delta\rho(\omega) = \frac{d\delta n(\omega)}{d\omega} \tag{4.17}$$

In the following we will apply the above formalism for antisite defects in MoS$_2$, namely Mo$_S$ and S$_{Mo}$, without including the variation in the force constant matrix $\triangle\Phi_{\alpha\beta}(lk;l'k') = 0$. We will show the formulas for the Mo$_S$ defect but for the other antisite is the same, just inverting the mass of sulfur and molybdenum. The perturbation matrix (4.6) is a diagonal 3x3 matrix:

$$\delta\mathfrak{l}(\omega^2) = \begin{bmatrix} \omega^2 M_S\epsilon & 0 & 0 \\ 0 & \omega^2 M_S\epsilon & 0 \\ 0 & 0 & \omega^2 M_S\epsilon \end{bmatrix} \tag{4.18}$$

Where $\epsilon = 1 - M_{Mo}/M_S$. The Green's function is also diagonal because we are interested in the contribution within the same cell and due to the orthogonality of the eigenvectors, the off-diagonal terms are zero. Consequently, the product of the perturbation matrix with the Green's function is:

$$\delta\mathfrak{l}\,\mathbf{g}(\omega^2) = \begin{bmatrix} \omega^2 M_S\epsilon g_{xx}(S;\omega^2) & 0 & 0 \\ 0 & \omega^2 M_S\epsilon g_{yy}(S;\omega^2) & 0 \\ 0 & 0 & \omega^2 M_S\epsilon g_{zz}(S;\omega^2) \end{bmatrix} \tag{4.19}$$

Where the $g_{\alpha\alpha}(S;\omega^2)$ refers to the Green's function with both atom indices for the sulfur atom. Therefore since the matrix is diagonal, the three eigenvalues for the three cartesian components are:

$$\lambda_\alpha = \epsilon M_S \omega^2 g_{\alpha\alpha}(S;\omega^2) \tag{4.20}$$

In Fig. (4.2) we show the variations $\delta n(\omega)$ obtained by the formula (4.16) and variation of the phonon density of state $\delta\rho(\omega)$. We show also the pristine phonon density of states ρ_0 in gray. The total density of state is obtained by the sum [75]:

$$\rho(\omega) = \rho_0(\omega) + c_d \delta\rho(\omega) \tag{4.21}$$

Where c_d is the concentration of defects taken as 0.1% for both defects which a typical value for antisites [57]. All the densities of states are rescaled in order to plot them together with $\delta n(\omega)$. The frequencies of the localized modes in the phononic gap and above the maximum frequency are checked against the frequencies obtained by a phonons pristine supercell calculation and replacing the mass of an atom in the dynamical matrix.

When the heavier molybdenum atom is substituting sulfur, all the frequencies are decreasing and therefore, $\delta n(\omega)$ is always positive and three modes from the bottom of the optical branch go in the phononic gap. From the relative height of the variation of density, we can also predict the degeneracy of the modes. The two double degenerate modes are the λ_x and λ_y, when $\delta n(\omega) = 2$ while the single degenerate is λ_z.

For the other case of sulfur substituting molybdenum, all the frequencies are pushed to higher energy and the mode variation $\delta n(\omega)$ is always negative. Two acoustic modes go in the gap and three modes go above the maximum frequency of the unperturbed spectrum. Here it is possible to see also a small resonance mode around 24 meV.

Of course, this is a very crude model, and a more realistic calculation should include a variation of the force constants matrix by displacing only atoms around the defect. In principle, this could save a lot of computational time compared to a normal phonons calculation, like in the method of embedding the force constants matrix [71] or machine learning based method [76].
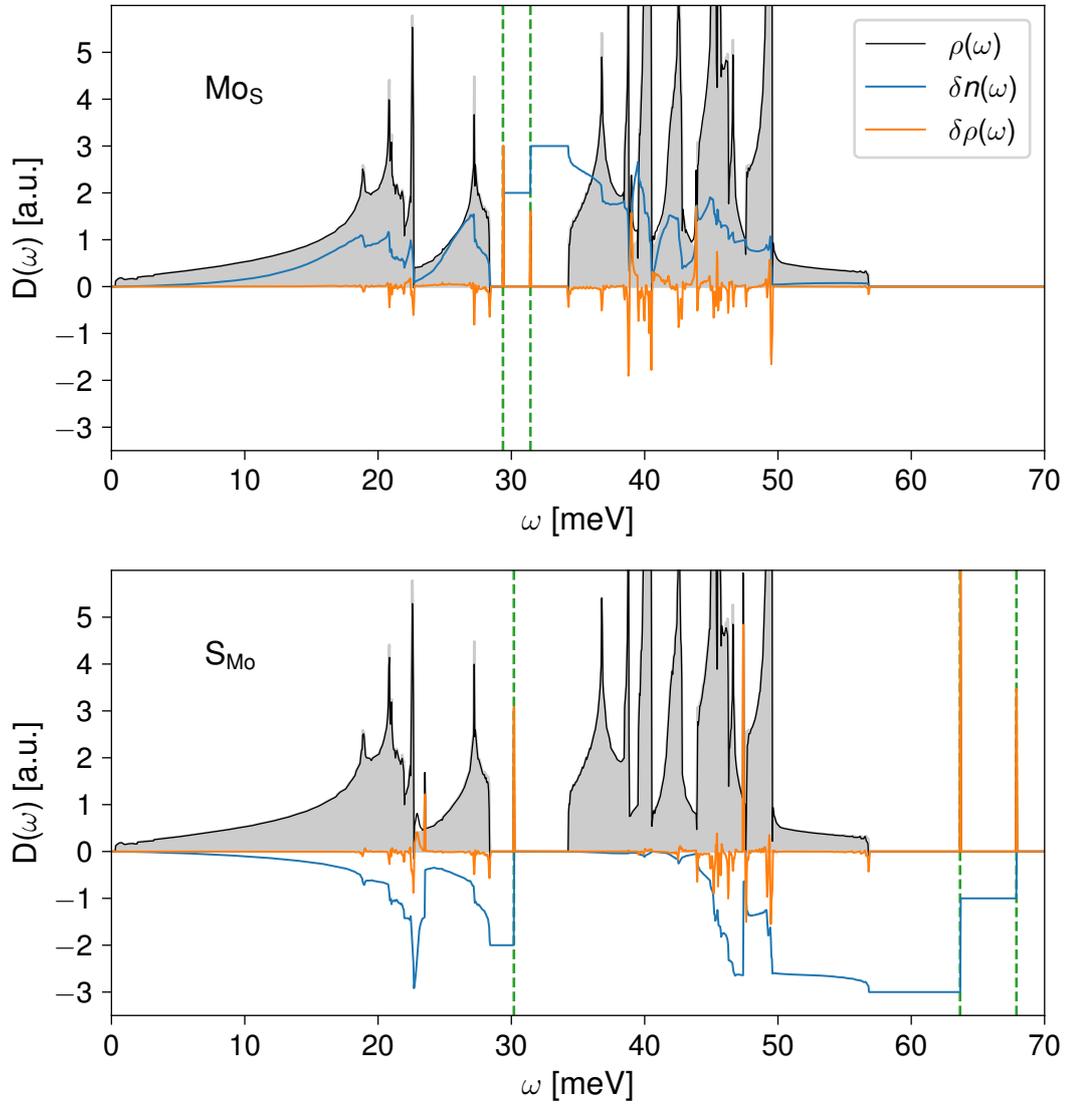
Figure 4.2: For the heavier Mo$_S$ defect the frequencies are decreasing and three localized mode go in the gap from the bottom of the optical branch (upper panel). While for the lighter S$_{Mo}$ defect the frequencies are increasing and two modes go in the gap from the acoustic branch and three mode go above the top of the optical branch (lower panel).

4.2 | **Nonradiative recombination at defects states**

The presence of electronic states in the band gap due to defects can affect the performance of a material [77]. In particular, for the transport properties, charged photo-excited carriers can recombine at defect states via nonradiative recombination, without emitting light but with the emission of phonons. This is known as Shockley-Read-Hall recombination [78] and it is the most important loss mechanism for application in opto-electronic devices.

Nonradiative recombination is a process where an electron (hole) in the conduction (valence) band recombines at defect state in the gap (Fig. 4.3).

Experimentally this phenomenon is well studied and the observable quantity is the time it takes for the carrier to recombine on the midgap state. The time is measured with the so-called capacitance spectroscopy technique [79], where the material is excited by a pulse of light to create the charge carrier and the capacitance of the material is measured over time. Taking into account the carrier on the defect states are localized and they do not participate to the conduction of the material, the time it takes for a carrier to recombine can be measured [80].

The inverse of this recombination time is the nonradiative rate, which links experiments with ab initio simulations. The rates for electrons R_n and for holes R_p are:

$$R_n = C_n\, n\, N_D$$
$$R_p = C_p\, p\, N_A$$

(4.22)

Where n (p) is density of electron (hole) in the conduction (valence) band, N_D (N_A) is density of donor (acceptor) defects and C_n (C_p) is the electron (hole) capture coefficient. The capture coefficient for a carrier C is a quantity that links experiments with ab initio simulations. A standard approach to calculate the capture coefficient is Fermi's golden rule:

$$C = V \frac{2\pi}{\hbar} g \sum_m w_m \sum_n |\Delta H^{elph}_{im;jn}|^2 \, \delta(E_{in} - E_{jm})$$

(4.23)

Where the perturbation that drives the transition is the electron-phonon interaction. A certain number of phonons is emitted during the transition, giving the process its name: multi phonon emission [81, 82].

In the Fermi's golden rule: V is the volume of the supercell, w_m is the thermal occupation of the vibrational state m of the initial state, E_{im} and E_{jm} are the

energies of the initial and the final states and g is the degeneracy factor of the final state.

A first approximation is needed to go further, which is the static approximation [83], that makes it possible to decouple the electron and phonon part of the perturbation, like in the Born-Oppenheimer approximation. Under this assumption, the perturbation can be expanded around the configuration Q_0 with only keeping the linear term:

$$\Delta\hat{H}^{el-ph}_{im;fn} = \sum_\lambda \underbrace{\langle \Psi_i | \partial\hat{H}/\partial Q_\lambda | \Psi_f \rangle}_{W^\lambda_{ij}} \langle \chi_{im} | Q_\lambda - Q_{0;\lambda} | \chi_{fn} \rangle \tag{4.24}$$

Where $\Psi_{i/f}$ are the electronic wavefunctions of the initial and final state, and $\chi_{im/fn}$ the ionic wavefunctions. The sum runs over all phonon modes Q_λ , and $Q_{0;\lambda}$ is the projection of the initial atomic configuration $\{Q_0\}$ along each of the phonon coordinates. W^λ_{ij} is the electron-phonon coupling matrix element to the phonon mode λ.

This is already posing a challenge because defects are simulated in a supercell with N atoms, and the calculation of the electron-phonon matrix elements requires 6N displacements (plus and minus for the three directions of the N atoms). There are attempts in the literature [84–86] where all the matrix elements are obtained at PBE level and then (4.24) is evaluated with the methods in Ref. [87]. Unfortunately, hybrid functionals are required in order to give a reasonable description of defects. This makes it impossible to calculate all of the matrix elements within this model.

A method that avoids the phonon calculation is the one-dimensional (1D) approach [44, 88–90], where an effective 1D mode is built from the initial and final geometries of the transition and then all the input parameters for the formula of the capture coefficient are converted in terms of the 1D mode. Here we will benchmark the 1D approach for the C_N defect in GaN. We will follow the notation in [44].

The atomic (α) and cartesian (t) components of the one-dimensional mode are:

$$Q_{\alpha,t} = m_\alpha^{1/2} \Delta R_{\alpha,t} = m_\alpha^{1/2}(R^{(f)}_{\alpha,t} - R^{(i)}_{\alpha,t}) \tag{4.25}$$

where $R^{(i)}_{\alpha,t}$ and $R^{(f)}_{\alpha,t}$ are the initial and final geometries of the transition. The 1D mode has magnitude:

$$\Delta Q^2 = \sum_{\alpha,t} m_\alpha \Delta R^2_{\alpha,t} \tag{4.26}$$
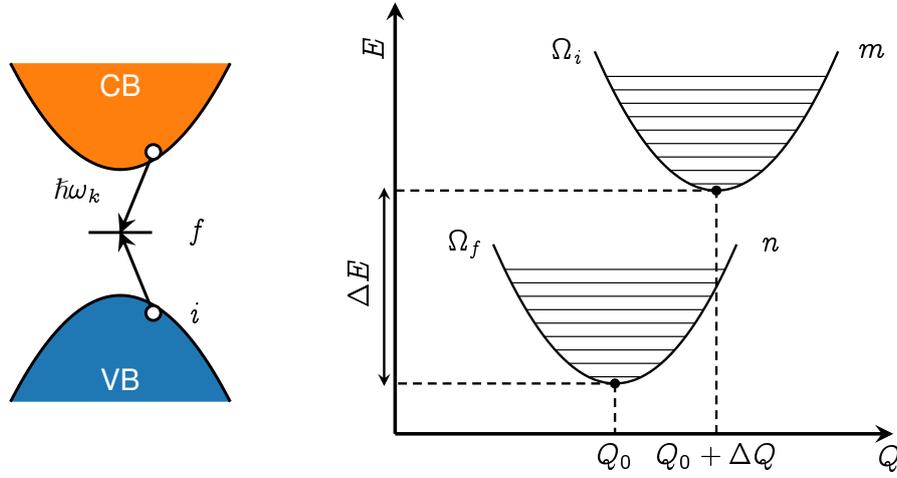
Figure 4.3: The nonradiative transition for a carrier (left): an electron (hole) jumps from the conduction (valence) band to the midgap state. On the right, the inputs for the capture coefficient as a function of the one-dimensional mode.

The $\triangle Q$ together with the reorganization energy $\lambda_{i/f}$ for the initial or final configurations:

$$\lambda_{i/f} = \frac{1}{2}\Omega_{i/f}^2\Delta Q^2 \tag{4.27}$$

They define the Huang-Rhys (HR) factor [82]:

$$S_{i/f} = \frac{\lambda_{i/f}}{\hbar\Omega_{i/f}} = \frac{1}{2}\Omega_{i/f}\Delta Q^2 \tag{4.28}$$

Given by the reorganization energy divided by the energy of the one-dimensional mode. It can also be seen as the number of phonons emitted during the relaxation from the initial to the final state and it quantifies the strength of the electron-phonon interaction for a given defect [50].

   The remaining quantities that need to be converted in the 1D approach are the electron-phonon matrix elements. Now there is only one matrix element, which is the interaction to the 1D mode and the wavefunctions of the initial and final state:

$$W_{if} = \langle\psi_i|\frac{\partial\hat{h}}{\partial Q}|\psi_f\rangle \tag{4.29}$$

43

Where the $\psi_{i/f}$ are single-particle Kohm-Sham wavefunctions and $\hat{h}$ their single-particle Hamiltonian.

Finally the the capture coefficient C in the 1D approach can be rewritten as:

$$C = \frac{2\pi}{\hbar} g V W_{if}^2 \sum_m w_m \sum_n |\langle \chi_{im}|Q - Q_0|\chi_{fn}\rangle|^2 \, \delta(\Delta E_{if} - m\hbar\Omega_i + n\hbar\Omega_f) \quad (4.30)$$

To sum up, the quantities needed for the nonradiative capture coefficient are:

- ΔE_{if} charge transition level of the transition

- $\Omega_{i/f}$ frequencies of 1D mode in the initial and final states

- W_{if} electron-phonon coupling to the 1D mode

The charge transition level can be obtained with the method of chapter 2. The frequencies $\Omega_{i/f}$ can be obtained by displacing the structure in the initial/final geometry along the one-dimensional mode and the frequencies are extracted with a quadratic fit.

4.3 | Benchmarking nonradiative capture for C_N in bulk GaN

In this section, we will benchmark the nonradiative recombination of the C_N defect in bulk GaN. This is a hole recombination and the nonradiative capture coefficient is available both experimentally and from ab initio calculations. This was first published as results in [44] and presented as a tutorial in [91].

In this benchmark, we used GPAW with the HSE functional as in the references and we tuned the mixing parameter to reproduce the experimental band gap of GaN (3.5 eV [92]). We used α = 0.31 for the mixing as in the references, but we note that using the same screening length ω = 0.2 Å, we did not get the same gap of 3.5 eV. Consequently, we used the default HSE value of ω = 0.11 Å. The difference is probably due to the different generations of the PAW setup between VASP and GPAW.

The nonradiative transition for the C_N defect can be understood by calculating the formation energy of the defect and see which charge transition level is inside the gap. For the C_N defect, it is already known that the $\varepsilon(0, -)$ is in the gap, and a hole is captured from the valence band at defect state via a nonradiative recombination. Therefore we need to relax only the neutral and the minus one state.
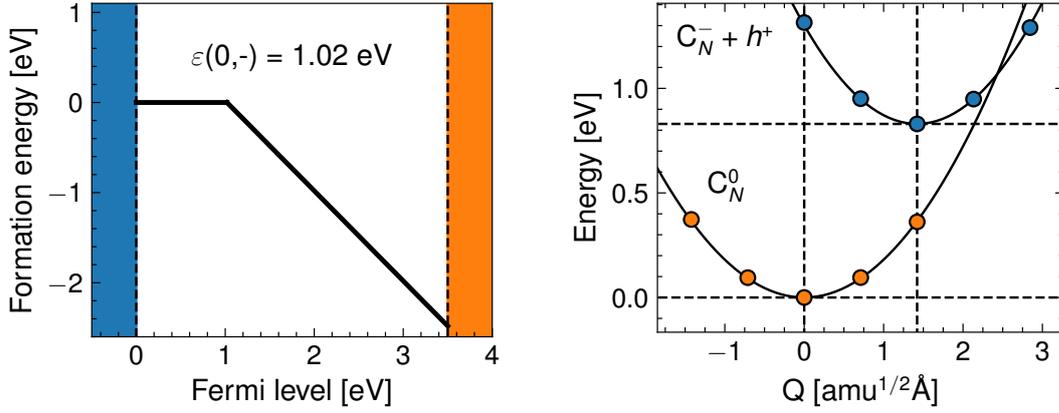
Figure 4.4: The formation energy for the C$_N$ defect in GaN as function of the Fermi level energy (left). The configuration coordinate diagram for the initial and final state.

Using the definition in chapter 2, the formation energy of the neutral defect is:

$$E_f[\mathrm{C_N^0}] = E_{tot}[\mathrm{C_N^0}] - E_{tot}[\mathrm{GaN}] - \mu_\mathrm{N} + \mu_\mathrm{C} \tag{4.31}$$

And no charge correction is needed. While for the minus one state:

$$E_f[\mathrm{C_N^-}] = E_{tot}[\mathrm{C_N^-}] - E_{tot}[\mathrm{GaN}] - \mu_\mathrm{N} + \mu_\mathrm{C} - (E_F + \varepsilon_v) + \Delta_{-1} \tag{4.32}$$

Since we are only interested in the CTL of the defect, in Fig. (4.4) we do not include the effect of the chemical potential.

The ΔE_{if} for the transition, in this case, is exactly the CTL referenced to valence band maximum:

$$\Delta E_{if} = \epsilon(0, -1) = E_f[C_\mathrm{N}{}^0; E_F\!=\!0] - E_f[C_\mathrm{N}{}^{-1}; E_F\!=\!0] \tag{4.33}$$

The other input is the frequency of the initial state, where the defect is negatively charged and the hole in the valence band and the final state, where the defect is neutral because the hole has recombined from the valence band. The configuration coordinate diagram of the initial and final state, separated by ΔE_{if} is shown in Fig. (4.4), where the geometries were displaced along the 1D mode and the frequencies $\Omega_{i/f}$ obtained with a quadratic fit.

The last quantity is the electron-phonon matrix element between the initial and final state with the 1D mode. This is evaluated with the formula introduced in chapter 2:

$$W_{if} = (\varepsilon_f - \varepsilon_i) \langle \psi_i | \frac{\partial \psi_f}{\partial Q} \rangle \tag{4.34}$$

This is obtained like in Ref. [44], from the overlap integral $S(Q) = \langle \psi_i(0) | \psi_f(Q) \rangle$
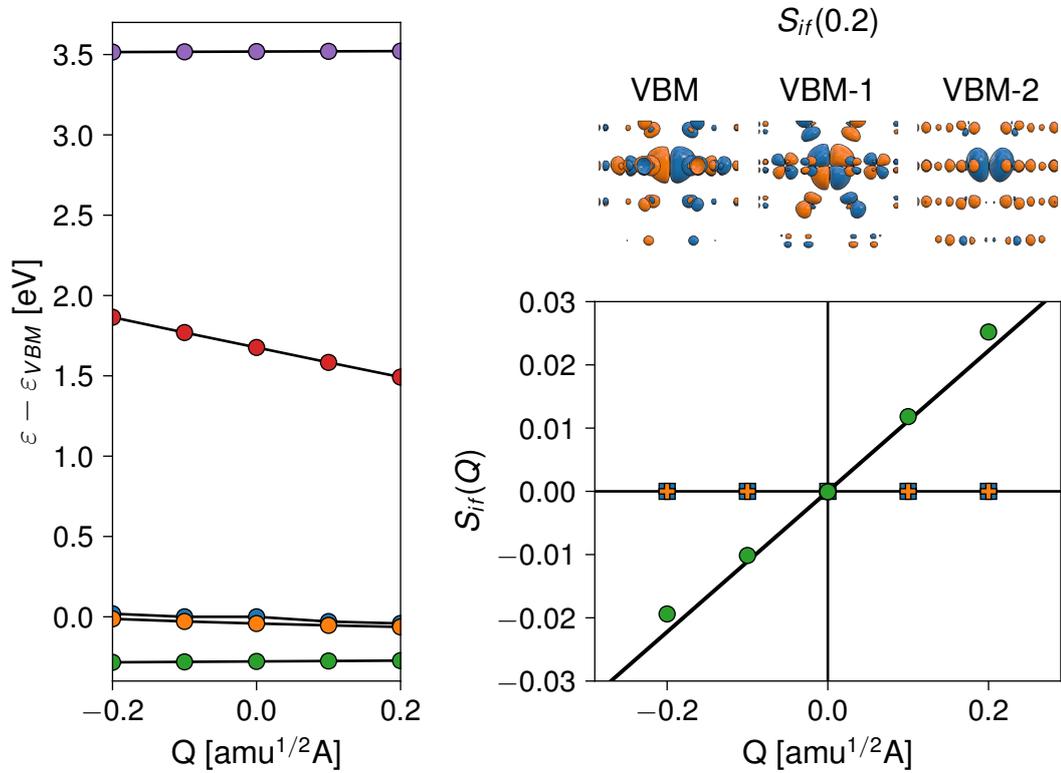
Figure 4.5: The Kohm-Sham eigenvalues of the three valence bands states, the defect state and the conduction band as function of the 1D mode (left). The overlap function $S(Q)$ as function of Q for the three states (right) and isosurfaces for $Q = 0.2$ amu$^{1/2}$ Å.

as function of Q and then making a linear fit to obtain the $\langle \psi_i | \frac{\partial \psi_f}{\partial Q} \rangle$ and then

multiplying for the energy difference $\varepsilon_f - \varepsilon_i$. The integral for the overlap function is evaluated with the all-electron wavefunctions. The matrix element can be either calculated in the initial or in the final state, but for the charged initial case, additional corrections are needed. Therefore here, it is calculated only in the final neutral state. Three matrix elements are present since the top of the valence band is degenerate three times.

Unlike in [44] where all three matrix elements are different from zero, in our case, only one is non-vanishing. In principle, this could be understood with group theory by considering the transformation proprieties of the valence band state and the defect state and then decomposing the 1D-mode in the different irreps. However, without going into this level of detail, it can be explained simply from the overlap function $S(Q)$. We show this in Fig. (4.5), where the isosurface for the overlap function is plotted for $Q = 0.2$ amu$^{1/2}$ Å. As we can see, only for the lowest band (VMB-2), the overlap function does not have alternating positive (orange) and negative (blue) regions, which results in a non-zero integral. This discrepancy can be due to a better conservation of the symmetry during the relaxation, for our case.

To sum up, we show a comparison of the values obtained with GPAW and two references in Table (4.1).

Table 4.1: Input parameters for the nonradiative capture coefficient for C_N and defect Zn_{Ga}-V_N in GaN.

		ΔE_{if} [eV]	$\|Q_i - Q_f\|$ [amu$^{1/2}$ Å]	$\hbar\Omega_i$ [meV]	S_i	$\hbar\Omega_f$ [meV]	S_f	W_{if} [eV/amu$^{1/2}$ Å]
C_N	GPAW(PBE)	0.53	0.89	36.0	3.4	41.5	3.9	0.108
	GPAW(HSE)	1.02	1.42	44.2	10.7	38.9	9.4	0.091
	Ref. [44]	1.02	1.61	42	-	36	10	0.064
	Ref. [91]	1.06	1.69	37.5	-	33.6	10	0.050
Zn_{Ga}-V_N	GPAW(HSE)	0.85	3.33	25	33	22	30	1.1×10^{-2}
	Ref. [44]	0.88	3.33	-	-	22	30	1.0×10^{-2}
	Ref. [85]	0.91	3.13	-	-	23	27	1.0×10^{-2}

We also calculated the input parameters for another defect, a Zn_{Ga}-V_N in GaN

[44, 85], which also exhibits the same hole recombination.

We see a good agreement between our benchmark and the references. However, it is also essential to mention the importance of using hybrid functionals. Indeed PBE is failing to reproduce the ΔQ and the CTL. On the other hand, the electron-phonon matrix element is somehow similar due to a cancellation of errors: the overlap is larger, but the difference in energy is smaller. Unlike the configuration coordinate diagram, the electron-phonon matrix element should be evaluated only for small displacement, where a linear behavior is expected. Small deviations from linear regime are clear in Fig. (4.5) for $Q = \pm 0.2$ amu$^{1/2}$ Å as in Ref. [44].
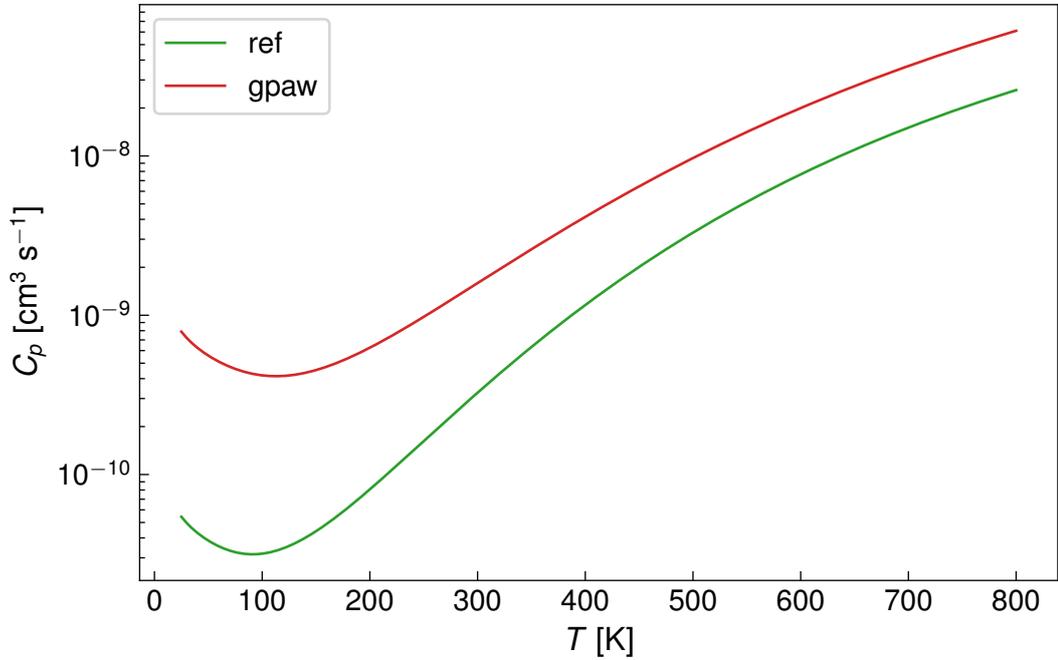
Figure 4.6: The capture coefficient C_p as function of the temperature for the C_N defect, for the Ref. [91] (green) and this benchmark (red).

The capture coefficient C_p in (4.30) has been implemented in ASR in the `asr.nonradiative` recipe following the method in Ref. [91]. The delta functions in (4.30) are approximated with Gaussians with smearing, while the overlap with

the 1D mode is calculated numerically using the ionic wavefunction:

$$\chi_{in}(Q) = \frac{1}{\sqrt{2^n n!}} \left(\frac{m\Omega_i}{\pi\hbar} \right)^{1/4} e^{-\frac{m\Omega_i Q^2}{2\hbar}} \mathcal{H}_n \left(\sqrt{\frac{m\Omega_i}{\hbar}} Q \right) \tag{4.35}$$

Where $\mathcal{H}_n(Q)$ is nth-order Hermite polynomial.

The capture coefficient C_p for the C$_N$ defect is shown in Fig. (4.6). There is one order of magnitude of difference between the reference and this benchmark due to the discrepancy in the input parameters. The parameter that has the dominant effect is the ΔE, because the capture coefficient is thermally activated. When the temperature is increasing, more vibrational excited states of the initial configuration are populated, and the nonradiative transition is more likely to happen.

49

4.4 │ Radiative transitions between defect states

Here we will discuss the effect of the defect on the radiative properties of the material. The ideas behind the methods are presented and we show one result for each defect. The methods have also been applied to the most promising defects in Paper IV.

In the first method, the 1D approach introduced before will be used to obtain the emission lineshape for a transition between two ground and excited states of the defect (Fig. 4.7). The 1D lineshape reproduces well the experimental lineshape in the case of large electron-phonon interaction, or with the terminology of the previous section, in the case of large HR factor (S > 10) [93].
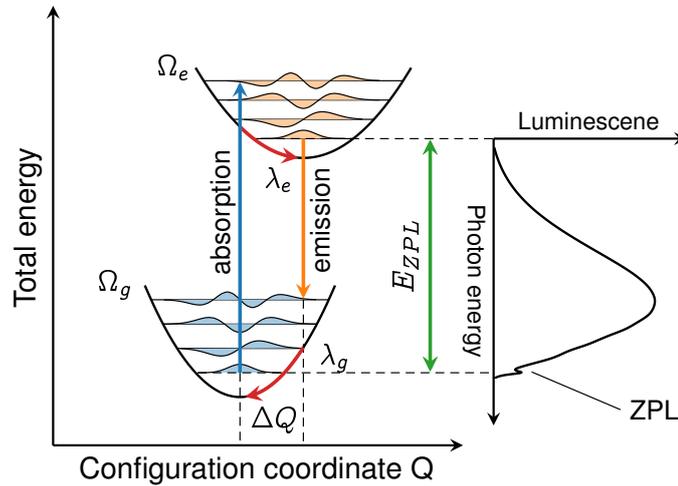
Figure 4.7: The emission and absorption processes from the ground and excited state of the defect.

The radiative transition differs from the non-radiative case, in that, the transition takes place between the ground and the excited state of the defect, with the emission of light. For the nonradiative case, the transition was from a band to the defect state, with the emission of phonons. If the separation in energy between the ground and excited state is E_{ZPL}, i.e. the zero-phonon line [82], the

luminescence for the emission process within the 1D approach is given by [50]:

$$L_{em}(\omega) = \sum_n \frac{e^{-S_g} S_g^n}{n!} \delta(\omega - n\Omega_g + \Omega_e - E_{ZPL}) \qquad (4.36)$$

Where S_g is the HR factor of the ground state, obtained from the displacement $\triangle Q$ and the reorganization energy in the ground state λ_g. Within the 1D approach, it is straightforward to calculate the absorption spectra with the HR factor of the excited state S_e:

$$L_{abs}(\omega) = \sum_n \frac{e^{-S_e} S_e^n}{n!} \delta(\omega - n\Omega_e + \Omega_g - E_{ZPL}) \qquad (4.37)$$

The 1D method fails to describe defects with a small HR factor and these systems are the most interesting ones for single-photon applications [64].

The other method developed for paper IV is the generating function approach, introduced in [94]. In this case, the contribution of all phonon modes of the system is included and therefore there is not a single HR factor, but there are the partial HR factors:

$$S_\lambda = \frac{\omega_\lambda q_\lambda^2}{2} \qquad (4.38)$$

Where q_λ is the projection of the 1D mode over the phonon mode λ:

$$q_\lambda = \sum_{\alpha t} m_\alpha^{1/2} (R_{\alpha t}^{(e)} - R_{\alpha t}^{(g)}) \cdot e_{\alpha t}^\lambda \qquad (4.39)$$

Then the electron-phonon function is defined as:

$$S(\omega) = \sum_\lambda S_\lambda \delta(\omega - \omega_\lambda) \qquad (4.40)$$

That gives an idea of which phonon mode contributes the most for the transition from the excited to the ground state of the defect. In this approach, the lineshape is proportional to the spectral function:

$$\mathcal{A}(E_{ZPL} - \omega) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} dt G(t) e^{i\omega t - \gamma|t|} \qquad (4.41)$$

That can be obtained by the generating function $G(t)$, which is the Fourier transform of the electron-phonon function $S(\omega)$:

$$G(t) = e^{S(t) - S(0)} \qquad (4.42)$$

Where $S(0)$ is:

$$S(0) = \sum_{\lambda} S_{\lambda} \tag{4.43}$$

The smearing parameter γ is needed to converge the integral and it has the effect of giving the width of the zero-phonon line peak. A standard approach has become to tune this parameter to match the experimental lineshape [94].
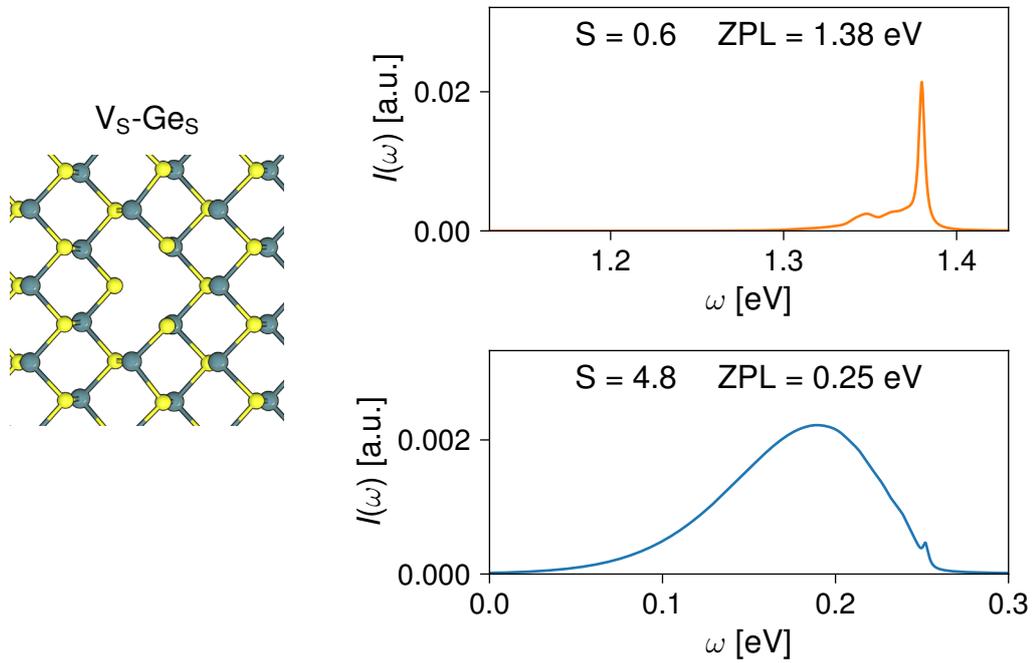
Figure 4.8: The emission lineshape for the transition in the spin-up channel (upper) and in the spin-down channel (lower) for the S_{Ge}-V_S in Ge_2S_2.

We have applied the two methods for some defects in Paper IV and they are available in ASR in the `asr.luminescence` recipe. Here we show as an example, a defect that is not included in Paper IV. This is the S_{Ge}-V_S defect, where both excitations in the two spin-channels are performed and the generating function approach is applied for the emission lineshape (Fig. 4.8).

5 | High-throughput computational studies

In this chapter, we will briefly present the idea of a high-throughput approach for material discovery and we will show examples of this procedure applied to 2D materials.

High-throughput in materials design refers to the systematic calculation of material properties to find a set of ideal candidates for the experimental realization [95, 96]. Online material databases can help perform such studies, where materials can also be screened for specific target properties. Examples of databases with materials with any dimensionality are: AFLOW [97], ICSD [98], Materials Project [99], OQDM [100] and NOMAD [101]. Databases that are only oriented to 2D materials and with advanced methodology are the Computational 2D Materials Database (C2DB) [102] and the database from Mounet et. al. [103].
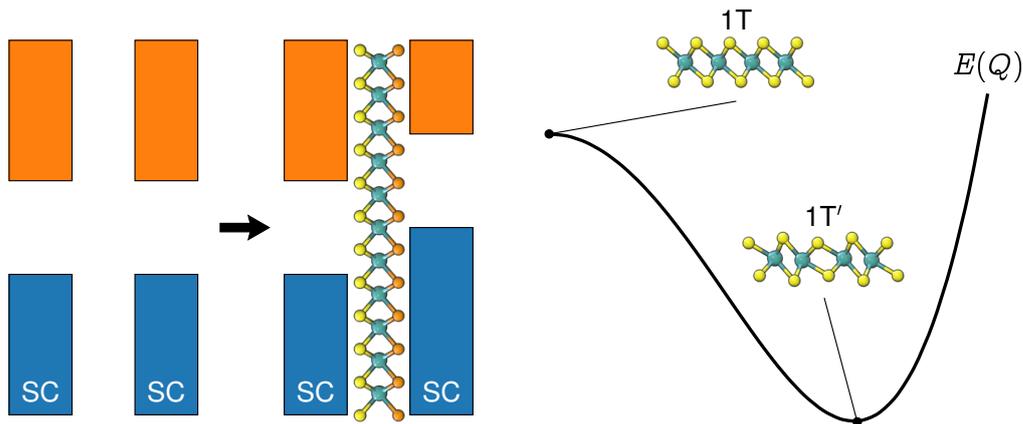
Figure 5.1: The Janus monolayer (MoSSe) can change the band alignment in a semiconductor-semiconductor interface (left). Displacing along the distortion (Q) of the 1T-MoS$_2$ is possible to get the 1T'-MoS$_2$ (right).

The first example of this high-throughput procedure was the calculations performed on the Janus monolayers in Paper I and II. This class of 2D material posses a finite dipole along the out-of-plane direction due to the difference in electronegativity of the chemical elements on the two sides of the layer. In Paper I, it was proved that this finite dipole could be used in heterostructures to control

the band alignment in semiconductor/semiconductor interface or to control the Schottky barrier height in semiconductor/metal interface. Consequently, a high-throughput study was conducted in Paper I. We found a set of 47 new stable Janus monolayers and included it in the updated version of C2DB (Paper II).

In paper III, another high-throughput screening study was performed where we investigated structural instabilities of 2D materials. Usually, the stability of a calculated material is inferred by the energy above the convex hull ΔH_{hull} [102]. Another criterion is the stability against structural distortions, which are manifested by the presence of imaginary phonon modes. A prototypical example is the T-phase of MoS_2 (Fig. 5.1), which has an in-plane instability. Within C2DB, the dynamical stability is tested by evaluating the stiffness tensor and the Hessian matrix of a 2x2 supercell of the primitive cell. In paper III, it was checked that this test for dynamical stability is reliable against a complete full phonon calculation in a larger supercell. In addition, we took a set of dynamically unstable materials in C2DB and we displaced them along the unstable mode. This procedure has the effect of finding a new more stable material, like in the case of the 1T'-phase of MoS_2, which is obtained by pushing the atoms along the distortion of 1T-phase. The workflow of Paper II was then applied for the new stable materials.

Finally, a high-throughput approach was applied for defects in Paper IV, where we selected the most stable non-magnetic monolayers in C2DB, with a band gap larger than 1 eV, for a total of 82 materials. Then we created vacancies and antisites, for a total of 500 defects and every defect was relaxed in different charged states.

All the point groups found in Paper IV are shown in Fig. (5.2), where they are ordered by decreasing symmetry. An example of a defect for every point group is also shown.

Starting from the most symmetric defects, we recovered the well-known V_B^{-1} in hexagonal boron nitride. This is a well-established defect in BN and one of the first observed color centers in 2D materials [104].

With $\mathbf{C}_{3v}$ symmetry, $SiCH_2$ is particularly interesting due to its structural similarity to bulk SiC. However, monolayer SiC is not thermodynamically stable (ΔH_{hull} = 0.5 eV/atom), whereas $SiCH_2$ is stable (ΔH_{hull} 0.06 eV/atom). Therefore, $SiCH_2$ can be considered the stable analogue of bulk SiC, with a large bandgap of 3.78 eV at PBE level.

Another important defect is the antisite S_W in WS_2. This defect is particularly interesting because if the relaxation is carried out without breaking the symmetry of the supercell, it leads to a local minimum in the structure. Whereas, if the
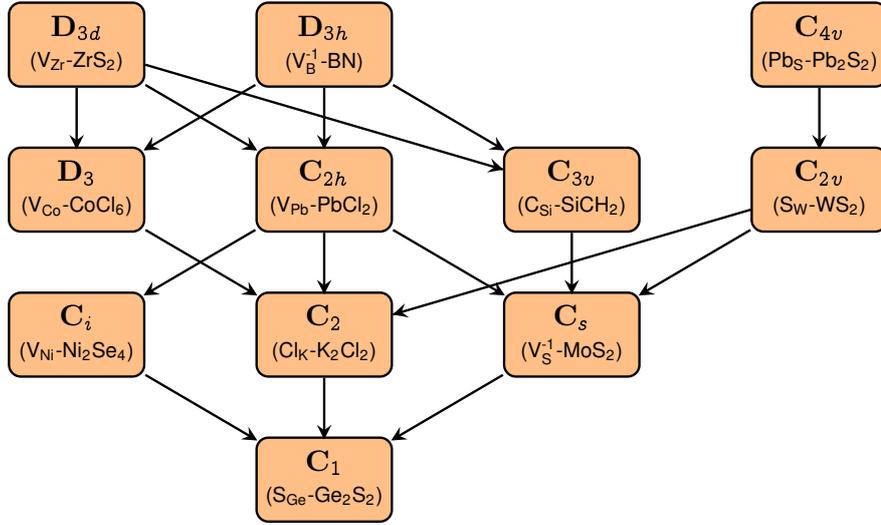
Figure 5.2: The point groups for the defects in paper IV. The arrows represent the subgroup hierarchy among all point groups.

symmetry is broken before the relaxation, it leads to a more stable structure. This consolidates the approach in Paper IV of breaking the symmetry of the supercell to find stable defect structures.

Our symmetry-based defect classification approach is further validated by the right classification V_S^{-1} in MoS_2, which reduces the symmetry from neutral (C_{3v}) to the -1 charged (C_s) [58] due to the Jahn-Teller distortion in the double degenerate defect states.

6 | Conclusion and Outlook

As explained in the thesis, we have developed ab initio approaches to describe defects in two 2D materials, and ways of engineering defects to optimize the materials' properties.

We stress the importance of the localized defect states inside the gap for the excited state property. Then we explain how to systematically perform the symmetry analysis for such states and characterizing them according to the irreps of the point groups. This approach is applied for thousand of defect states in 2D materials.

Then we presented two applications of the symmetry analysis: to study the symmetry of the multiplet state made from the occupied defect states, and to predict the effect of spin-orbit coupling on the defect states.

We also explore the role that phonons play in the emission and absorption processes from the defect states. We first make some general considerations on the effect of defects on the phonon frequencies of the pristine material. We show how defects can produce localized modes and resonance modes, using antisite defects in monolayer MoS_2.

Regarding the optical properties, we implement a method to quantify the nonradiative recombination of carriers at defect states and benchmark using the C_N defect in bulk GaN. We also implement the state-of-the-art methods to study the emission lineshape for transitions between defect states.

Finally, we explore the structural instabilities in the materials in C2DB in a high-throughput approach study.

As concluded above, the approaches developed and implemented in this thesis are fairly general. Therefore, one can potentially apply these to explore the different properties of numerous materials for various applications.

A | Appendix

A.1 | Characters and direct product tables for $\bar{\mathrm{C}}_{3v}$

$\bar{\mathrm{C}}_{3v}$	E	$\bar{E}$	C_3 $C_3^2\bar{E}$	C_3^2 $C_3\bar{E}$	σ_v	$\sigma_v\bar{E}$
A_1	1	1	1	1	1	1
A_2	1	1	1	1	-1	-1
E	2	2	-1	-1	0	0
$E_{1/2}$	2	-2	1	-1	0	0
$E_{3/2}^1$	1	-1	-1	1	i	$-i$
$E_{3/2}^2$	1	-1	-1	1	$-i$	i

Table A.1: Characters tables for the $\bar{\mathrm{C}}_{3v}$ double group.

$\bar{\mathrm{C}}_{3v}$	A_1	A_2	E	$E_{1/2}$	$E_{3/2}^1$	$E_{3/2}^2$
A_1	A_1	A_2	E	$E_{1/2}$	$E_{1/2}^1$	$E_{1/2}^2$
A_2	A_2	A_1	E	$E_{1/2}$	$E_{1/2}^2$	$E_{1/2}^1$
E	E	E	$A_1 \oplus A_2 \oplus E$	$E_{1/2} \oplus E_{1/2}^1 \oplus E_{1/2}^2$	$E_{1/2}$	$E_{1/2}$
$E_{1/2}$	$E_{1/2}$	$E_{1/2}$	$E_{1/2} \oplus E_{1/2}^1 \oplus E_{1/2}^2$	$A_1 \oplus A_2 \oplus E$	E	E
$E_{3/2}^1$	$E_{3/2}^1$	$E_{3/2}^2$	$E_{1/2}$	E	A_2	A_1
$E_{3/2}^2$	$E_{3/2}^2$	$E_{3/2}^1$	$E_{1/2}$	E	A_1	A_2

Table A.2: Direct product of the irreps for the $\bar{\mathrm{C}}_{3v}$ double group.

Papers

Paper I: Engineering atomically sharp potential steps and band alignment at solid interfaces using 2D Janus layers

A. C. Riis-Jensen, <u>Simone Manti</u>, Kristian S. Thygesen

Article

Engineering Atomically Sharp Potential Steps and Band Alignment at Solid Interfaces using 2D Janus Layers

Anders C. Riis-Jensen, Simone Manti, and Kristian S. Thygesen*

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: The possibility of stacking two-dimensional (2D) materials into van der Waals (vdW) heterostructures has recently created new opportunities for band structure engineering at the atomic level. However, despite the weak vdW interaction, controlling the electrostatic potential governing the band lineup at the 2D interfaces is still posing a significant challenge. Here, we demonstrate that 2D Janus monolayers, possessing an intrinsic out-of-plane dipole moment, can be used to control the band alignment at semiconductor−semiconductor and metal−semiconductor interfaces in a highly predictive manner. Using density functional theory (DFT), we calculate the band structure of a wide range of different vdW interfaces. We find that upon insertion of a Janus structure the band line-ups and Schottky barriers can be controlled to high accuracy. The main result of this work is that the out-of-plane dipole moment of the Janus structure changes little upon insertion in the interface. As a consequence, the effect on the electrostatic potential at the interface can be predicted from the properties of the freestanding Janus structure. In addition to this, we predict 47 stable Janus monolayers, covering a wide range of dipole moments and band edge positions and thus providing a comprehensive library of 2D building blocks for manipulating the band alignment at interfaces.
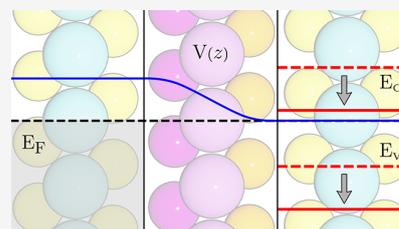
INTRODUCTION

More than a decade after the discovery of graphene,[1] the class of atomically thin two-dimensional (2D) crystals remains one of the hottest topics in physics. An important reason is that such materials open new possibilities for studying and manipulating electronic quantum states directly at the atomic length scale. By stacking different 2D layers into van der Waals (vdW) heterostructures, it is, at least in principle, possible to design the energy landscape of the electrons, i.e., the band structure, with a precision far beyond what is possible with conventional epitaxial growth.[2] In this paper, we introduce the idea of using 2D Janus monolayers to control the electrostatic potential and band line up at vdW interfaces.

Many solid-state devices, including semiconductor (SC) lasers,[3−5] solar cells,[6,7] and transistors,[8] rely on heterostructure materials with electronic energy levels carefully aligned across their interfaces. This makes the interfacial band lineup problem one of the most critical challenges for the semiconductor industry. Quite generally, the band alignment at SC heterojunctions can be categorized into three types according to the relative position of valence and conduction bands on the two sides of the interface. In Figure 1, we show examples of type-I and type-II band alignments. It is exceedingly difficult if not impossible by means of conventional techniques employed in the semiconductor industry (doping, strain, alloying) to change the band alignment at an interface without changing the composition or structure of the material(s) in the vicinity of the interface. This clearly implies a risk of degrading other of the material properties, e.g., the carrier mobility, carrier lifetime, etc. A key result of this paper
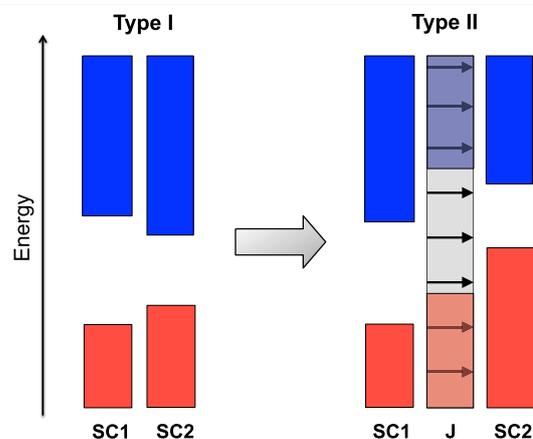
Figure 1. Left: sketch of a type-I band alignment between two different semiconductors. Right: sketch of the achieved type-II band alignment between the two semiconductors by sandwiching a Janus monolayer between them, which shifts the electrostatic potential on either side of the Janus monolayer. In a type III band alignment, the valence band maximum (VBM) of one of the structures is located above the conduction band minimum (CBM) of the other structure.

59

is that the band alignment can be controlled elegantly and with high precision in vdW heterojunctions without affecting the material structure or composition.

For metal/SC interfaces, the single most critical parameter used to characterize the band alignment is the height of the Schottky barrier (SB), i.e., the distance between the metal Fermi level and the nearest band edge of the SC (valence band or conduction band). Apart from the position of the bands in the intrinsic materials, i.e., the work function of the metal and the electron affinity and ionization potential of the SC, a number of effects influence the SB including Pauli repulsion and metal-induced gap states (MIGSs).[9−13] Such effects lead to Fermi-level pinning (FLP)[10,14−16,18−22] and deviations from the Anderson rule,[23] which makes it difficult to control (in practice minimize) the SB. This is further complicated by the unpredictable effect of the chemical interactions at the interface.[17] One specific and important motivation for developing means to control the SB at metal−2D semiconductor interfaces comes from the huge interest in atomically thin field-effect transistors, e.g., based on transition-metal dichalcogenides (TMDs). For such devices, the SB at the source/drain contacts remains a performance-limiting factor that must be improved for them to become competitive.[24−33]

Due to the presence of MIGS, it has been proven ineffective to dope SC/metals structures to lower the SB. This has previously been explained by the fact that while originating on the metal the states are still localized deeply into the SC,[34] effectively neutralizing the effect of doping. Approaches to overcome this includes inserting a layer of boron nitride between the metal and the SC[22,35] or different oxide layers;[36−39] however, the exact effect of all of the above-mentioned mechanisms is extremely difficult to predict.

In this work, we propose an alternative method to control and adjust the band lineup at solid interfaces in a highly predictable manner. By sandwiching a 2D material with a finite out-of-plane dipole moment, i.e., a Janus monolayer, between two materials, it is possible to introduce an atomically sharp potential step across the interface. For SC/SC interfaces, this opens the possibility of controlling the band lineup and even changing the type of band alignment, while for metal/SC interfaces, the SB can essentially be made to vanish. Importantly, we find that the dipole moment of the Janus layer does not change significantly upon insertion in the heterostructure, meaning that the resulting shift in the band offset can be predicted to high accuracy from the properties of the isolated Janus monolayer, at least for systems with small interlayer charge transfer. Beyond pure vdW-bonded structures, we find that the intrinsic dipole moment is unaltered, even when placed on conventional bulk metal surfaces. Using high-throughput density functional theory (DFT) computations, we further predict 47 new Janus monolayers that we find to be stable, have a finite band gap, and a finite out-of-plane dipole moment. The induced shift in electrostatic potential achievable by these novel 2D Janus monolayers range from close to zero to about 2 eV, showing the high degree of band offset tunability offered by the proposed concept.

■ COMPUTATIONAL DETAILS

All calculations were performed with the GPAW code.[40] Monolayer structures were relaxed using the Perdew−Burke−Ernzerhof (PBE) functional[41] on a Monkhorst−Pack k-point

grid[42] with a k-point density of 6.0 and a planewave cutoff of 800 eV. The unit cells had 15 Å of vacuum in the perpendicular direction, and a Fermi smearing of 0.05 eV was used. The band structure and the size of the dipole moment were calculated using a k-point density of 12.0. Spin−orbit coupling was not included. For the multilayer structures, the interlayer binding distances were determined using the BEEF-vdW functional.[43] All structures were relaxed until the maximum force on any atom was below 0.01 eV/Å and the maximum stress on the unit cell was 0.002 eV/Å³. Since the layers in the heterostructures, in general, will have incommensurable lattices, it is necessary to use larger supercells and strain one or both materials slightly. Table 1

Table 1. Computational Details about the Supercells Used for the Heterostructure Calculations Including the Cell Size, Strain, and Smallest Interlayer Distance d

structure	cell size	strain (%)	d (Å)
graphene/H-MoSSe/hBN	graphene: 4 × 4	1.62	3.82/3.69
	H-MoSSe: 3 × 3	−0.46	
	hBN: 4 × 4	3.31	
T-MoS₂/H-MoSSe/H-MoS₂	T-MoS₂: 1 × 1	0	3.51/3.67
	H-MoSSe: 1 × 1	−2.06	
	H-MoS₂: 1 × 1	0	
H-MoS₂/H-MoSSe/H-MoS₂	H-MoS₂: 1 × 1	0	3.54/3.63
	H-MoSSe: 1 × 1	−2.06	
H-WS₂/H-MoSSe/H-WS₂	H-WS₂: 1 × 1	0	3.52/3.66
	H-MoSSe: 1 × 1	−2.06	
H-MoS₂/H-CrSSe/H-MoS₂	H-MoS₂: 1 × 1	0	3.61/3.75
	H-CrSSe: 1 × 1	1.76	
H-MoSe₂/H-MoSTe/H-MoSe₂	H-MoSe₂: 1 × 1	0	3.61/3.74
	H-MoSTe: 1 × 1	<0.01	
H-MoSe₂/H-TaSSe/H-MoSe₂	H-MoSe₂: 1 × 1	0	3.58/3.81
	H-TaSSe: 1 × 1	−2.59	
H-MoSe₂/H-TiSSe/H-MoSe₂	H-MoSe₂: 1 × 1	0	3.63/3.86
	H-TiSSe: 1 × 1	−2.93	
H-MoSe₂/H-WSeTe/H-MoSe₂	H-MoSe₂: 1 × 1	0	3.66/3.58
	H-WSeTe: 1 × 1	−3.29	
H-MoS₂/H-WSSe/H-MoS₂	H-MoS₂: 1 × 1	0	3.54/3.65
	H-WSSe: 1 × 1	−2.10	
H-MoSe₂/H-WSTe/H-MoSe₂	H-MoSe₂: 1 × 1	0	3.64/3.68
	H-WSTe: 1 × 1	−1.25	
H-MoSe₂/H-ZrBrCl/H-MoSe₂	H-MoSe₂: 1 × 1	0	3.69/3.57
	H-ZrBrCl: 1 × 1	−4.82	
H-MoSe₂/H-TiBrCl/H-MoSe₂	H-MoSe₂: 1 × 1	0	3.78/3.64
	H-TiBrCl: 1 × 1	−1.75	

presents a summary of the supercell size and strain applied to each layer in each heterostructure. For H-phase/H-phase interfaces, AB stacking was used, while for H-phase/T-phase interfaces, AA stacking was used, i.e., the metal atoms were stacked on top of each other. These stacking configurations were found to be energetically most stable. To judge the stability (dynamic and thermodynamic) of the new Janus monolayers resulting from the high-throughput study, we followed the criteria of the Computational 2D Materials Database.[44]

■ RESULTS

Recently, both MoSSe[45,46] and BiTeI[47] have been realized in the monolayer form experimentally. These structures possess
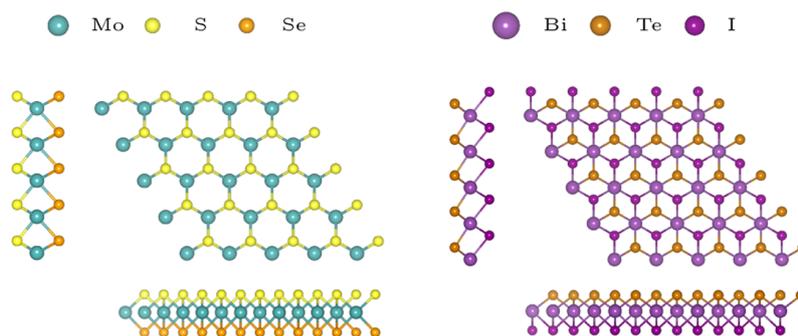
Figure 2. Top and side views of the atomic structures of the MoSSe prototype (left) and the BiTeI prototype (right).
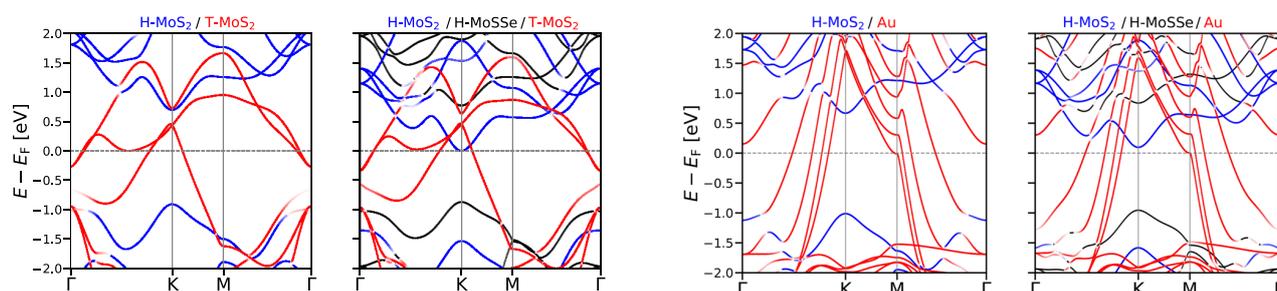
Figure 3. Left: band structure of a bilayer consisting of MoS_2 in the H- and T-phases, respectively. Right: band structure of the same structure with the Janus monolayer MoSSe sandwiched in between the two layers. For both plots, red/blue is the projection of the bands onto the T- and H-phase of MoS_2, respectively, while black is the projection onto MoSSe. Notice how the insertion of MoSSe shifts the band edges of H-MoS_2 with respect to the Fermi level of the structure, effectively lowering the SB.

Figure 5. Left: band structure of H-MoS_2 at the bulk Au metal surface. Right: band structure of the same structure with the Janus monolayer MoSSe sandwiched in between. For both plots, red/blue is the projection of the bands onto the Au and H-MoS_2, respectively, while black is the projection onto MoSSe.

an out-of-plane dipole moment created by the difference in electronegativity of the S and Se (I and Te) atoms,[48] which are located on different sides of the central metal layer, see Figure 2. Such structures with broken mirror symmetry and a finite out-of-plane dipole moment are known as Janus monolayers.

In Figure 1, we illustrate the concept of band lineup control at vdW heterojunction interfaces. On the left is shown an example of a bilayer consisting of two (possibly) different semiconducting 2D monolayers having a type-I band alignment. On the right is shown the same two monolayers, with a Janus monolayer sandwiched in-between. By choosing a Janus monolayer with an appropriate size of the out-of-plane dipole moment, it is possible to shift, for instance, a type-I band alignment into a type-II band alignment as shown. Another obvious application is to use Janus monolayers to lower the SB at metal/SC interfaces. For a general metal/SC interface, where the Fermi level of the metal is located between the valence band maximum (VBM) and conduction band minimum (CBM) of the SC, the bands originating from the
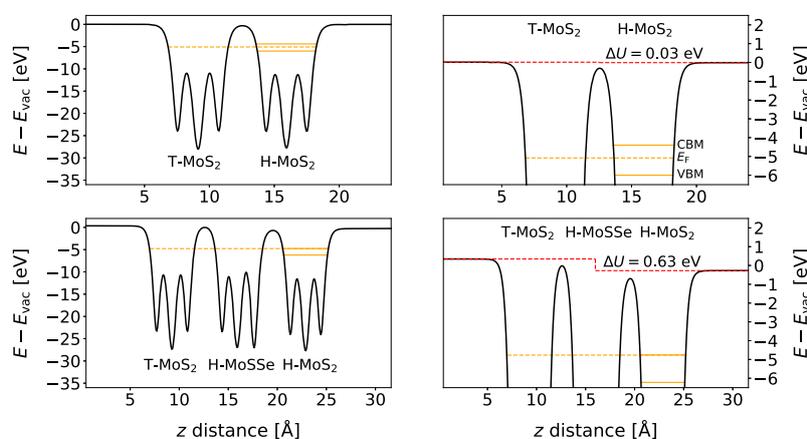
Figure 4. Electrostatic potential, with respect to the vacuum energy, for the bilayer and trilayer structures from Figure 3 in the direction perpendicular to the layers. The right column is a close-up of the top of the figures in the left column, showing the vacuum-level shift ΔU on either side of the heterostructures. In orange is shown the VBM, CBM, and E_F.
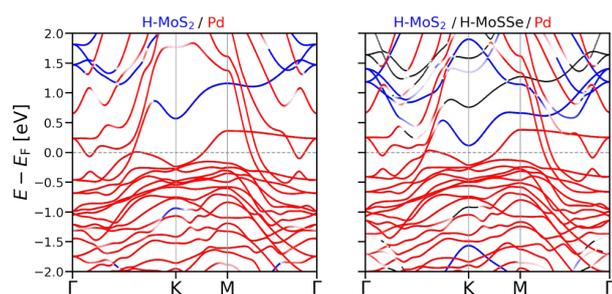
Figure 6. Same as Figure 5 with a Pd(111) surface instead of a Au(111) surface.
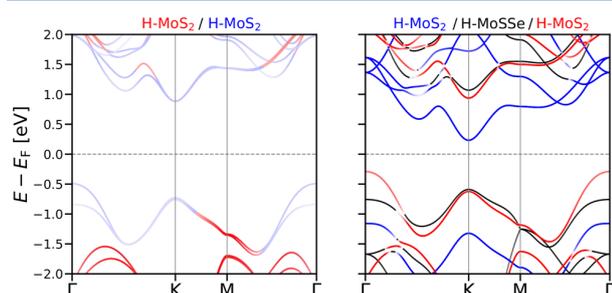
Figure 7. Left: band structure of a bilayer H-MoS$_2$. Right: band structure of the same structure with the Janus monolayer MoSSe sandwiched in between the two layers. For both plots, red/blue is the projection of the bands onto the two MoS$_2$ layers, while black is the projection onto MoSSe. Notice how the insertion of MoSSe splits the bands of the two MoS$_2$ layers at the K-point.
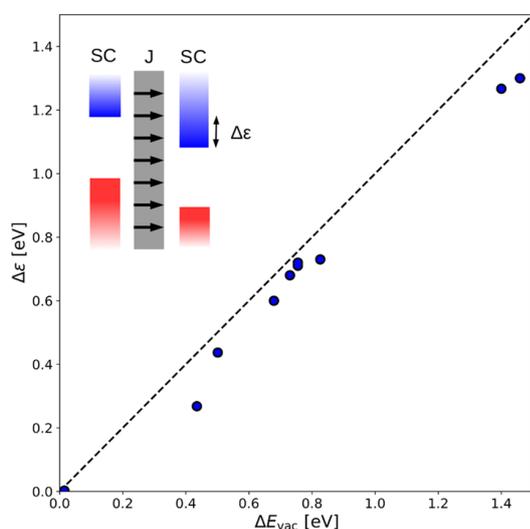
Figure 8. Splitting of the conduction bands $\Delta\epsilon$ for a SC bilayer, upon insertion of a Janus monolayer, as a function of the shift in potential across the freestanding Janus monolayer ΔE_{vac}. Each data point represents a different Janus monolayer sandwiched between two identical SCs. From left to right, the Janus monolayers are TiSSe, ZrBrCl, TiBrCl, WSeTe, WSSe, MoSSe, MoSSe, CrSSe, WSte, and MoSTe. The exact trilayer composition can then be deducted from Table 1. We stress that there is little-to-no difference between trilayers H-MoS$_2$/H-MoSSe/H-MoS$_2$ and H-WS$_2$/H-MoSSe/H-WS$_2$. The fact that the shift in band energies is very close to the shift of the freestanding Janus monolayer shows the highly predictive nature of the proposed strategy for band alignment engineering.
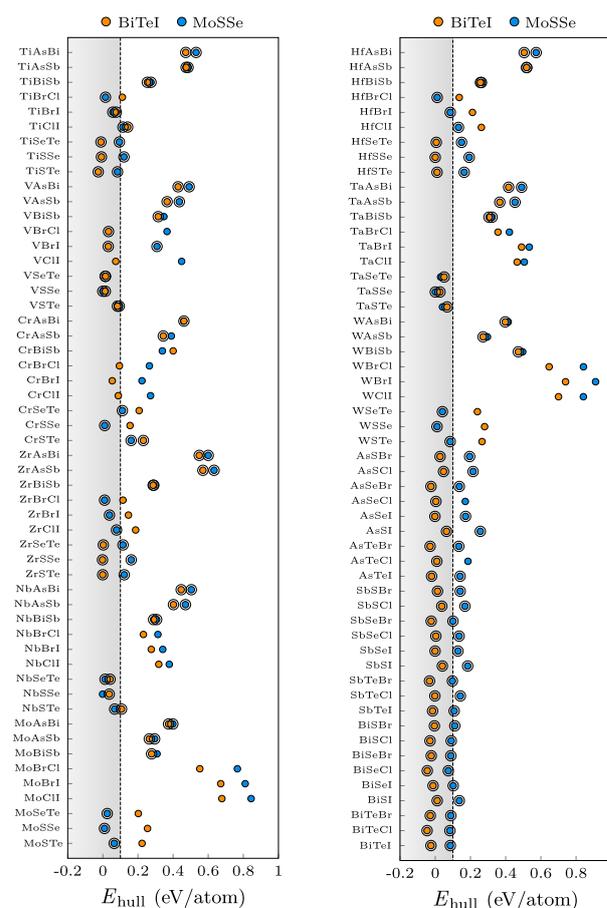
Figure 9. Energy above the convex hull for all Janus monolayers calculated in this study colored according to their phase. Structures with an energy above the convex hull lower than 0.1 eV/atom are considered thermodynamically stable. Dynamical stability is shown by a black circle.

SC will bend at the interface, creating a Schottky barrier. Below we show that by inserting a Janus monolayer between the metal and the SC, one can control the SB.

Schottky Barriers at Metal/SC Interfaces. To illustrate the concept in practice, we have calculated the band structure of both SC/SC and metal/SC vdW heterostructures with and without a Janus monolayer sandwiched in-between. First, consider the two band structures shown in Figure 3. On the left is shown the band structure of bilayer T-MoS$_2$/H-MoS$_2$. We note that MoS$_2$ is well known to be metallic in the T-phase and semiconducting in the H-phase. We note that the T-phase is in fact dynamically unstable and undergoes a transition to the T′-phase; however, this is unimportant for the present discussion. On the right is shown the band structure of the same structure with H-MoSSe sandwiched in-between. For the two cases, we consider an AB and ABA stacking, respectively. The red color represents the bands projected onto the T-MoS$_2$ layer, blue is projected onto the H-MoS$_2$ layer, and black is projected onto the Janus H-MoSSe layer. For the bilayer, we see that the bands from T-MoS$_2$ are crossing the Fermi level, which overall renders the bilayer metallic, while the H-MoS$_2$ more or less preserves the size of its direct band gap at the K-point. The distance from the CBM of H-MoS$_2$ to the Fermi level is approximately 0.6 eV,
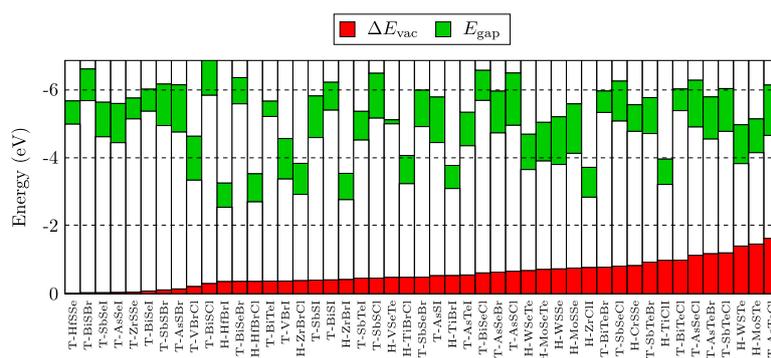
Figure 10. Overview of all stable Janus monolayers in the MoSSe (H-phase) and BiTeI (T-phase) prototypes, sorted according to the vacuum-level shift (red). The band gap is shown in green on an absolute scale relative to vacuum, such that the valence band maximum and conduction band minimum can be found from the white/green boundaries. All energies refer to PBE calculations.

which equals the SB of the interface. On the right, after insertion of the MoSSe layer, we see how the SB has essentially vanished. Now, the CBM of the H-MoS_2 layer just touches the Fermi level, thus reducing the SB to almost zero. We note that the SB of T-MoS_2/H-MoS_2 has previously been studied, where it was shown that the SB can be shifted by 0.3 eV by varying the stacking pattern.[49] While this is an interesting result, the method proposed in our study gives the possibility to vary the SB, in a highly controlled manner, anywhere between the electronic band gap and essentially 0 eV, by varying the Janus monolayer material, which will be demonstrated later.

A more detailed picture of the electrostatic discontinuity induced by the Janus monolayer is shown in Figure 4. The top panel shows the electrostatic potential perpendicular to the T-MoS_2/H-MoS_2 bilayer (averaged over the in-plane directions). The lower panel shows the same for the trilayer structure including H-MoSSe. Here, the right-hand side plot is a close-up of the upper part of the left plot. For the bilayer, we see a minor change in the electrostatic potential across the bilayer of 0.03 eV. Since both T-MoS_2 and H-MoS_2 have a zero dipole moment in the out-of-plane direction in their freestanding monolayer form, this shift originates from a small interfacial dipole created at the interface. This phenomenon is well known for metal−SC interfaces and has also been seen at graphene/SC interfaces.[50] After insertion of H-MoSSe, we now find a shift of 0.63 eV. This large shift arises because of the internal dipole in the H-MoSSe layer and gives the shift in the band structure we observed in Figure 3. We here stress that we in this study consider the manipulation of SB defined by the SC band edges relative to the Fermi level and do not consider the tunneling barrier evident from Figure 4, which for some systems can impose an additional barrier. Modeling of such additional tunneling barriers are system-specific and will be the scope of future studies.

It can be interesting to study how the intrinsic dipole moment of the Janus monolayers are affected by strain effects. To do this, we calculate the vacuum-level shift and band gap of strained monolayer H-MoSSe. The applied strain ranges from -3% to +3% (see Figure S1 in the Supporting Information), which we believe are experimentally realizable values. We find that the vacuum-level shift across monolayer MoSSe can be varied between 0.74 and 0.77 eV. While this gives rise to only little variation in the intrinsic dipole moment of the Janus monolayers, it shows that the proposed setup is very robust to strain effects. On the other hand, the band gap

changes over a range from 1.1 to 1.6 eV, showing a great possible tunability of the electronic properties of Janus monolayers.

To further illustrate the concept, we consider the band structure of a bulk metal/SC interface and calculate the band structure of H-MoS_2 at the bulk Au(111) surface and at the bulk Pd(111) surface. The bulk metal surface is modeled by a slab of four atomic layers. The calculation has been performed in a 1×1 cell, with H-MoS_2 unstrained and Au and Pd strained significantly to match the lattice constant of H-MoS_2. Despite the very large strain, we find that the typical characteristics of both the Au and Pd band structures can be recognized, i.e., the wide 6sp bands crossing the Fermi surface and the flat 5d bands at around −1.5 eV for Au and the 5d bands crossing the Fermi level for Pd. This can be seen in Figures 5 and 6. Notice that the same color coding has been used as that for the previous band structure plot. Comparing the SB with and without H-MoSSe, we find a reduction of the SB of about 0.6 eV for the Au surface, which is close to the reduction found for T-MoS_2/H-MoS_2. For the Pd surface, the reduction is closer to 0.5 eV, showing a larger cancellation of the internal dipole of MoSSe in this case. We note here that no chemical interactions are taking place at the MoS_2/Au interface, while a considerable chemical bond strength is present at the MoS_2/Pd interface.[51] Thus, this stronger interaction explains the larger cancellation of the internal dipole. This is further stressed by the interlayer distance between H-MoS_2 and Au/Pd. For H-MoS_2/Au, the smallest out-of-plane interlayer distance is 3.4 Å, while it is only 2.0 Å for H-MoS_2/Pd. This shows that the reduction of the out-of-plane dipole moment is rather modest, even at metal surfaces that couple strongly to the Janus monolayer.

Band Alignment at SC/SC Interfaces. Next, we consider how the band alignment can be tuned for SC/SC interfaces. Specifically, we consider bilayer H-MoS_2 with and without a H-MoSSe layer sandwiched in between. The band structure of both cases is depicted in Figure 7 using the same coloring as in Figure 3. We first notice the direct to indirect band gap transition in going from the monolayer to bilayer H-MoS_2, where the VBM is located at the K-point for monolayer H-MoS_2 and at the Γ-point for bilayer H-MoS_2. The two layers hybridize strongly around the Γ point, leading to significant band splitting. In contrast, the interlayer hybridization is negligible around the K-point and the two bands are essentially degenerate. Therefore, the relative band positions at the K-point can be used to deduce the effect of electrostatic

Table 2. Material Properties of All Semiconducting Stable Material Candidates[a]

material	phase	a (Å)	$E_{gap}^{(PBE)}$	$E_{gap}^{(HSE)}$	ΔE_{vac}	$E_{center}^{(PBE)}$	ΔH_f	ΔH_{full}
HfSSe	T	3.71	0.68	1.51	0.01	5.32	−1.47	−0.01
BiSBr	T	4.11	0.93	1.60	0.03	6.13	−0.56	−0.00
SbSeI	T	4.17	1.02	1.53	0.03	5.12	−0.31	−0.00
AsSeI	T	3.93	1.15	1.68	0.04	5.01	−0.16	−0.00
ZrSSe	T	3.74	0.61	1.51	0.05	5.44	−1.48	−0.01
BiSeI	T	4.27	0.65	1.22	0.08	5.68	−0.44	−0.01
SbSBr	T	3.99	1.22	1.84	0.11	5.55	−0.42	0.01
AsSBr	T	3.72	1.39	2.04	0.14	5.44	−0.26	0.03
VBrCl	T	3.75	1.29	3.77	0.21	3.98	−1.01	0.00
BiSCl	T	4.07	1.04	1.74	0.30	6.35	−0.66	−0.03
HfBrI	H	3.64	0.71	1.16	0.36	2.90	−0.89	−0.00
BiSeBr	T	4.19	0.77	1.36	0.36	5.96	−0.58	−0.02
HfBrCl	H	3.43	0.82	1.36	0.36	3.11	−1.34	−0.08
BiTeI	T	4.42	0.45	0.93	0.37	5.43	−0.37	−0.02
VBrI	T	3.99	1.19	3.49	0.37	3.96	−0.66	0.01
ZrBrCl	H	3.49	0.91	1.40	0.38	3.37	−1.51	−0.02
SbSI	T	4.08	1.22	1.78	0.40	5.20	−0.28	0.04
BiSI	T	4.19	0.82	1.47	0.41	5.80	−0.41	0.01
ZrBrI	H	3.70	0.77	1.17	0.42	3.15	−1.09	0.02
SbTeI	T	4.32	0.84	1.32	0.45	4.93	−0.25	−0.02
SbSCl	T	3.94	1.32	1.99	0.46	5.82	−0.51	0.04
VSeTe	H	3.47	0.12	0.75	0.48	5.05	−0.50	0.00
TiBrCl	H	3.38	0.83	1.30	0.48	3.65	−1.33	−0.05
SbSeBr	T	4.08	1.07	1.65	0.48	5.44	−0.44	−0.02
AsSI	T	3.84	1.34	1.95	0.53	5.11	−0.11	0.06
TiBrI	H	3.63	0.68	1.04	0.53	3.43	−0.88	0.04
AsTeI	T	4.09	0.98	1.50	0.55	4.84	−0.12	−0.02
BiSeCl	T	4.15	0.88	1.51	0.61	6.12	−0.68	−0.05
AsSeBr	T	3.82	1.23	1.84	0.63	5.34	−0.29	−0.02
AsSCl	T	3.65	1.54	2.26	0.66	5.71	−0.35	0.05
WSeTe	H	3.43	1.04	1.49	0.68	4.17	−0.32	0.04
MoSeTe	H	3.43	1.14	1.59	0.72	4.47	−0.48	0.02
WSSe	H	3.25	1.40	1.91	0.73	4.50	−0.72	0.00
MoSSe	H	3.25	1.45	1.95	0.75	4.85	−0.81	0.00
ZrClI	H	3.64	0.88	1.32	0.77	3.27	−1.23	0.07
BiTeBr	T	4.34	0.63	1.16	0.78	5.64	−0.50	−0.03
SbSeCl	T	4.03	1.17	1.80	0.81	5.66	−0.54	0.00
CrSSe	H	3.13	0.78	1.24	0.83	5.16	−0.57	−0.00
SbTeBr	T	4.24	1.06	1.62	0.93	5.23	−0.38	−0.03
TiClI	H	3.55	0.75	1.19	0.98	3.58	−1.01	0.07
BiTeCl	T	4.30	0.64	1.20	0.99	5.69	−0.60	−0.05
AsSeCl	T	3.76	1.37	2.06	1.13	5.58	−0.38	0.00
AsTeBr	T	3.98	1.24	1.84	1.18	5.16	−0.24	−0.03
SbTeCl	T	4.19	1.26	1.87	1.20	5.40	−0.47	−0.00
WSTe	H	3.36	1.14	1.68	1.40	4.39	−0.42	0.08
MoSTe	H	3.36	0.99	1.59	1.46	4.64	−0.55	0.06
AsTeCl	T	3.93	1.48	2.15	1.62	5.39	−0.32	0.01

[a]The table contains information about elements, structural phase, in-plane lattice constant [Å], PBE and HSE electronic band gaps [eV], shift in the electrostatic potential across the Janus monolayer [eV], the PBE band gap center relative to vacuum [eV], the heat of formation [eV/atom], and the energy above the convex hull [eV].

shift induced by the Janus monolayer. After insertion of the H-MoSSe layer, we see the splitting of both the conduction and valence bands at the K-point, effectively turning the perfect band alignment into a type-II band alignment. For the H-MoS$_2$/H-MoSSe/H-MoS$_2$ heterostructure, we find a splitting of 0.72 eV, which equals that of the T-MoS$_2$/H-MoSSe/H-MoS$_2$ heterostructure to within about 0.1 eV. Upon further stacking of Janus monolayers in a multilayer structure, recent studies[48,52] have shown that the effect is to

close the band gap, effectively creating a naturally doped p−n junction. This also shows that more than one Janus monolayer can be inserted if a larger shift in the potential is desired.

It is also interesting to see how the interface can be controlled for a SC at a graphene monolayer because of the semimetallic nature of graphene, and we therefore now consider the two graphene/hBN and graphene/H-MoSSe/ hBN systems, where hBN is hexagonal boron nitride. In the same manner as for T-MoS$_2$/H-MoSSe/H-MoS$_2$, we find for

graphene/H-MoSSe/hBN that the bands of hBN shift down by 0.69 relative to the Fermi level after the inclusion of H-MoSSe. This is larger than the shift observed for T-MoS$_2$/H-MoSSe/H-MoS$_2$; however, the comparison to this system is not completely justified because in the latter case the conduction band gets pinned at the Fermi level. It is more interesting to observe that the shift is slightly smaller than that for H-MoS$_2$/H-MoSSe/H-MoS$_2$. This observation can be explained by the larger cancellation of the internal dipole of H-MoSSe by graphene as compared to that of H-MoS$_2$. The larger cancellation of the internal dipole in H-MoSSe on graphene compared to that in H-MoS$_2$ can be understood from the simple point that the S and Se atoms in H-MoSSe are spatially longer separated from the valence states of H-MoS$_2$ (located on the Mo atoms) compared to the states in graphene.

To investigate whether the induced band offset for a SC/Janus/SC vdWH can be predicted from the dipole moment of the freestanding Janus monolayer, we have calculated the band structure for ten different SC/Janus/SC vdWHs. For a given trilayer, the same SC is used on both sides (in analogy to the H-MoS$_2$/H-MoSSe/H-MoS$_2$ structure studied above). For the SC, we use either H-MoS$_2$ or H-MoSe$_2$ (depending on what fits the lattice constant of the Janus monolayer better) and then investigate the effect of changing the Janus monolayer. The Janus monolayers have been chosen such that they span a wide range of dipole moments and their lattice constants are close to those of MoS$_2$ or MoSe$_2$ to minimize the applied strain (see the next section for a discussion on the other Janus monolayers). We also include the structure H-WS$_2$/H-MoSSe/H-WS$_2$ to study the effect of having the same Janus monolayer sandwiched between two different pairs of SCs. We use an in-plane 1 × 1 unit cell for the calculations, and only apply strains to the Janus monolayer to keep the band structure of the SC unchanged. We calculate the shift of the electrostatic potential on either side of the freestanding strained Janus monolayer and the induced band shift of the SC in the SC/Janus/SC configuration. These values are plotted against each other in Figure 8. It is evident that the induced band shift of the bands between the two SCs in the SC/J/SC trilayer can be predicted from the out-of-plane dipole moment of the freestanding Janus monolayer. We note in passing that there is a linear relationship between the shift in electrostatic potential and the internal out-of-plane dipole moment of the Janus monolayer. One more point is worth noticing from Figure 8. First, we see to within 0.01 eV the same band shift for the H-MoS$_2$/MoSSe/H-MoS$_2$ and H-WS$_2$/H-MoSSe/H-WS$_2$, showing that the band shift induced by a Janus monolayer is only little affected by the choice of SC. This effect can be ascribed to the fact that the bands get pinned at the Fermi level, limiting the possible band shift (as seen above for T-MoS$_2$/H-MoSSe/H-MoS$_2$), and second, for metallic heterostructures, the free carriers can more effectively screen the internal dipole of the Janus monolayer. This is also the reason why the shift in potential is smaller in the SC/Janus/SC system compared to the potential shift for the freestanding Janus monolayer for all systems since there will always be a small finite cancellation of the internal dipole moment. If the concentration of S and Se atoms on either side is altered, this will affect the intrinsic out-of-plane dipole moment. To achieve the interface manipulations discussed in this chapter experimentally, we stress that the proposed method is highly dependent on the quality of the transfer of the Janus monolayer into (or onto) the desired structures.

Janus Monolayer Library. To further expand the prospects of the ideas put forward in this study, we consider the set of Janus monolayers from our previous study by Riis-Jensen et al.[53] In this study, an initial set of 216 Janus monolayers with the chemical formula MXY in the H- and T-phases commonly known from the transition-metal dichalcogenides (TMDs) were investigated.

In the following, we adopt the notation used in the Computational 2D Materials Database (C2DB)[54] and in Riis-Jensen et al. and refer to the H and T crystal structures as the MoSSe and BiTeI prototypes. All structures and properties shown in this work are available in the C2DB. The 216 candidate materials were constructed by combinatorial lattice decoration of the MoSSe and BiTeI prototype structures using elements with similar chemical properties. Specifically, for both MoSSe and BiTeI prototypes, all possible combinations using one of the transition metals from groups V, VI, and VII (for the central metal atom in the prototypes) in combination with two elements from either the pnictogens (As, Sb, Bi), or chalcogens (S, Se, Te), or halogens (Cl, Br, I) were considered. In addition, the study considered all combinations using one of the pnictogens as the metal atom, one element from the chalcogens, and one element from the halogens. This makes a total of 108 candidate structures for each prototype. To assess the thermodynamic stability, we calculate the heat of formation and the energy above the convex hull, with the latter defined as the most stable elementaries and binaries, see C2DB[44] for more details. For the dynamic stability, we calculate the Γ-point phonons of the 2 × 2 cell as well as the elastic tensor. An imaginary phonon frequency or negative elastic constant implies a dynamically unstable material. In Figure 9, we show the calculated energy above the convex hull for all 216 structures. The MoSSe prototype is shown in blue, and the BiTeI prototype is shown in orange. Taking uncertainties in the calculated heat of formation into account, we consider a material to be thermodynamically stable if its energy above the convex hull is less than 0.1 eV/atom, as marked by the gray area. Points that have a black circle indicate materials that are dynamically stable and have a finite band gap.

We predict 47 materials that are both semiconducting, thermodynamically stable, and dynamically stable. Out of these, 27 are in the BiTeI phase with a pnictogen atom as the central metal atom, and among these is the experimentally realized BiTeI. For the 47 materials that we find to be semiconducting and predicted to be stable, we plot the shift in the electrostatic potential across the material (red), together with the band edges (white/green boundaries) in Figure 10. All stable semiconducting materials, their in-plane lattice constant, electronic band gap, band gap center, the shift in the potential, heat of formation, and the energy above the convex hull are summarized in Table 2. The large variation in the band edge position and dipole strength underlines the flexibility of the proposed concept.

■ CONCLUSIONS

We have proposed, and critically assessed, a new method for tuning the band lineup at solid-state interfaces by the insertion of a Janus monolayer in the interface. Due to its out-of-plane dipole moment, the Janus monolayer creates a step in the electrostatic potential, which gives rise to a relative shift of the

65

band energies on the two sides of the interface. Our DFT calculations show that the dipole of the Janus monolayer is almost unperturbed by the interface. Consequently, the shift in band alignment is determined by the intrinsic dipole of the Janus monolayer and therefore can be predicted to high accuracy. This important finding is a result of the inertness of the Janus monolayer and stands in contrast to existing methods that involve the formation of chemical bonds and complex charge transfer processes whose effects on the band energies are difficult to predict. Finally, we conducted a computational screening for new Janus monolayers and identified 47 (meta)stable candidates with a large range of out-of-plane dipole moments, providing great flexibility for tuning of band alignment.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpcc.0c01286.

Vacuum-level shift and band gap; band structure of bilayer graphene/hBN; band structure of bilayer graphene/H-MoSSe/hBN (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Kristian S. Thygesen − *CAMD, Department of Physics and Center for Nanostructured Graphene (CNG), Department of Physics, Technical University of Denmark, DK-2800 Kongens Lyngby, Denmark;* ● orcid.org/0000-0001-5197-214X; Email: thygesen@fysik.dtu.dk

Authors

Anders C. Riis-Jensen − *CAMD, Department of Physics, Technical University of Denmark, DK-2800 Kongens Lyngby, Denmark;* ● orcid.org/0000-0002-4914-5820

Simone Manti − *CAMD, Department of Physics, Technical University of Denmark, DK-2800 Kongens Lyngby, Denmark*

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.jpcc.0c01286

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) Novoselov, K. S.; Geim, A. K.; Morozov, S. V.; Jiang, D.; Zhang, Y.; Dubonos, S. V.; Grigorieva, I. V.; Firsov, A. A. Electric Field Effect in Atomically Thin Carbon Films. *Science* **2004**, *306*, 666−669.

(2) Geim, A. K.; Grigorieva, I. V. Van der Waals Heterostructures. *Nature* **2013**, *499*, 419.

(3) Wu, S.; Buckley, S.; Schaibley, J. R.; Feng, L.; Yan, J.; Mandrus, D. G.; Hatami, F.; Yao, W.; Vučković, J.; Majumdar, A.; et al. Monolayer Semiconductor Nanocavity Lasers with Ultralow Thresholds. *Nature* **2015**, *520*, 69.

(4) Ye, Y.; Wong, Z. J.; Lu, X.; Ni, X.; Zhu, H.; Chen, X.; Wang, Y.; Zhang, X. Monolayer Excitonic Laser. *Nat. Photonics* **2015**, *9*, 733.

(5) Salehzadeh, O.; Djavid, M.; Tran, N. H.; Shih, I.; Mi, Z. Optically Pumped Two-Dimensional MoS2 Lasers Operating at Room-Temperature. *Nano Lett.* **2015**, *15*, 5302−5306.

(6) Linghu, J.; Yang, T.; Luo, Y.; Yang, M.; Zhou, J.; Shen, L.; Feng, Y. P. High-Throughput Computational Screening of Vertical 2D van der Waals Heterostructures for High-Efficiency Excitonic Solar Cells. *ACS Appl. Mater. Interfaces* **2018**, *10*, 32142−32150.

(7) Bernardi, M.; Palummo, M.; Grossman, J. C. Semiconducting Monolayer Materials as a Tunable Platform for Excitonic Solar Cells. *ACS Nano* **2012**, *6*, 10082−10089.

(8) Chhowalla, M.; Jena, D.; Zhang, H. Two-Dimensional Semiconductors for Transistors. *Nat. Rev. Mater.* **2016**, *1*, No. 16052.

(9) Tersoff, J. Schottky Barriers and Semiconductor Band Structures. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1985**, *32*, 6968.

(10) Gong, C.; Colombo, L.; Wallace, R. M.; Cho, K. The Unusual Mechanism of Partial Fermi Level Pinning at Metal-MoS2 Interfaces. *Nano Lett.* **2014**, *14*, 1714−1720.

(11) Henck, H.; Ben Aziza, Z.; Zill, O.; Pierucci, D.; Naylor, C. H.; Silly, M. G.; Gogneau, N.; Oehler, F.; Collin, S.; Brault, J.; et al. Interface Dipole and Band Bending in the Hybrid p-n Hetero-junction MoS₂/GaN(0001). *Phys. Rev. B* **2017**, *96*, No. 115312.

(12) Sahoo, A.; Nafday, D.; Paul, T.; Ruiter, R.; Roy, A.; Mostovoy, M.; Banerjee, T.; Saha-Dasgupta, T.; Ghosh, A. Out-of-Plane Interface Dipoles and Anti-Hysteresis in Graphene-Strontium Titanate Hybrid Transistor. *npj 2D Mater. Appl.* **2018**, *2*, No. 9.

(13) Si, C.; Lin, Z.; Zhou, J.; Sun, Z. Controllable Schottky Barrier in GaSe/Graphene Heterostructure: the Role of Interface Dipole. *2D Mater.* **2017**, *4*, No. 015027.

(14) Kim, C.; Moon, I.; Lee, D.; Choi, M. S.; Ahmed, F.; Nam, S.; Cho, Y.; Shin, H.-J.; Park, S.; Yoo, W. J. Fermi Level Pinning at Electrical Metal Contacts of Monolayer Molybdenum Dichalcogenides. *ACS Nano* **2017**, *11*, 1588−1596.

(15) Allain, A.; Kang, J.; Banerjee, K.; Kis, A. Electrical Contacts to Two-Dimensional Semiconductors. *Nat. Mater.* **2015**, *14*, 1195.

(16) Das, S.; Chen, H.-Y.; Penumatcha, A. V.; Appenzeller, J. High Performance Multilayer MoS2 Transistors with Scandium Contacts. *Nano Lett.* **2013**, *13*, 100−105.

(17) Kang, J.; Liu, W.; Banerjee, K. High-Performance MoS2 Transistors with Low-Resistance Molybdenum Contacts. *Appl. Phys. Lett.* **2014**, *104*, No. 093106.

(18) Kang, J.; Liu, W.; Sarkar, D.; Jena, D.; Banerjee, K. Computational Study of Metal Contacts to Monolayer Transition-Metal Dichalcogenide Semiconductors. *Phys. Rev. X* **2014**, *4*, No. 031005.

(19) Guo, Y.; Liu, D.; Robertson, J. 3D Behavior of Schottky Barriers of 2D Transition-Metal Dichalcogenides. *ACS Appl. Mater. Interfaces* **2015**, *7*, 25709−25715.

(20) Guo, Y.; Liu, D.; Robertson, J. Chalcogen Vacancies in Monolayer Transition Metal Dichalcogenides and Fermi Level Pinning at Contacts. *Appl. Phys. Lett.* **2015**, *106*, No. 173106.

(21) Mönch, W. Role of Virtual Gap States and Defects in Metal-Semiconductor Contacts. *Phys. Rev. Lett.* **1987**, *58*, 1260−1263.

(22) Bokdam, M.; Brocks, G.; Katsnelson, M. I.; Kelly, P. J. Schottky Barriers at Hexagonal Boron Nitride/Metal Interfaces: A First-Principles Study. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2014**, *90*, No. 085415.

(23) Anderson, R. Germanium-Gallium Arsenide Heterojunctions [letter to the editor]. *J. Res. Dev.* **1960**, *4*, 283−287.

(24) Amani, M.; Chin, M.; Glen Birdwell, A.; P. O'Regan, T.; Najmaei, S.; Liu, Z.; M. Ajayan, P.; Lou, J.; Dubey, M. Electrical Performance of Monolayer MoS2 Field-Effect Transistors Prepared by Chemical Vapor Deposition. *Appl. Phys. Lett.* **2013**, *102*, No. 193107.

(25) Wang, Q. H.; Kalantar-Zadeh, K.; Kis, A.; Coleman, J. N.; Strano, M. S. Electronics and Pptoelectronics of Two-Dimensional Transition Metal Dichalcogenides. *Nat. Nanotechnol.* **2012**, *7*, 699.

(26) Allain, A.; Kang, J.; Banerjee, K.; Kis, A. Electrical Contacts to Two-Dimensional Semiconductors. *Nat. Mater.* **2015**, *14*, 1195.

(27) Duan, X.; Wang, C.; Pan, A.; Yu, R.; Duan, X. Two-Dimensional Transition Metal Dichalcogenides as Atomically Thin Semiconductors: Opportunities and Challenges. *Chem. Soc. Rev.* **2015**, *44*, 8859−8876.

(28) Lembke, D.; Bertolazzi, S.; Kis, A. Single-Layer MoS2 Electronics. *Acc. Chem. Res.* **2015**, *48*, 100−110.

(29) Chhowalla, M.; Jena, D.; Zhang, H. Two-dimensional Semiconductors for Transistors. *Nat. Rev. Mater.* **2016**, *1*, No. 16052.

(30) Manzeli, S.; Ovchinnikov, D.; Pasquier, D.; V. Yazyev, O.; Kis, A. 2D Transition Metal Dichalcogenides. *Nat. Rev. Mater.* **2017**, *2*, No. 17033.

(31) Ma, Y.; Kou, L.; Li, X.; Dai, Y.; Heine, T. Two-Dimensional Transition Metal Dichalcogenides with a Hexagonal Lattice: Room-Temperature Quantum Spin Hall Insulators. *Phys. Rev. B* **2016**, *93*, No. 035442.

(32) Kou, L.; Frauenheim, T.; Chen, C. Nanoscale Multilayer Transition-Metal Dichalcogenide Heterostructures: Band Gap Modulation by Interfacial Strain and Spontaneous Polarization. *J. Phys. Chem. Lett.* **2013**, *4*, 1730−1736.

(33) Radisavljevic, B.; Radenovic, A.; Brivio, J.; Giacometti, V.; Kis, A. Single-Layer MoS2 Transistors. *Nat. Nanotechnol.* **2011**, *6*, 147.

(34) Min, K.-A.; Park, J.; Wallace, R. M.; Cho, K.; Hong, S. Reduction of Fermi Level Pinning at Au-MoS2 Interfaces by Atomic Passivation on Au Surface. *2D Mater.* **2017**, *4*, No. 015019.

(35) Farmanbar, M.; Brocks, G. Controlling the Schottky Barrier at MoS2/Metal Contacts by Inserting a BN Monolayer. *Phys. Rev. B* **2015**, *91*, No. 161304.

(36) Chuang, S.; Battaglia, C.; Azcatl, A.; McDonnell, S.; Kang, J. S.; Yin, X.; Tosun, M.; Kapadia, R.; Fang, H.; Wallace, R. M.; et al. MoS2 P-type Transistors and Diodes Enabled by High Work Function MoOx Contacts. *Nano Lett.* **2014**, *14*, 1337−1342.

(37) McDonnell, S.; Azcatl, A.; Addou, R.; Gong, C.; Battaglia, C.; Chuang, S.; Cho, K.; Javey, A.; Wallace, R. M. Hole Contacts on Transition Metal Dichalcogenides: Interface Chemistry and Band Alignments. *ACS Nano* **2014**, *8*, 6265−6272.

(38) Chen, J.-R.; Odenthal, P. M.; Swartz, A. G.; Floyd, G. C.; Wen, H.; Luo, K. Y.; Kawakami, R. K. Control of Schottky Barriers in Single Layer MoS2 Transistors with Ferromagnetic Contacts. *Nano Lett.* **2013**, *13*, 3106−3110.

(39) Nam, S.-G.; Cho, Y.; Lee, M.-H.; Shin, K. W.; Kim, C.; Yang, K.; Jeong, M.; Shin, H.-J.; Park, S. Barrier Height Control in Metal/Silicon Contacts with Atomically Thin MoS2 and WS2 Interfacial Layers. *2D Mater.* **2018**, *5*, No. 041004.

(40) Enkovaara, J.; Rostgaard, C.; Mortensen, J. J.; Chen, J.; Dułak, M.; Ferrighi, L.; Gavnholt, J.; Glinsvad, C.; Haikola, V.; Hansen, H. A.; et al. Electronic Structure Calculations with GPAW: a Real-Space Implementation of the Projector Augmented-Wave Method. *J. Phys.: Condens. Matter* **2010**, *22*, No. 253202.

(41) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865−3868.

(42) Monkhorst, H. J.; Pack, J. D. Special Points for Brillouin-Zone Integrations. *Phys. Rev. B* **1976**, *13*, 5188−5192.

(43) Wellendorff, J.; Lundgaard, K. T.; Møgelhøj, A.; Petzold, V.; Landis, D. D.; Nørskov, J. K.; Bligaard, T.; Jacobsen, K. W. Density Functionals for Surface Science: Exchange-Correlation Model Development with Bayesian Error Estimation. *Phys. Rev. B* **2012**, *85*, No. 235149.

(44) Haastrup, S.; Strange, M.; Pandey, M.; Deilmann, T.; Schmidt, P. S.; Hinsche, N. F.; Gjerding, M. N.; Torelli, D.; Larsen, P. M.; Riis-Jensen, A. C.; et al. The Computational 2D Materials Database: high-throughput modeling and discovery of atomically thin crystals. *2D Mater.* **2018**, *5*, No. 042002.

(45) Zhang, J.; Jia, S.; Kholmanov, I.; Dong, L.; Er, D.; Chen, W.; Guo, H.; Jin, Z.; Shenoy, V. B.; Shi, L.; et al. Janus Monolayer Transition-Metal Dichalcogenides. *ACS Nano* **2017**, *11*, 8192−8198.

(46) Lu, A.-Y.; Zhu, A.; Xiao, J.; Chu, C.-P.; Han, Y.; Chu, M.-H.; Cheng, C.-C.; Yang, C.-W.; Wei, R.-H.; Yang, Y.; et al. Janus Monolayers of Transition Metal Dichalcogenides. *Nat. Nanotechnol.* **2017**, *12*, 744−749.

(47) Fülöp, B.; Tajkov, Z.; Pető, J.; Kun, P.; Koltai, J.; Oroszlány, L.; Tóvári, E.; Murakawa, H.; Tokura, Y.; Bordács, S.; et al. Exfoliation of Single Layer BiTeI Flakes. *2D Mater.* **2017**, No. 031013.

(48) Riis-Jensen, A. C.; Pandey, M.; Thygesen, K. S. Efficient Charge Separation in 2D Janus van der Waals Structures with Built-in Electric Fields and Intrinsic p-n Doping. *J. Phys. Chem. C* **2018**, *122*, 24520−24526.

(49) Hu, X.; Wang, Y.; Shen, X.; Krasheninnikov, A. V.; Sun, L.; Chen, Z. 1T phase as an Efficient Hole Injection Layer to TMDs Transistors: A Universal Approach to Achieve p-Type Contacts. *2D Mater.* **2018**, No. 031012.

(50) Jin, C.; Rasmussen, F. A.; Thygesen, K. S. Tuning the Schottky Barrier at the Graphene/MoS2 Interface by Electron Doping: Density Functional Theory and Many-Body Calculations. *J. Phys. Chem. C* **2015**, *119*, 19928−19933.

(51) Fontana, Marcio.; Deppe, Tristan.; Boyd, Anthony K.; Rinzan, Mohamed.; Liu, Amy Y.; Paranjape, Makarand.; Barbara, Paola. Electron-Hole Transport and Photovoltaic Effect in Gated MoS2 Schottky Junctions. *Sci. Rep.* **2013**, *3*, No. 1634.

(52) Guan, Z.; Ni, S.; Hu, S. Tunable Electronic and Optical Properties of Monolayer and Multilayer Janus MoSSe as a Photocatalyst for Solar Water Splitting: A First-Principles Study. *J. Phys. Chem. C* **2018**, *122*, 6209−6216.

(53) Riis-Jensen, A. C.; Deilmann, T.; Olsen, T.; Thygesen, K. S. Classifying the Electronic and Optical Properties of Janus Monolayers. *ACS Nano* **2019**, *13*, 13354−13364.

(54) C2DB Webpage. 2019. https://cmr.fysik.dtu.dk/c2db/c2db.html (accessed Jan 27, 2019).

67

Paper II: Recent progress of the computational 2D materials database (C2DB)

Morten Niklas Gjerding, Alireza Taghizadeh, Asbjørn Rasmussen, Sajid Ali, Fabian Bertoldo, Thorsten Deilmann, Nikolaj Rørbæk Knøsgaard, Mads Kruse, Ask Hjorth Larsen, <u>Simone Manti</u>, Thomas Garm Pedersen, Urko Petralanda, Thorbjørn Skovhus, Mark Kamper Svendsen, Jens Jørgen Mortensen, Thomas Olsen and Kristian Sommer Thygesen

Recent progress of the computational 2D materials database (C2DB)

To cite this article: Morten Niklas Gjerding *et al* 2021 *2D Mater.* **8** 044002

View the article online for updates and enhancements.

2D Materials

Recent progress of the computational 2D materials database (C2DB)

Morten Niklas Gjerding[1] , Alireza Taghizadeh[1,2] , Asbjørn Rasmussen[1] , Sajid Ali[1] , Fabian Bertoldo[1] , Thorsten Deilmann[3] , Nikolaj Rørbæk Knøsgaard[1] , Mads Kruse[1] , Ask Hjorth Larsen[1] , Simone Manti[1] , Thomas Garm Pedersen[2] , Urko Petralanda[1] , Thorbjørn Skovhus[1] , Mark Kamper Svendsen[1] , Jens Jørgen Mortensen[1] , Thomas Olsen[1]

and Kristian Sommer Thygesen[1,*]

[1] Computational Atomic-scale Materials Design (CAMD), Department of Physics, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark
[2] Department of Materials and Production, Aalborg University, 9220 Aalborg Øst, Denmark
[3] Institut für Festkörpertheorie, Westfälische Wilhelms-Universität Münster, 48149 Münster, Germany
* Author to whom any correspondence should be addressed.

E-mail: thygesen@fysik.dtu.dk

Abstract

The Computational 2D Materials Database (C2DB) is a highly curated open database organising a wealth of computed properties for more than 4000 atomically thin two-dimensional (2D) materials. Here we report on new materials and properties that were added to the database since its first release in 2018. The set of new materials comprise several hundred monolayers exfoliated from experimentally known layered bulk materials, (homo)bilayers in various stacking configurations, native point defects in semiconducting monolayers, and chalcogen/halogen Janus monolayers. The new properties include exfoliation energies, Bader charges, spontaneous polarisations, Born charges, infrared polarisabilities, piezoelectric tensors, band topology invariants, exchange couplings, Raman spectra and second harmonic generation spectra. We also describe refinements of the employed material classification schemes, upgrades of the computational methodologies used for property evaluations, as well as significant enhancements of the data documentation and provenance. Finally, we explore the performance of Gaussian process-based regression for efficient prediction of mechanical and electronic materials properties. The combination of open access, detailed documentation, and extremely rich materials property data sets make the C2DB a unique resource that will advance the science of atomically thin materials.

1. Introduction

The discovery of new materials, or new properties of known materials, to meet a specific industrial or scientific requirement, is an exciting intellectual challenge of the utmost importance for our environment and economy. For example, the successful transition to a society based on sustainable energy sources and the realisation of quantum technologies (e.g. quantum computers and quantum communication) depend critically on new materials with novel functionalities. First-principles quantum mechanical calculations, e.g. based on density functional theory (DFT) [1], can predict the properties of materials with high accuracy even before they are made

in the lab. They provide insight into mechanisms at the most fundamental (atomic and electronic) level and can pinpoint and calculate key properties that determine the performance of the material at the macroscopic level. Powered by high-performance computers, atomistic quantum calculations in combination with data science approaches, have the potential to revolutionise the way we discover and develop new materials.

Atomically thin, two-dimensional (2D) crystals represent a fascinating class of materials with exciting perspectives for both fundamental science and technology [2–5]. The family of 2D materials has been growing steadily over the past decade and counts about a hundred materials that have been realised

70

in single-layer or few-layer form [6–10]. While some of these materials, including graphene, hexagonal boron nitride (hBN), and transition metal dichalcogenides (TMDCs), have been extensively studied, the majority have only been scarcely characterised and remain poorly understood. Computational studies indicate that around 1000 already known layered crystals have sufficiently weak interlayer (IL) bonding to allow the individual layers to be mechanically exfoliated [11, 12]. Supposedly, even more 2D materials could be realised beyond this set of already known crystals. Adding to this the possibility of stacking individual 2D layers (of the same or different kinds) into ultrathin van der Waals (vdW) crystals [13], and tuning the properties of such structures by varying the relative twist angle between adjacent layers [14, 15] or intercalating atoms into the vdW gap [16, 17], it is clear that the prospects of tailor made 2D materials are simply immense. To support experimental efforts and navigate the vast 2D materials space, first-principles calculations play a pivotal role. In particular, FAIR[5] [18] databases populated by high-throughput calculations can provide a convenient overview of known materials and point to new promising materials with desired (predicted) properties. Such databases are also a fundamental requirement for the successful introduction and deployment of artificial intelligence in materials science.

Many of the unique properties exhibited by 2D materials have their origin in quantum confinement and reduced dielectric screening. These effects tend to enhance many-body interactions and lead to profoundly new phenomena such as strongly bound excitons [19–21] with nonhydrogenic Rydberg series [22–24], phonons and plasmons with anomalous dispersion relations [25, 26], large dielectric band structure renormalisations [27, 28], unconventional Mott insulating and superconducting phases [14, 15], and high-temperature exciton condensates [29]. Recently, it has become clear that long range magnetic order can persist [30, 31] and (in-plane) ferroelectricity even be enhanced [32], in the single layer limit. In addition, first-principles studies of 2D crystals have revealed rich and abundant topological phases [33, 34]. The peculiar physics ruling the world of 2D materials entails that many of the conventional theories and concepts developed for bulk crystals break down or require special treatments when applied to 2D materials [26, 35, 36]. This means that computational studies must be performed with extra care, which in turn calls for well-organised and well-documented 2D property data sets that can form the basis for the development, benchmarking, and consolidation of physical theories and numerical implementations.

The Computational 2D Materials Database (C2DB) [6, 37] is a highly curated and fully open database containing elementary physical properties of around 4000 2D monolayer crystals. The data has been generated by automatic high-throughput calculations at the level of DFT and many-body perturbation theory as implemented in the GPAW [38, 39] electronic structure code. The computational workflow is constructed using the atomic simulation recipes (ASR) [40]—a recently developed Python framework for high-throughput materials modelling building on the atomic simulation environment (ASE) [41]—and managed/executed using the MyQueue task scheduler [42].

The C2DB differentiates itself from existing computational databases of bulk [43–45] and low-dimensional [11, 12, 46–50] materials, by the large number of physical properties available, see table 1. The use of beyond-DFT theories for excited state properties (GW band structures and Bethe–Salpeter equation (BSE) absorption for selected materials) and Berry-phase techniques for band topology and polarisation quantities (spontaneous polarisation, Born charges, piezoelectric tensors), are other unique features of the database.

The C2DB can be downloaded in its entirety or browsed and searched online. As a new feature, all data entries presented on the website are accompanied by a clickable help icon that presents a scientific documentation ('what does this piece of data describe?') and technical documentation ('how was this piece of data computed?'). This development enhances the usability of the database and improves the reproducibility and provenance of the data contained in C2DB. As another novelty it is possible to download all property data pertaining to a specific material or a specific type of property, e.g. the band gap, for all materials thus significantly improving data accessibility.

In this paper, we report on the significant C2DB developments that have taken place during the past two years. These developments can be roughly divided into four categories: (1) General updates of the workflow used to select, classify, and stability assess the materials. (2) Computational improvements for properties already described in the 2018 paper. (3) New properties. (4) New materials. The developments, described in four separate sections, cover both original work and review of previously published work. In addition, we have included some outlook discussions of ongoing work. In the last section we illustrate an application of statistical learning to predict properties directly from the atomic structure.

71

2. Selection, classification, and stability

Figure 1 illustrates the workflow behind the C2DB. In this section we describe the first part of the workflow

---

[5] FAIR data are data which meet principles of findability, accessibility, interoperability, and reusability.

Table 1. Properties calculated by the C2DB monolayer workflow. The computational method and the criteria used to decide whether the property should be evaluation for a given material is also shown. A '*' indicates that spin–orbit coupling (SOC) is included. All calculations are performed with the GPAW code using a plane wave basis except for the Raman calculations, which employ a double-zeta polarised basis of numerical atomic orbitals [51].

Property	Method	Criteria	Count
Bader charges	PBE	None	3809
Energy above convex hull	PBE	None	4044
Heat of formation	PBE	None	4044
Orbital projected band structure	PBE	None	2487
Out-of-plane dipole	PBE	None	4044
Phonons (Γ and BZ corners)	PBE	None	3865
Projected density of states	PBE	None	3332
Stiffness tensor	PBE	None	3968
Exchange couplings	PBE	Magnetic	538
Infrared polarisability	PBE	$E_{\mathrm{gap}}^{\mathrm{PBE}} > 0$	784
Second harmonic generation	PBE	$E_{\mathrm{gap}}^{\mathrm{PBE}} > 0$, non-magnetic, non-centrosymmetric	375
Electronic band structure PBE	PBE*	None	3496
Magnetic anisotropies	PBE*	Magnetic	823
Deformation potentials	PBE*	$E_{\mathrm{gap}}^{\mathrm{PBE}} > 0$	830
Effective masses	PBE*	$E_{\mathrm{gap}}^{\mathrm{PBE}} > 0$	1272
Fermi surface	PBE*	$E_{\mathrm{gap}}^{\mathrm{PBE}} = 0$	2505
Plasma frequency	PBE*	$E_{\mathrm{gap}}^{\mathrm{PBE}} = 0$	3144
Work function	PBE*	$E_{\mathrm{gap}}^{\mathrm{PBE}} = 0$	4044
Optical polarisability	RPA@PBE	None	3127
Electronic band structure	HSE06@PBE*	None	3155
Electronic band structure	G_0W_0@PBE*	$E_{\mathrm{gap}}^{\mathrm{PBE}} > 0$, $N_{\mathrm{atoms}} < 5$	357
Born charges	PBE, Berry phase	$E_{\mathrm{gap}}^{\mathrm{PBE}} > 0$	639
Raman spectrum	PBE, LCAO basis set	Non-magnetic, dyn. stable	708
Piezoelectric tensor	PBE, Berry phase	$E_{\mathrm{gap}}^{\mathrm{PBE}}$, non-centrosym.	353
Optical absorbance	BSE@G_0W_0*	$E_{\mathrm{gap}}^{\mathrm{PBE}} > 0$, $N_{\mathrm{atoms}} < 5$	378
Spontaneous polarisation	PBE, Berry phase	$E_{\mathrm{gap}}^{\mathrm{PBE}} > 0$, nearly centrosym. polar space group	151
Topological invariants	PBE*, Berry phase	$0 < E_{\mathrm{gap}}^{\mathrm{PBE}} < 0.3$ eV	242

Figure 1. The workflow behind the C2DB. After the structural relaxation, the dimensionality of the material is checked and it is verified that the material is not already present in the database. Next, the material is classified according to its chemical composition, crystal structure, and magnetic state. Finally, the thermodynamic and dynamic stabilities are assessed from the energy above the convex hull and the sign of the minimum eigenvalues of the dynamical matrix and stiffness tensor. Unstable materials are stored in the database; stable materials are subject to the property workflow. The C2DB monolayer database is interlinked with databases containing structures and properties of multilayer stacks and point defects in monolayers from the C2DB.

until the property calculations (red box), focusing on aspects related to selection criteria, classification, and stability assessment, that have been changed or updated since the 2018 paper.

2.1. Structure relaxation

Given a prospective 2D material, the first step is to carry out a structure optimisation. This calculation is performed with spin polarisation and with the symmetries of the original structure enforced. The latter is done to keep the highest level of control over the resulting structure by avoiding 'uncontrolled' symmetry breaking distortions. The prize to pay is a higher risk of generating dynamically unstable structures.

2.2. Selection: dimensionality analysis

A dimensionality analysis [52] is performed to identify and filter out materials that have disintegrated into non-2D structures during relaxation. Covalently bonded clusters are identified through an analysis of the connectivity of the structures where two atoms are considered to belong to the same cluster if their distance is less than some scaling of the sum of their covalent radii, i.e. $d < k(r_i^{cov} + r_j^{cov})$, where i and j are atomic indices. A scaling factor of $k = 1.35$ was determined empirically. Only structures that consist of a single 2D cluster after relaxation are further processed. Figure 2 shows three examples (graphene, Ge_2Se_2, and Pb_2O_6) of structures and their cluster dimensionalities before and after relaxation. All structures initially consist of a single 2D cluster, but upon relaxation Ge_2Se_2 and Pb_2O_6 disintegrate into two 2D clusters as well as one 2D and two 0D clusters, respectively. On the other hand, the relaxation of graphene decreases the in-plane lattice constant but does not affect the dimensionality. According to the criterion defined above only graphene will enter the database.

2.3. Selection: ranking similar structures

Maintaining a high-throughput database inevitably requires a strategy for comparing similar structures and ranking them according to their relevance. In particular, this is necessary in order to identify different representatives of the same material e.g. resulting from independent relaxations, and thereby avoid duplicate entries and redundant computations. The C2DB strategy to this end involves a combination of structure clustering and Pareto analysis.

First, a single-linkage clustering algorithm is used to group materials with identical reduced chemical formula and 'similar' atomic configurations. To quantify configuration similarity a slightly modified version of PyMatGen's [53] distance metric is employed where the cell volume normalisation is removed to make it applicable to 2D materials surrounded by vacuum. Roughly speaking, the metric measures the maximum distance an atom must be moved (in units of Å) in order to match the two

Before relaxation   **After relaxation**

$N_{3D} = 0, N_{2D} = 1$ $N_{3D} = 0, N_{2D} = 1$
$N_{1D} = 0, N_{0D} = 0$ $N_{1D} = 0, N_{0D} = 0$

$N_{3D} = 0, N_{2D} = 1$ $N_{3D} = 0, N_{2D} = 2$
$N_{1D} = 0, N_{0D} = 0$ $N_{1D} = 0, N_{0D} = 0$

$N_{3D} = 0, N_{2D} = 1$ $N_{3D} = 0, N_{2D} = 1$
$N_{1D} = 0, N_{0D} = 0$ $N_{1D} = 0, N_{0D} = 2$
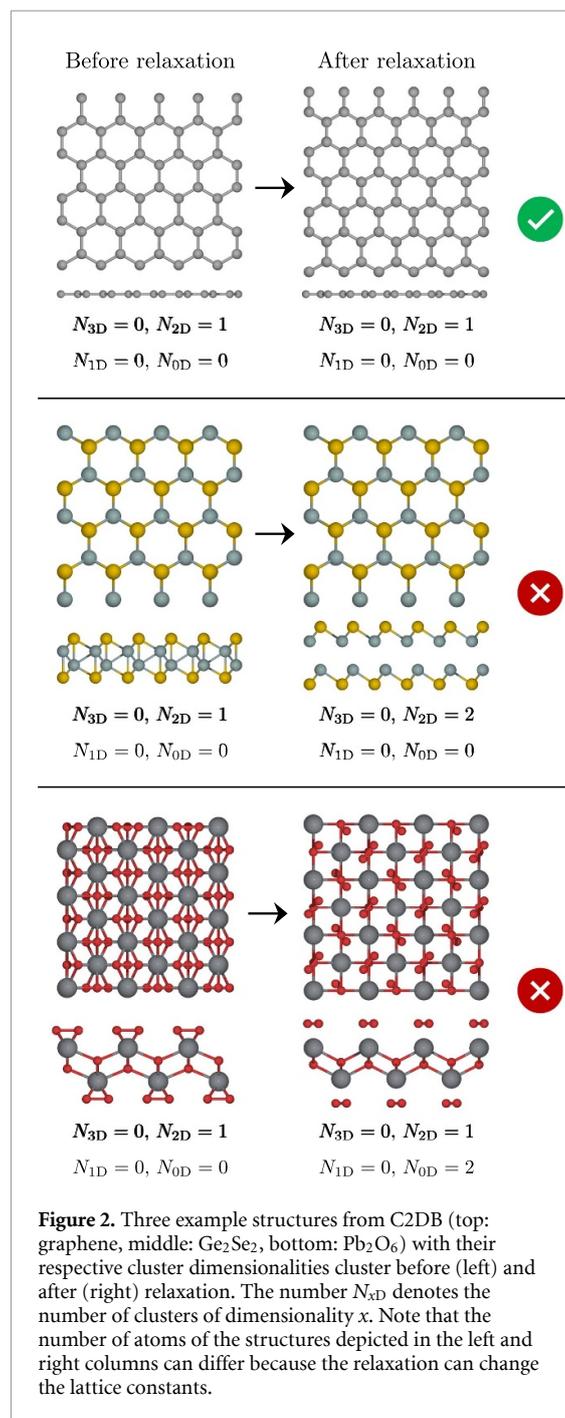
Figure 2. Three example structures from C2DB (top: graphene, middle: Ge_2Se_2, bottom: Pb_2O_6) with their respective cluster dimensionalities cluster before (left) and after (right) relaxation. The number N_{xD} denotes the number of clusters of dimensionality x. Note that the number of atoms of the structures depicted in the left and right columns can differ because the relaxation can change the lattice constants.

atomic configurations. Two atomic configurations belong to the same cluster if their distance is below an empirically determined threshold of 0.3 Å.

At this point, the simplest strategy would be to remove all but the most stable compound within a cluster. However, this procedure would remove many high symmetry crystals for which a more stable distorted version exists. For example, the well known T-phase of MoS_2 would be removed in favour of the more stable T$'$-phase. This is undesired as high-symmetry structures, even if dynamically unstable at $T = 0$, may provide useful information and might in fact become stabilised at higher temperatures [54]. Therefore, the general strategy adopted for the C2DB,

Figure 3. Illustration of the Pareto analysis used to filter out duplicates or irrelevant structures from the C2DB. All points represent materials with the same reduced chemical formula (in this case ReS_2) that belong to the same cluster defined by the structure metric. Only structures lying on the $(N, \Delta H)$-Pareto front are retained (black circles) while other materials are excluded (red circles). The philosophy behind the algorithm is to keep less stable materials if they contain fewer atoms per unit cell than more stable materials and thus represent structures of higher symmetry.

is to keep a material that is less stable than another material of the same cluster if it has fewer atoms in its primitive unit cell (and thus typically higher symmetry). Precisely, materials within a given cluster are kept only if they represent a defining point of the $(N, \Delta H)$-Pareto front, where N is the number of atoms in the unit cell and ΔH is the heat of formation. A graphical illustration of the Pareto analysis is shown in figure 3 for the case of ReS_2.

2.4. Classification: crystal structure
The original C2DB employed a *crystal prototype* classification scheme where specific materials were promoted to prototypes and used to label groups of materials with the same or very similar crystal structure. This approach was found to be difficult to maintain (as well as being non-transparent). Instead, materials are now classified according to their *crystal type* defined by the reduced stoichiometry, space group number, and the alphabetically sorted labels of the occupied Wyckoff positions. As an example, MoS_2 in the H-phase has the crystal type: AB2-187-bi.

2.5. Classification: magnetic state
In the new version of the C2DB, materials are classified according to their magnetic state as either *non-magnetic* or *magnetic*. A material is considered magnetic if any atom has a local magnetic moment greater than $0.1\,\mu_B$.

In the original C2DB, the *magnetic* category was further subdivided into ferromagnetic (FM) and antiferromagnetic (AFM). But since the simplest antiferromagnetically ordered state typically does not represent the true ground state, all material entries with an AFM state have been removed from the C2DB and replaced by the material in its FM state. Although the latter is less stable, it represents a

more well defined state of the material. Crucially, the nearest neighbour exchange couplings for all magnetic materials have been included in the C2DB (see section 5.8). This enables a more detailed and realistic description of the magnetic order via the Heisenberg model. In particular, the FM state of a material is not expected to represent the true magnetic ground if the exchange coupling $J < 0$.

2.6. Stability: thermodynamic
The heat of formation, ΔH, of a compound is defined as its energy per atom relative to its constituent elements in their standard states [55]. The thermodynamic stability of a compound is evaluated in terms of its energy above the *convex hull*, ΔH_{hull}, which gives the energy of the material relative to other competing phases of the same chemical composition, including mixed phases [6], see figure 4 for an example. Clearly, ΔH_{hull} depends on the pool of reference phases, which in turn defines the convex hull. The original C2DB employed a pool of reference phases comprised by 2807 elemental and binary bulk crystals from the convex hull of the Open Quantum Materials Database (OQMD) [55]. In the new version, this set has been extended by approximately 6783 ternary bulk compounds from the convex hull of OQMD, making a total of 9590 stable bulk reference compounds.

As a simple indicator for the thermodynamic stability of a material, the C2DB employs three labels (low, medium, high) as defined in table 2. These indicators are unchanged from the original version of the C2DB. In particular, the criterion $\Delta H_{hull} < 0.2$ eV atom^{-1}, defining the most stable category, was established based on an extensive analysis of 55 experimentally realised monolayer crystals [6].
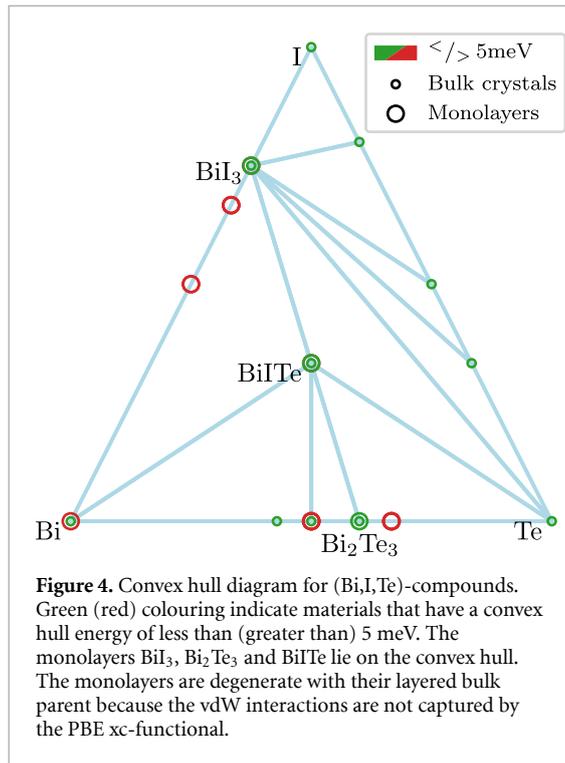
74

Figure 4. Convex hull diagram for (Bi,I,Te)-compounds. Green (red) colouring indicate materials that have a convex hull energy of less than (greater than) 5 meV. The monolayers BiI_3, Bi_2Te_3 and BiITe lie on the convex hull. The monolayers are degenerate with their layered bulk parent because the vdW interactions are not captured by the PBE xc-functional.

Table 2. Thermodynamic stability indicator assigned to all materials in the C2DB. ΔH and ΔH_{hull} denote the heat of formation and energy above the convex hull, respectively.

Thermodynamic stability indicator	Criterion (eV atom^{-1})
Low	$\Delta H > 0.2$
Medium	$\Delta H < 0.2$ and $\Delta H_{hull} > 0.2$
High	$\Delta H < 0.2$ and $\Delta H_{hull} < 0.2$

It should be emphasised that the energies of both monolayers and bulk reference crystals are calculated with the Perdew-Burke-Ernzerhof (PBE) xc-functional [56]. This implies that some inaccuracies must be expected, in particular for materials with strongly localised d-electrons, e.g. certain transition metal oxides, and materials for which dispersive interactions are important, e.g. layered van der Waals crystals. The latter implies that the energy of a monolayer and its layered bulk parent (if such exists in the pool of references) will have the same energy. For further details and discussions see reference [6].

2.7. Stability: dynamical

Dynamically stable materials are situated at a local minimum of the potential energy surface and are thus stable to small structural perturbations. Structures resulting from DFT relaxations can end up in saddle point configurations because of imposed symmetry constraints or an insufficient number of atoms in the unit cell.

In C2DB, the dynamical stability is assessed from the signs of the minimum eigenvalues of (1) the stiffness tensor (see section 3.1) and (2) the Γ-point

Hessian matrix for a supercell containing 2×2 repetitions of the unit cell (the structure is not relaxed in the 2×2 supercell). If one of these minimal eigenvalues is negative the material is classified as dynamically unstable. This indicates that the energy can be reduced by displacing an atom and/or deforming the unit cell, respectively. The use of two categories for dynamical stability, i.e. stable/unstable, differs from the original version of the C2DB where an intermediate category was used for materials with negative but numerically small minimal eigenvalue of either the Hessian or stiffness tensors.

3. Improved property methodology

The new version of the C2DB has been generated using a significantly extended and improved workflow for property evaluations. This section focuses on improvements relating to properties that were already present in the original version of the C2DB while new properties are discussed in the next section.

3.1. Stiffness tensor

The stiffness tensor, C, is a rank-4 tensor that relates the stress of a material to the applied strain. In Mandel notation (a variant of Voigt notation) C is expressed as an $N \times N$ matrix relating the N independent components of the stress and strain tensors. For a 2D material $N = 3$ and the tensor takes the form:

$$\mathbf{C} = \begin{bmatrix} C_{xxxx} & C_{xxyy} & \sqrt{2}C_{xxxy} \\ C_{xxyy} & C_{yyyy} & \sqrt{2}C_{yyxy} \\ \sqrt{2}C_{xxxy} & \sqrt{2}C_{yyxy} & 2C_{xyxy} \end{bmatrix}, \quad (1)$$

where the indices on the matrix elements refer to the rank-4 tensor. The factors multiplying the tensor elements account for their multiplicities in the full rank-4 tensor. In the C2DB workflow, C is calculated as a finite difference of the stress under an applied strain with full relaxation of atomic coordinates. A negative eigenvalue of C signals a dynamical instability, see section 2.7.

In the first version of the C2DB only the diagonal elements of the stiffness tensor were calculated. The new version also determines the shear components such that the full 3×3 stiffness tensor is now available. This improvement also leads to a more accurate assessment of dynamical stability [57].

3.2. Effective masses with parabolicity estimates

For all materials with a finite band gap the effective masses of electrons and holes are calculated for bands within 100 meV of the conduction band minimum and valence band maximum, respectively. The Hessian matrices at the band extrema (PBE) are determined by fitting a second order polynomium to the PBE band structure including SOC, and the effective masses are obtained by subsequent diagonalisation of the Hessian. The main fitting-procedure is unaltered
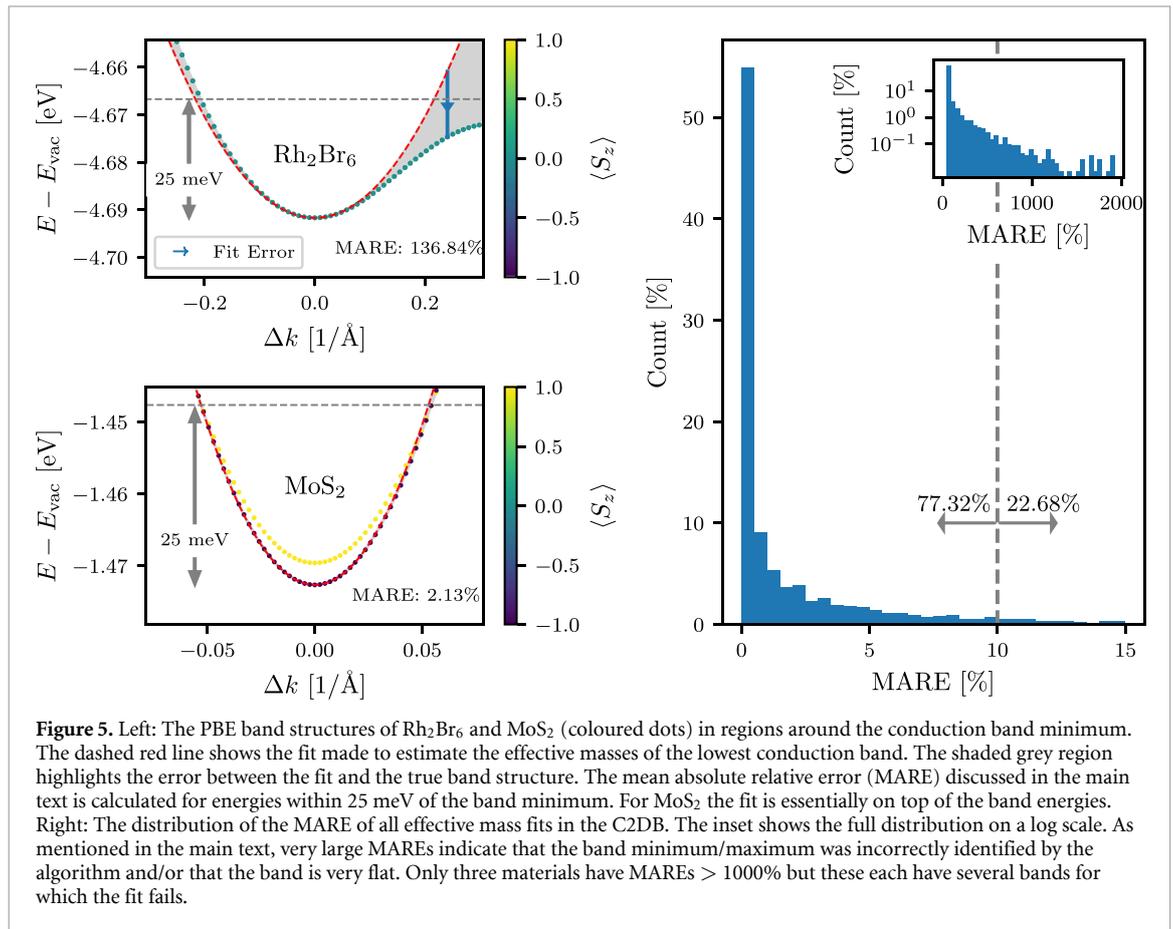
Figure 5. Left: The PBE band structures of Rh_2Br_6 and MoS_2 (coloured dots) in regions around the conduction band minimum. The dashed red line shows the fit made to estimate the effective masses of the lowest conduction band. The shaded grey region highlights the error between the fit and the true band structure. The mean absolute relative error (MARE) discussed in the main text is calculated for energies within 25 meV of the band minimum. For MoS_2 the fit is essentially on top of the band energies. Right: The distribution of the MARE of all effective mass fits in the C2DB. The inset shows the full distribution on a log scale. As mentioned in the main text, very large MAREs indicate that the band minimum/maximum was incorrectly identified by the algorithm and/or that the band is very flat. Only three materials have MAREs > 1000% but these each have several bands for which the fit fails.

from the first version of C2DB, but two important improvements have been made.

The first improvement consists in an additional k-mesh refinement step for better localisation of the BE in the Brillouin zone. After the location of the BE has been estimated based on a uniformly sampled band structure with k-point density of 12 Å, another one-shot calculation is performed with a denser k-mesh around the estimated BE positions. This ensures a more accurate and robust determination of the location of the BE, which can be important in cases with a small but still significant spin–orbit splitting or when the band is very flat or non-quadratic around the BE. The second refinement step is the same as in the first version of C2DB, i.e. the band energies are calculated on a highly dense k-mesh in a small disc around the BE, and the Hessian is obtained by fitting the band energies in the range up to 1 meV from the BE.

The second improvement is the calculation of the mean absolute relative error (MARE) of the polynomial fit in a 25 meV range from the BE. The value of 25 meV corresponds to the thermal energy at room temperature and is thus the relevant energy scale for many applications. To make the MARE independent of the absolute position of the band we calculate the average energy of the band over the 25 meV and compare the deviation of the fit to this energy scale. The MARE provides a useful measure of the parabolicity

of the energy bands and thus the validity of the effective mass approximation over this energy scale.

Figure 5 shows two examples of band structures with the effective mass fits and corresponding fit errors indicated. Additionally, the distribution of MARE for all the effective mass fits in the C2DB are presented. Most materials have an insignificant MARE, but a few materials have very large errors. Materials with a MARE above a few hundreds of percentages fall into two classes. For some materials the algorithm does not correctly find the position of the BE. An example is Ti_2S_2 in the space group C2/m. For others, the fit and BE location are both correct, but the band flattens away from the BE which leads to a large MARE as is the case for Rh_2Br_6 shown in the figure or Cl_2Tl_2 in the space group P-1. In general a small MARE indicates a parabolic band while materials with large MARE should be handled on a case-by-case basis.

3.3. Orbital projected band structure
To facilitate a state-specific analysis of the PBE Kohn–Sham wave functions, an orbital projected band structure (PBS) is provided to complement the projected density of states (PDOS). In the PAW methodology, the all-electron wave functions are projected onto atomic orbitals inside the augmentation spheres centred at the position of each atom. The PBS resolves these atomic orbital contributions to the

Figure 6. Orbital projected band structure and orbital projected density of states of MoS$_2$ in the H-phase. The pie chart symbols indicate the fractional atomic orbital character of the Kohn–Sham wave functions.
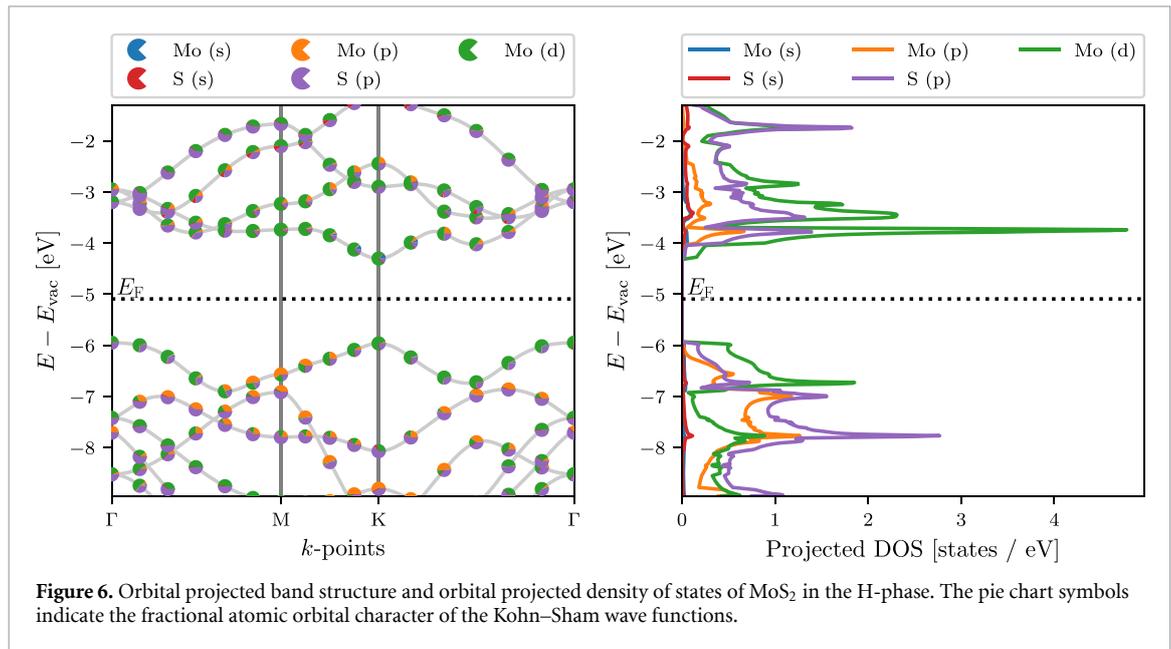
wave functions as a function of band and k-point whereas the PDOS resolves the atomic orbital character of the total density of states as a function of energy. The SOC is not included in the PBS or PDOS, as its effect is separately visualised by the spin-PBS also available in the C2DB.

As an example, figure 6 shows the PBS (left) and PDOS (right) of monolayer MoS$_2$ calculated with PBE. The relative orbital contribution to a given Bloch state is indicated by a pie chart symbol. In the present example, one can deduce from the PBS that even though Mo-p orbitals and S-p orbitals contribute roughly equally to the DOS in the valence band, the Mo-p orbital contributions are localised to a region in the BZ around the M-point, whereas the S-p orbitals contribute throughout the entire BZ.

3.4. Corrected G_0W_0 band structures
The C2DB contains G_0W_0 quasiparticle (QP) band structures of 370 monolayers covering 14 different crystal structures and 52 chemical elements. The details of these calculations can be found in the original C2DB paper [6]. A recent in-depth analysis of the 61.716 G_0W_0 data points making up the QP band structures led to several important conclusions relevant for high-throughput G_0W_0 calculations. In particular, it identified the linear QP approximation as a significant error source in standard G_0W_0 calculations and proposed an extremely simple correction scheme (the *empirical Z* (empZ) scheme), that reduces this error by a factor of two on average.

The empZ scheme divides the electronic states into two classes according to the size of the QP weight, Z. States with $Z \in [0.5, 1.0]$ are classified as QP consistent (QP-c) while states with $Z \notin [0.5, 1.0]$ are classified as QP inconsistent (QP-ic). With this definition, QP-c states will have at least half of their spectral weight in the QP peak. The distribution of

the $60.000+$ Z-values is shown in figure 7. It turns out that the linear approximation to the self-energy, which is the gist of the QP approximation, introduces significantly larger errors for QP-ic states than for QP-c states. Consequently, the empZ method replaces the calculated Z of QP-ic states with the mean of the Z-distribution, $Z_0 \approx 0.75$. This simple replacement reduces the average error of the linear approximation from 0.11 to 0.06 eV.

An illustration of the method applied to MoS$_2$ is shown in figure 7. The original uncorrected G_0W_0 band structure is shown in blue while the empZ corrected band structure is shown in orange. MoS$_2$ has only one QP-ic state in the third conduction band at the K-point. Due to a break-down of the QP approximation for this state, the G_0W_0 correction is greatly overestimated leading to a local discontinuity in the band structure. The replacement of Z by Z_0 for this particular state resolves the problem. All G_0W_0 band structures in the C2DB are now empZ corrected.

3.5. Optical absorbance
In the first version of the C2DB, the optical absorbance was obtained from the simple expression [6]

$$A(\omega) \approx \frac{\omega \mathrm{Im} \alpha^{2D}(\omega)}{\epsilon_0 c}, \qquad (2)$$

where α^{2D} is the long wavelength limit of the in-plane sheet polarisability density (note that the equation is written here in SI units). The sheet polarisability is related to the sheet conductivity via $\sigma^{2D}(\omega) = -i\omega\alpha^{2D}(\omega)$. The expression (2) assumes that the electric field inside the layer equals the incoming field (i.e. reflection is ignored), and hence, it may overestimate the absorbance.

In the new version, the absorbance is evaluated from $A = 1 - R - T$, where R and T are the reflected and transmitted powers of a plane wave at normal
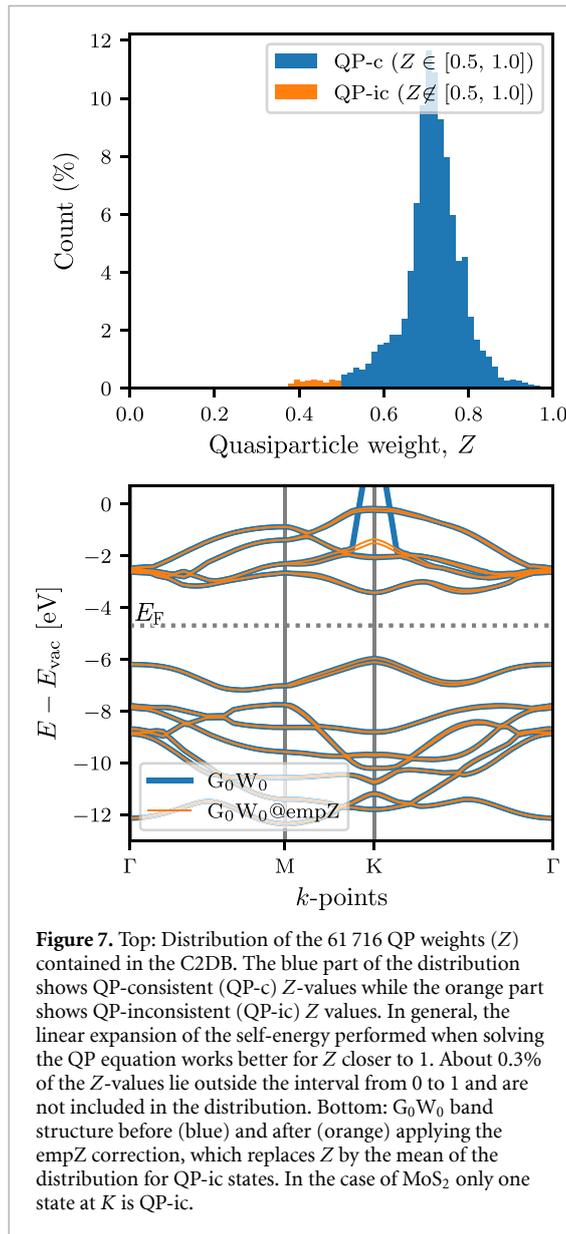
Figure 7. Top: Distribution of the 61 716 QP weights (Z) contained in the C2DB. The blue part of the distribution shows QP-consistent (QP-c) Z-values while the orange part shows QP-inconsistent (QP-ic) Z values. In general, the linear expansion of the self-energy performed when solving the QP equation works better for Z closer to 1. About 0.3% of the Z-values lie outside the interval from 0 to 1 and are not included in the distribution. Bottom: G_0W_0 band structure before (blue) and after (orange) applying the empZ correction, which replaces Z by the mean of the distribution for QP-ic states. In the case of MoS_2 only one state at K is QP-ic.

incidence, respectively. These can be obtained from the conventional transfer matrix method applied to a monolayer suspended in vacuum. The 2D material is here modelled as an infinitely thin layer with a sheet conductivity. Alternatively, it can be modelled as quasi-2D material of thickness d with a 'bulk' conductivity of $\sigma = \sigma^{2D}/d$ [58], but the two approaches yield very similar results, since the optical thickness of a 2D material is much smaller than the optical wavelength. Within this model, the expression for the absorbance of a suspended monolayer with the sheet conductivity σ^{2D} reads:

$$A(\omega) = \text{Re} \left\{ \sigma^{2D}(\omega)\eta_0 \right\} \left| \frac{2}{2 + \sigma^{2D}(\omega)\eta_0} \right|^2, \quad (3)$$

78

where $\eta_0 = 1/(\epsilon_0 c) \approx 377 \ \Omega$ is the vacuum impedance.

If the light–matter interaction is weak, i.e. $|\sigma^{2D}\eta_0| \ll 1$, equation (3) reduces to equation (2).

Nonetheless, due the strong light–matter interaction in some 2D materials, this approximation is not reliable in general. In fact, it can be shown that the maximum possible absorption from equation (3) is 50%, which is known as the upper limit of light absorption in thin films [59]. This limit is not guaranteed by equation (2), which can even yield an absorbance above 100%.

As an example, figure 8 shows the absorption spectrum of monolayer MoS_2 for in- and out-of-plane polarised light as calculated with the exact equation (3) and the approximate equation (2), respectively. In all cases the sheet polarisability is obtained from the BSE to account for excitonic effects [6]. For weak light–matter interactions, e.g. for the z-polarised light, the two approaches agree quite well, but noticeable differences are observed in regions with stronger light–matter interaction.

4. New materials in the C2DB

In this section we discuss the most significant extensions of the C2DB in terms of new materials. The set of materials presented here is not complete, but represents the most important and/or well defined classes. The materials discussed in sections 4.1 and 4.2 (MXY Janus monolayers and monolayers extracted from experimental crystal structure databases) are already included in the C2DB. The materials described in sections 4.3 and 4.4 (homo-bilayers and monolayer point defect systems) will soon become available as separate C2DB-interlinked databases.

4.1. MXY Janus monolayers

The class of TMDC monolayers of the type MX_2 (where M is the transition metal and X is a chalcogen) exhibits a large variety of interesting and unique properties and has been widely discussed in the literature [60]. Recent experiments have shown that it is not only possible to synthesise different materials by changing the metal M or the chalcogen X, but also by exchanging the X on one side of the layer by another chalcogen (or halogen) [61–63]. This results in a class of 2D materials known as MXY Janus monolayers with broken mirror symmetry and finite out-of-plane dipole moments. The prototypical MXY crystal structures are shown in figure 9 for the case of MoSSe and BiTeI, which have both been experimentally realised [61–63]. Adopting the nomenclature from the TMDCs, the crystal structures are denoted as H- or T-phase, depending on whether X and Y atoms are vertically aligned or displaced, respectively.

In a recent work [64], the C2DB workflow was employed to scrutinise and classify the basic electronic and optical properties of 224 different MXY Janus monolayers. All data from the study is available in the C2DB. Here we provide a brief discussion of the Rashba physics in these materials and refer the

Figure 8. Optical absorption of standalone monolayer MoS$_2$ for *x/y*-polarisation (left) and *z*-polarisation (right) at normal incident in the BSE framework, obtained using equation (2) (blue) or equation (3) (orange). The crystal structure cross-sectional views are shown in the inset with the definition of directions.

Figure 9. Atomic structure of the MXY Janus monolayers in the H-phase (left) and T-phase (right). The two prototype materials MoSSe and BiTeI are examples of experimentally realised monolayers adopting these crystal structures (not to scale).

interested reader to [64] for more details and analysis of other properties.

A key issue when considering hypothetical materials, i.e. materials not previously synthesised, is their stability. The experimentally synthesised MoSSe and BiTeI are both found to be dynamically stable and lie within 10 meV of the convex hull confirming their thermodynamic stability. Out of the 224 initial monolayers 93 are classified as stable according to the C2DB criteria (dynamically stable and $\Delta H_{hull} < 0.2$ eV atom^{-1}). Out of the 93 stable materials, 70 exhibit a finite band gap when computed with the PBE xc-functional.

The Rashba effect is a momentum dependent splitting of the band energies of a 2D semiconductor in the vicinity of a band extremum arising due to the combined effect of spin–orbit interactions and a broken crystal symmetry in the direction perpendicular to the 2D plane. The simplest model used to describe the Rashba effect is a 2D electron gas in a perpendicular electric field (along the *z*-axis). Close to

the band extremum, the energy of the two spin bands is described by the Rashba Hamiltonian [65, 66]:

$$H = \alpha_R (\boldsymbol{\sigma} \times \mathbf{k}) \cdot \hat{\mathbf{e}}_z, \qquad (4)$$

where $\boldsymbol{\sigma}$ is the vector of Pauli matrices, $\mathbf{k} = \mathbf{p}/\hbar$ is the wave number, and the Rashba parameter is proportional to the electric field strength, $\alpha_R \propto E_0$.

Although the Rashba Hamiltonian is only meant as a qualitative model, it is of interest to test its validity on the Janus monolayers. The electric field of the Rashba model is approximately given by $E_0 = \Delta V_{vac}/d$, where ΔV_{vac} is the shift in vacuum potential on the two sides of the layer (see left inset of figure 10) and d is the layer thickness. Assuming a similar thickness for all monolayers, the electric field is proportional to the potential shift. Not unexpected, the latter is found to correlate strongly with the difference in electronegativity of the X and Y atoms, see left panel of figure 10.

The Rashba energy, E_R, can be found by fitting $E(k) = \hbar^2 k^2/2m^* \pm \alpha_R k$ to the band structure (see right inset of figure 10) and should scale with the electric field strength. However, as seen from the right panel of figure 10, there is no correlation between the two quantities. Hence we conclude that the simple Rashba model is completely inadequate and that the strength of the perpendicular electric field cannot be used to quantify the effect of spin–orbit interactions on band energies.

4.2. Monolayers from known layered bulk crystals

The C2DB has been extended with a number of monolayers that are likely exfoliable from experimentally known layered bulk compounds. Specifically, the Inorganic Crystal Structure Database (ICSD) [67] and Crystallography Open Database (COD) [68]

Figure 10. Left: Correlation between the electronegativity difference of *X* and *Y* in MXY Janus monolayers and the vacuum level shift across the layer. Right: Correlation between the Rashba energy and the vacuum level shift. Structures in the H-phase (e.g. MoSSe) are shown in black while structures in the T-phase (e.g. BiTeI) are shown in orange. The linear fit has the slope 1.35 eV/$\Delta\chi$ (Pauling scale). The insets show the definition of the vacuum level shift and the Rashba energy, respectively. Modified from [64].

have first been filtered for corrupted, duplicate and theoretical compounds, which reduce the initial set of 585.485 database entries to 167.767 unique materials. All of these have subsequently been assigned a 'dimensionality score' based on a purely geometrical descriptor. If the 2D score is larger than the sum of 0D, 1D and 3D scores we regard the material as being exfoliable and we extract the individual 2D components that comprise the material (see also section 2.2). We refer to the original work on the method for details [52] and note that similar approaches were applied in [11, 12] to identify potentially exfoliable monolayers from the ICSD and COD.

The search has been limited to bulk compounds containing less than six different elements and no rare earth elements. This reduces the set of relevant bulk materials to 2991. For all of these we extracted the 2D components containing less than 21 atoms in the unit cell, which were then relaxed and sorted for duplicates following the general C2DB workflow steps described in sections 2.1–2.3. At this point 781 materials remain. This set includes most known 2D materials and 207 of the 781 were already present in the C2DB prior to this addition. All the materials (including those that were already in C2DB) have been assigned an ICSD/COD identifier that refers to the parent bulk compound from which the 2D material was computationally exfoliated. We emphasise that we have not considered exfoliation energies in the analysis and a subset of these materials may thus be rather strongly bound and challenging to exfoliate even if the geometries indicate van der Waals bonded structures of the parent bulk compounds.

Figure 11 shows the distribution of energies above the convex hull for materials derived from

parent structures in ICSD or COD as well as for the entire C2DB, which includes materials obtained from combinatorial lattice decoration as well. As expected, the materials derived from experimental bulk materials are situated rather close to the convex hull whereas those obtained from lattice decoration extend to energies far above the convex hull. It is also observed that a larger fraction of the experimentally derived materials are dynamically stable. There are, however, well known examples of van der Waals bonded structures where the monolayer undergoes a significant lattice distortion, which will manifest itself as a dynamical instability in the present context. For example, bulk MoS_2 exists in van der Waals bonded structures composed of either 2 H-MoS_2 or 1 T-MoS_2 layers, but a monolayer of the 1 T phase undergoes a structural deformation involving a doubling of the unit cell [69] and is thus categorised as dynamically unstable by the C2DB workflow. The dynamically stable materials derived from parent bulk structures in the ICSD and COD may serve as a useful subset of the C2DB that are likely to be exfoliable from known compounds and thus facilitate experimental verification. As a first application the subset has been used to search for magnetic 2D materials, which resulted in a total of 85 ferromagnets and 61 anti-ferromagnets [70].

4.3. Outlook: multilayers
The C2DB is concerned with the properties of covalently bonded monolayers (see discussion of dimensionality filtering in section 2.2). However, multilayer structures composed of two or more identical monolayers are equally interesting and often have properties that deviate from those of the monolayer. In fact, the synthesis of layered vdW structures with a
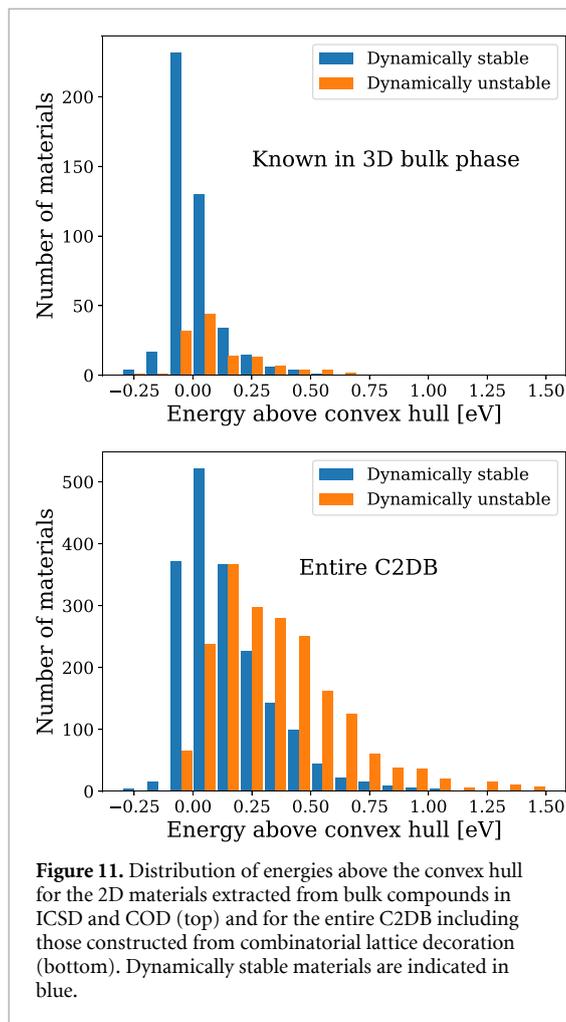
Figure 11. Distribution of energies above the convex hull for the 2D materials extracted from bulk compounds in ICSD and COD (top) and for the entire C2DB including those constructed from combinatorial lattice decoration (bottom). Dynamically stable materials are indicated in blue.

Figure 12. An illustration of the optimisation of the interlayer (IL) distance for MoS_2 in the AA stacking. The black crosses are the points sampled by the optimisation algorithm while the blue curve is a spline interpolation of the black crosses. The inset shows the MoS_2 AA stacking and the definition of the IL distance is indicated with a black double-sided arrow.

Table 3. Exfoliation energies for selected materials calculated with the PBE+D3 xc-functional as described in section 4.3 and compared with the DF2 and rVV10 results from [11]. The spacegroups are indicated in the column 'SG'. All numbers are in units of meV $Å^{-2}$.

Material	SG	PBE + D3	DF2	rVV10
MoS_2	P-6m2	28.9	21.6	28.8
$MoTe_2$	P-6m2	30.3	25.2	30.4
ZrNBr	Pmmn	18.5	10.5	18.5
C	P6/mmm	18.9	20.3	25.5
P	Pmna	21.9	38.4	30.7
BN	P-6m2	18.9	19.4	24.4
WTe_2	P-6m2	32.0	24.7	30.0
PbTe	P3m1	23.2	27.5	33.0

controllable number of layers represents an interesting avenue for atomic-scale materials design. Several examples of novel phenomena emerging in layered vdW structures have been demonstrated including direct-indirect band gap transitions in MoS_2 [71, 72], layer-parity selective Berry curvatures in few-layer WTe_2 [73], thickness-dependent magnetic order in CrI_3 [74, 75], and emergent ferroelectricity in bilayer hBN [76].

As a first step towards a systematic exploration of multilayer 2D structures, the C2DB has been used as basis for generating homobilayers in various stacking configurations and subsequently computing their properties following a modified version of the C2DB monolayer workflow. Specifically, the most stable monolayers (around 1000) are combined into bilayers by applying all possible transformations (unit cell preserving point group operations and translations) of one layer while keeping the other fixed. The candidate bilayers generated in this way are subject to a stability analysis, which evaluates the binding energy and optimal IL distance based on PBE-D3 [77] total energy calculations keeping the atoms of the monolayers fixed in their PBE relaxed geometry, see figures 12 and table 3.
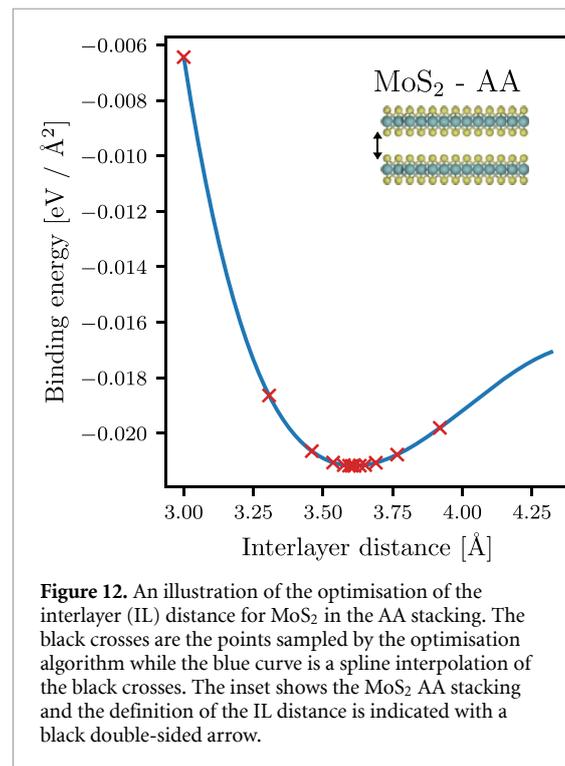
The calculated IL binding energies are generally in the range from a few to a hundred meV $Å^{-2}$ and IL distances range from 1.5 to 3.8 Å. A scatter plot of preliminary binding energies and IL distances is shown in figure 13. The analysis of homobilayers provides an estimate of the energy required to peel a monolayer off a bulk structure. In particular, the binding energy for the most stable bilayer configuration provides a measure of the *exfoliation energy* of the monolayer. This key quantity is now available for all monolayers in the C2DB, see section 5.1.

4.4. Outlook: point defects

The C2DB is concerned with the properties of 2D materials in their pristine crystalline form. However, as is well known the perfect crystal is an idealised model of real materials, which always contain defects in smaller or larger amounts depending on the intrinsic materials properties and growth conditions. Crystal defects often have a negative impact on

Figure 13. Scatter plot of the calculated interlayer distance and binding energies of (homo)bilayers of selected materials from C2DB. A few well known materials are highlighted: MoS_2, graphene (C_2), and hexagonal boron nitride (hBN). The bilayer binding energies provide an estimate of the monolayer exfoliation energies, see section 5.1.
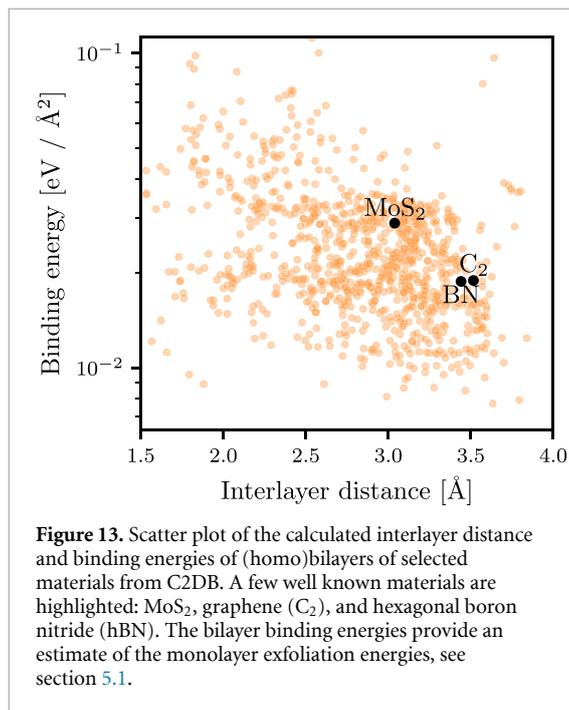
physical properties, e.g. they lead to scattering and life time-reduction of charge carriers in semiconductors. However, there are also important situations where defects play a positive enabling role, e.g. in doping of semiconductors, as colour centres for photon emission [78, 79] or as active sites in catalysis.

To reduce the gap between the pristine model material and real experimentally accessible samples, a systematic evaluation of the basic properties of the simplest native point defects in a selected subset of monolayers from the C2DB has been initiated. The monolayers are selected based on the stability of the pristine crystal. Moreover, only non-magnetic semi-conductors with a PBE band gap satisfying $E_{gap} > 1$ eV are currently considered as such materials are candidates for quantum technology applications like single-photon sources and spin qubits. Following these selection criteria around 300 monolayers are identified and their vacancies and intrinsic substitu-tional defects are considered, yielding a total of about 1500 defect systems.

Each defect system is subject to the same work-flow, which is briefly outlined below. To enable point defects to relax into their lowest energy configuration, the symmetry of the pristine host crystal is intention-ally broken by the chosen supercell, see figure 14 (a). In order to minimise defect–defect interaction, super-cells are furthermore chosen such that the minimum distance between periodic images of defects is larger than 15 Å. Unique point defects are created based on the analysis of equivalent Wyckoff positions for the host material. To illustrate some of the properties that will feature in the upcoming point defect database, we consider the specific example of monolayer CH_2Si.

First, the formation energy [80, 81] of a given defect is calculated from PBE total energies. Next,

Slater–Janak transition state theory is used to obtain the charge transition levels [82, 83]. By combining these results, one obtains the formation energy of the defect in all possible charge states as a function of the Fermi level. An example of such a diagram is shown in figure 14 (b) for the case of the V_C and C_{Si} defects in monolayer CH_2Si. For each defect and each charge state, the PBE single-particle energy level dia-gram is calculated to provide a qualitative overview of the electronic structure. A symmetry analysis [84] is performed for the defect structure and the individual defect states lying inside the band gap. The energy level diagram of the neutral V_{Si} defect in CH_2Si is shown in figure 14 (c), where the defect states are labelled according to the irreducible representations of the C_s point group.
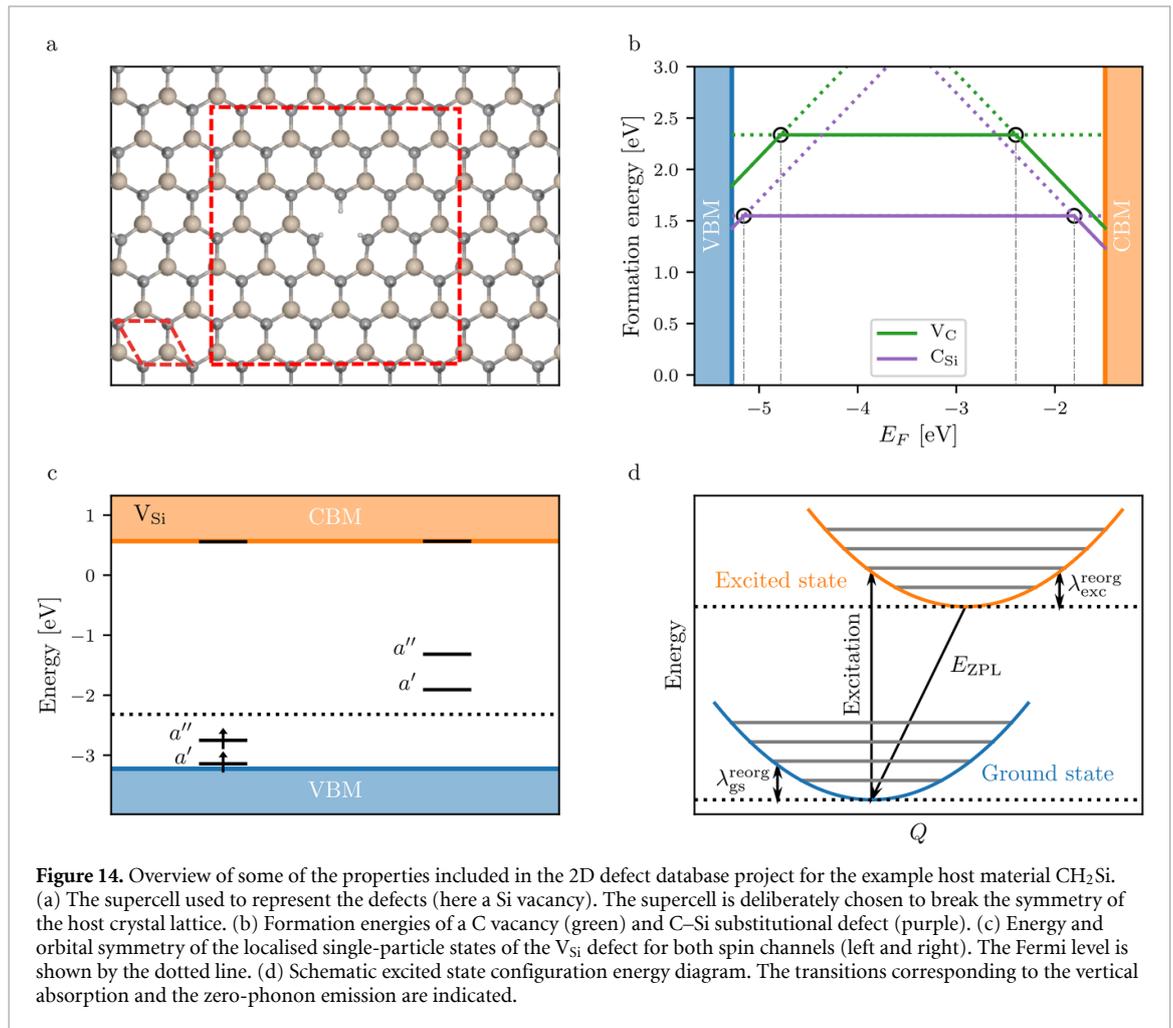
In general, excited electronic states can be mod-elled by solving the Kohn–Sham equations with non-Aufbau occupations. The excited-state solutions are saddle points of the Kohn–Sham energy functional, but common self-consistent field (SCF) approaches often struggle to find such solutions, especially when nearly degenerate states are involved. The calcula-tion of excited states corresponding to transitions between localised states inside the band gap is there-fore performed using an alternative method based on the direct optimisation (DO) of orbital rotations in combination with the maximum overlap method (MOM) [85]. This method ensures fast and robust convergence of the excited states, as compared to SCF. In figure 14 (d), the reorganisation energies for the ground and excited state, as well as the zero-phonon line (ZPL) energy are sketched. For the specific case of the Si vacancy in CH_2Si, the DO-MOM method yields $E_{ZPL} = 3.84$ eV, $\lambda_{gs}^{reorg} = 0.11$ eV and $\lambda_{exc}^{reorg} = 0.16$ eV. For systems with large electron-phonon coupling (i.e. Huang–Rhys factor > 1) a one-dimensional approx-imation for displacements along the main phonon mode is used to produce the configuration coordin-ate diagram (see figure 14 (d)). In addition to the ZPL energies and reorganisation energies, the Huang–Rhys factors, photoluminescence spectrum from the 1D phonon model, hyperfine coupling and zero field splitting are calculated.

5. New properties in the C2DB

This section reports on new properties that have become available in the C2DB since the first release. The employed computational methodology is described in some detail and results are compared to the literature where relevant. In addition, some interesting property correlations are considered along with general discussions of the general significance and potential application of the available data.

5.1. Exfoliation energy
The exfoliation energy of a monolayer is estimated as the binding energy of its bilayer in the most stable

Figure 14. Overview of some of the properties included in the 2D defect database project for the example host material CH_2Si. (a) The supercell used to represent the defects (here a Si vacancy). The supercell is deliberately chosen to break the symmetry of the host crystal lattice. (b) Formation energies of a C vacancy (green) and C–Si substitutional defect (purple). (c) Energy and orbital symmetry of the localised single-particle states of the V_{Si} defect for both spin channels (left and right). The Fermi level is shown by the dotted line. (d) Schematic excited state configuration energy diagram. The transitions corresponding to the vertical absorption and the zero-phonon emission are indicated.

stacking configuration (see also section 4.3). The binding energy is calculated using the PBE + D3 xc-functional [86] with the atoms of both monolayers fixed in the PBE relaxed geometry. Table 3 compares exfoliation energies obtained in this way to values from Mounet *et al* [11] for a representative set of monolayers.

5.2. Bader charges

For all monolayers we calculate the net charge on the individual atoms using the Bader partitioning scheme [87]. The analysis is based purely on the electron density, which we calculate from the PAW pseudo density plus compensation charges using the PBE xc-functional. Details of the method and its implementation can be found in Tang *et al* [88]. In section 5.4 we compare and discuss the relation between Bader charges and Born charges.

5.3. Spontaneous polarisation

The spontaneous polarisation (P_s) of a bulk material is defined as the charge displacement with respect to that of a reference centrosymmetric structure [89, 90]. Ferroelectric materials exhibit a finite value

of P_s that may be switched by an applied external field and have attracted a large interest for a wide range of applications [91–93].

The spontaneous polarisation in bulk materials can be regarded as electric dipole moment per unit volume, but in contrast to the case of finite systems this quantity is ill-defined for periodic crystals [89]. Nevertheless, one can define the formal polarisation density:

$$\mathbf{P} = \frac{1}{2\pi} \frac{e}{V} \sum_l \phi_l \mathbf{a}_l, \qquad (5)$$

where $\mathbf{a}_l$ (with $l \in \{1, 2, 3\}$) are the lattice vectors spanning the unit cell, V is the cell volume and e is the elementary charge. ϕ_l is the polarisation phase along the lattice vector defined by:

$$\phi_l = \sum_i Z_i \mathbf{b}_l \cdot \mathbf{u}_i - \phi_l^{\text{elec}}, \qquad (6)$$

where $\mathbf{b}_l$ is the reciprocal lattice vector satisfying $\mathbf{b}_l \cdot \mathbf{R}_l = 2\pi$ and $\mathbf{u}_i$ is the position of nucleus i with charge eZ_i. The electronic contribution to the polarisation phase is defined as:
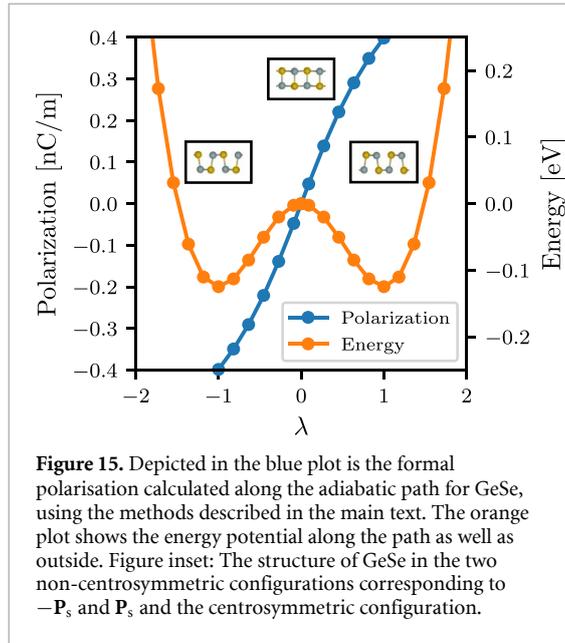
Figure 15. Depicted in the blue plot is the formal polarisation calculated along the adiabatic path for GeSe, using the methods described in the main text. The orange plot shows the energy potential along the path as well as outside. Figure inset: The structure of GeSe in the two non-centrosymmetric configurations corresponding to $-\mathbf{P}_s$ and $\mathbf{P}_s$ and the centrosymmetric configuration.

$$\phi_l^{\text{elec}} = \frac{1}{N_{k\perp\mathbf{b}_l}} \text{Im} \sum_{k\in\text{BZ}_{\perp\mathbf{b}_l}}$$
$$\times \ln \prod_{j=0}^{N_{k\|\mathbf{b}_l}-1} \det_{occ} \left[\langle u_{n\mathbf{k}+j\delta\mathbf{k}} | u_{m\mathbf{k}+(j+1)\delta\mathbf{k}} \rangle \right],$$
$$(7)$$

where $\text{BZ}_{\perp\mathbf{b}_l} = \{\mathbf{k}|\mathbf{k}\cdot\mathbf{b}_l = 0\}$ is a plane of $\mathbf{k}$-points orthogonal to $\mathbf{b}_l$, $\delta\mathbf{k}$ is the distance between neighbouring k-points in the $\mathbf{b}_l$ direction and $N_{k\|\mathbf{b}_l}$ ($N_{k\perp\mathbf{b}_l}$) is the number of $\mathbf{k}$-points along (perpendicular to) the $\mathbf{b}_l$ direction. These expression generalise straightforwardly to 2D.

The formal polarisation is only well-defined modulo $e\mathbf{R}_n/V$ where $\mathbf{R}_n$ is any lattice vector. However, changes in polarisation are well defined and the spontaneous polarisation may thus be obtained by:

$$\mathbf{P}_s = \int_0^1 \frac{d\mathbf{P}(\lambda)}{d\lambda} d\lambda, \qquad (8)$$

where λ is a dimensionless parameter that defines an adiabatic structural path connecting the polar phase ($\lambda = 1$) with a non-polar phase ($\lambda = 0$). The methodology has been implemented in GPAW and used to calculate the spontaneous polarisation of all stable materials in the C2DB with a PBE band gap above 0.01 eV and a polar space group symmetry. For each material, the centrosymmetric phase with smallest atomic displacement from the polar phase is constructed and relaxed under the constraint of inversion symmetry. The adiabatic path connecting the two phases is then used to calculate the spontaneous polarisation using equations (5)–(8). An example of a calculation for GeSe is shown in figure 15 where the polarisation along the path connecting two equivalent polar phases via the centrosymmetric phase is shown together with the total energy. The

spontaneous polarisation obtained from the path is 39.8 nC m^{-1} in good agreement with previous calculations [94].

5.4. Born charges
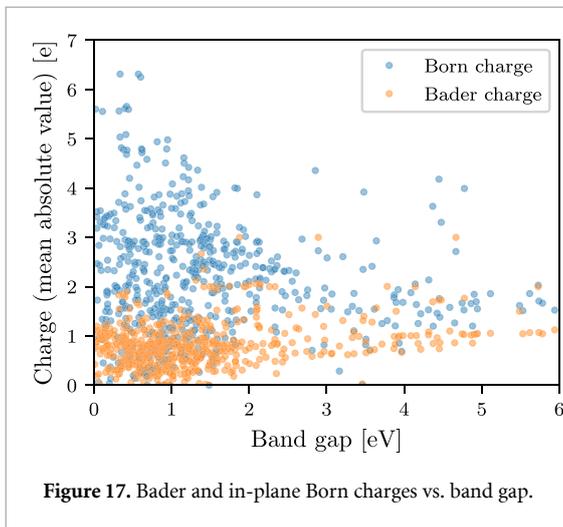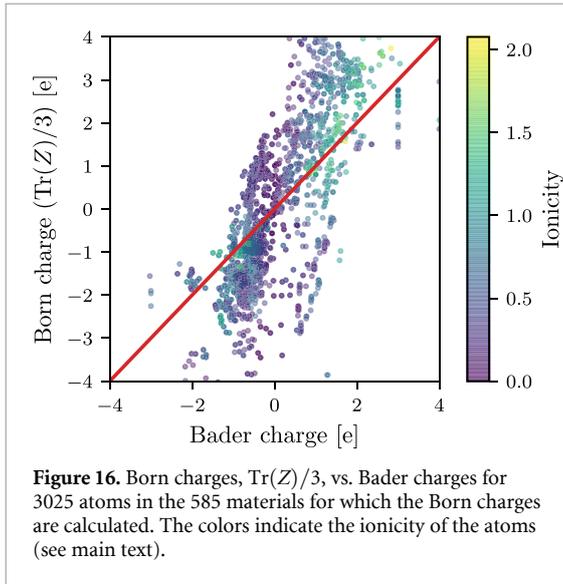The Born charge of an atom a at position $\mathbf{u}_a$ in a solid is defined as:

$$Z_{ij}^a = \frac{V}{e} \frac{\partial P_i}{\partial u_{aj}} \bigg|_{E=0}. \qquad (9)$$

It can be understood as an effective charge assigned to the atom to match the change in polarisation in direction i when its position is perturbed in direction j. Since the polarisation density and the atomic position are both vectors, the Born charge of an atom is a rank-2 tensor. The Born charge is calculated as a finite difference and relies on the Modern theory of polarisation [95] for the calculation of polarisation densities, see reference [96] for more details. The Born charge has been calculated for all stable materials in C2DB with a finite PBE band gap.

It is of interest to examine the relation between the Born charge and the Bader charge (see section 5.2). In materials with strong ionic bonds one would expect the charges to follow the atoms. On the other hand, in covalently bonded materials the hybridisation pattern and thus the charge distribution, depends on the atom positions in a complex way, and the idea of charges following the atom is expected to break down. In agreement with this idea, the (in-plane) Born charges in the strongly ionic hexagonal hBN ($\pm 2.71e$ for B and N, respectively) are in good agreement with the calculated Bader charges ($\pm 3.0e$). In contrast, (the in-plane) Born charges in MoS$_2$ ($-1.08e$ and $0.54e$ for Mo and S, respectively) deviate significantly from the Bader charges ($1.22e$ and $-0.61e$ for Mo and S, respectively). In fact, the values disagree even on the sign of the charges underlining the non-intuitive nature of the Born charges in covalently bonded materials.

Note that the out-of-plane Born charges never match the Bader charges, even for strongly ionic insulators, and are consistently smaller in value than the in-plane components. The smaller out-of-plane values are consistent with the generally smaller out-of-plane polarisability of 2D materials (for both electronic and phonon contributions) and agrees with the intuitive expectation that it is more difficult to polarise a 2D material in the out-of-plane direction as compared to the in-plane direction.

Figure 16 shows the average of the diagonal of the Born charge tensor, $\text{Tr}(Z^a)/3$, plotted against the Bader charges for all 585 materials in the C2DB for which the Born charges have been computed. The data points have been coloured according to the ionicity of the atom a defined as $I(a) = |\chi_a - \langle\chi\rangle|$, where χ_a and $\langle\chi\rangle$ are the Pauling electronegativity of atom a and the average electronegativity of all atoms in the

Figure 16. Born charges, $\mathrm{Tr}(Z)/3$, vs. Bader charges for 3025 atoms in the 585 materials for which the Born charges are calculated. The colors indicate the ionicity of the atoms (see main text).

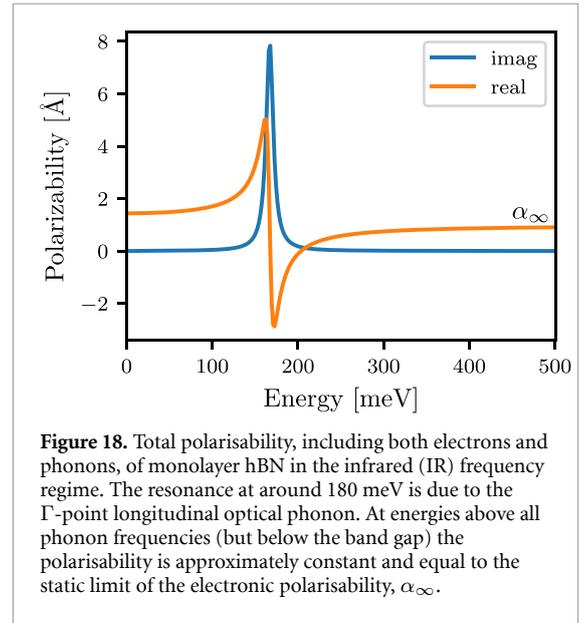**Figure 17.** Bader and in-plane Born charges vs. band gap.

unit cell, respectively. The ionicity is thus a measure of the tendency of an atom to donate/accept charge relative to the average tendency of atoms in the material. It is clear from figure 16 that there is a larger propensity for the Born and Bader charges to match in materials with higher ionicity.

Figure 17 plots the average (in-plane) Born charge and the Bader charge versus the band gap. It is clear that large band gap materials typically exhibit integer Bader charges, whereas there is no clear correlation between the Born charge and the band gap.

5.5. Infrared polarisability

The original C2DB provided the frequency dependent polarisability computed in the random phase approximation (RPA) with inclusion of electronic interband and intraband (for metals) transitions [6]. However, phonons carrying a dipole moment (so-called IR active phonons) also contribute to the polarisability at frequencies comparable to the frequency of optical phonons. This response is described by the IR polarisability:

Figure 18. Total polarisability, including both electrons and phonons, of monolayer hBN in the infrared (IR) frequency regime. The resonance at around 180 meV is due to the Γ-point longitudinal optical phonon. At energies above all phonon frequencies (but below the band gap) the polarisability is approximately constant and equal to the static limit of the electronic polarisability, α_∞.

$$\alpha^{\mathrm{IR}}(\omega) = \frac{e^2}{A} \mathbf{Z}^T \mathbf{M}^{-1/2} \left(\sum_i \frac{\mathbf{d}_i \mathbf{d}_i^T}{\omega_i^2 - \omega^2 - i\gamma\omega} \right) \mathbf{M}^{-1/2} \mathbf{Z},$$

(10)

where $\mathbf{Z}$ and $\mathbf{M}$ are matrix representations of the Born charges and atomic masses, ω_i^2 and d_i are eigenvectors and eigenvalues of the dynamical matrix, A is the in-plane cell area and γ is a broadening parameter representing the phonon lifetime and is set to 10 meV. The total polarisability is then the sum of the electronic polarisability and the IR polarisability.

The new C2DB includes the IR polarisability of all monolayers for which the Born charges have been calculated (stable materials with a finite band gap), see section (5.4). As an example, figure 18 shows the total polarisability of monolayer hexagonal hBN. For details on the calculation of the IR polarisability see reference [96].

5.6. Piezoelectric tensor

The piezoelectric effect is the accumulation of charges, or equivalently the formation of an electric polarisation, in a material in response to an applied mechanical stress or strain. It is an important material characteristic with numerous scientific and technological applications in sonar, microphones, accelerometers, ultrasonic transducers, energy conversion, etc [97, 98]. The change in polarisation originates from the movement of positive and negative charge centres as the material is deformed.

Piezoelectricity can be described by the (proper) piezoelectric tensor c_{ijk} with $i, j, k \in \{x, y, z\}$, given by [99]:

$$c_{ijk} = \frac{e}{2\pi V} \sum_l \frac{\partial \phi_{li}}{\partial \epsilon_{jk}},$$

(11)

which differs from equation (5) only by a derivative of the polarisation phase with respect to the strain tensor

Table 4. Comparison of computed piezoelectric tensor versus experimental values and previous calculations for hexagonal BN and a selected set of TMDCs (space group 187). All numbers are in units of nC/m. Experimental data for MoS_2 is obtained from [102].

Material	Exp.	Theory [101]	C2DB
BN	—	0.14	0.13
MoS_2	0.3	0.36	0.35
$MoSe_2$	—	0.39	0.38
$MoTe_2$	—	0.54	0.48
WS_2	—	0.25	0.24
WSe_2	—	0.27	0.26
WTe_2	—	0.34	0.34

ϵ_{jk}. Note that c_{ijk} does not depend on the chosen branch cut.

The piezoelectric tensor is a symmetric tensor with at most 18 independent components. Furthermore, the point group symmetry restricts the number of independent tensor elements and their relationships due to the well-known Neumann's principle [100]. For example, monolayer MoS_2 with point group D_{3h}, has only one non-vanishing independent element of c_{ijk}. Note that c_{ijk} vanishes identically for centrosymmetric materials. Using a finite-difference technique with a finite but small strain (1% in our case), equation (11) has been used to compute the proper piezoelectric tensor for all non-centrosymmetric materials in the C2DB with a finite band gap. Table 4 shows a comparison of the piezoelectric tensors in the C2DB with literature for a selected set of monolayer materials. Good agreement is obtained for all these materials.

5.7. Topological invariants

For all materials in the C2DB exhibiting a direct band gap below 1 eV, the k-space Berry phase spectrum of the occupied bands has been calculated from the PBE wave functions. Specifically, a particular k-point is written as $k_1 \mathbf{b}_1 + k_2 \mathbf{b}_2$ and the Berry phases $\gamma_n(k_2)$ of the occupied states on the path $k_1 = 0 \rightarrow k_1 = 1$ is calculated for each value of k_2. The connectivity of the Berry phase spectrum determines the topological properties of the 2D Bloch Hamiltonian [103, 104].

The calculated Berry phase spectra of the relevant materials are available for visual inspection on the C2DB webpage. Three different topological invariants have been extracted from these spectra and are reported in the C2DB: (1) The Chern number, C, takes an integer value and is well defined for any gapped 2D material. It determines the number of chiral edge states on any edge of the material. For any non-magnetic material the Chern number vanishes due to time-reversal symmetry. It is determined from the Berry phase spectrum as the number of crossings at any horizontal line in the spectrum. (2) The mirror Chern number, C_M, defined for gapped materials with a mirror plane in the atomic layer [105]. For such materials, all states may be chosen as mirror

eigenstates with eigenvalues $\pm i$ and the Chern numbers $C_\pm$ can be defined for each mirror sector separately. For a material with vanishing Chern number, the mirror Chern number is defined as $C_M = (C_+ - C_-)/2$ and takes an integer value corresponding to the number of edge states on any mirror symmetry preserving edge. It is obtained from the Berry phase spectrum as the number of chiral crossings in each of the mirror sectors. (3) The Z_2 invariant, ν, which can take the values 0 and 1, is defined for materials with time-reversal symmetry. Materials with $\nu = 1$ are referred to as quantum spin Hall insulators and exhibit helical edge states at any time-reversal conserving edge. It is determined from the Berry phase spectrum as the number of crossing points modulus 2 at any horizontal line in the interval $k_2 \in [0, 1/2]$.

Figure 19 shows four representative Berry phase spectra corresponding to the three cases of non-vanishing C, C_M and ν as well as a trivial insulator. The four materials are: $OsCl_3$ (space group 147)—a Chern insulator with $C = 1$, $OsTe_2$ (space group 14)—a mirror crystalline insulator with $C_M = 2$, SbI (spacegroup 1)—a quantum spin Hall insulator with $\nu = 1$ and BiITe (spacegroup 156)—a trivial insulator. Note that a gap in the Berry phase spectrum always implies a trivial insulator.

In [106] the C2DB was screened for materials with non-trivial topology. At that point it was found that the database contained 7 Chern insulators, 21 mirror crystalline topological insulators and 48 quantum spin Hall insulators. However, that does not completely exhaust the the topological properties of materials in the C2DB. In particular, there may be materials that can be topologically classified based on crystalline symmetries other than the mirror plane of the layer. In addition, second order topological effects may be present in certain materials, which imply that flakes will exhibit topologically protected corner states. Again, the Berry phase spectra may be used to unravel the second order topology by means of nested Wilson loops [107].

5.8. Exchange coupling constants

The general C2DB workflow described in sections 2.1–2.3 will identify the FM ground state of a material and apply it as starting point for subsequent property calculations, whenever it is more stable than the spin-paired ground state. In reality, however, the FM state is not guaranteed to comprise the magnetic ground state. In fact, AFM states often have lower energy than the FM one, but in general it is non-trivial to obtain the true magnetic ground state. We have chosen to focus on the FM state due to its simplicity and because its atomic structure and stability are often very similar to those of other magnetic states. Whether or not the FM state is the true magnetic ground state is indicated by the nearest neighbour exchange coupling constant as described below.
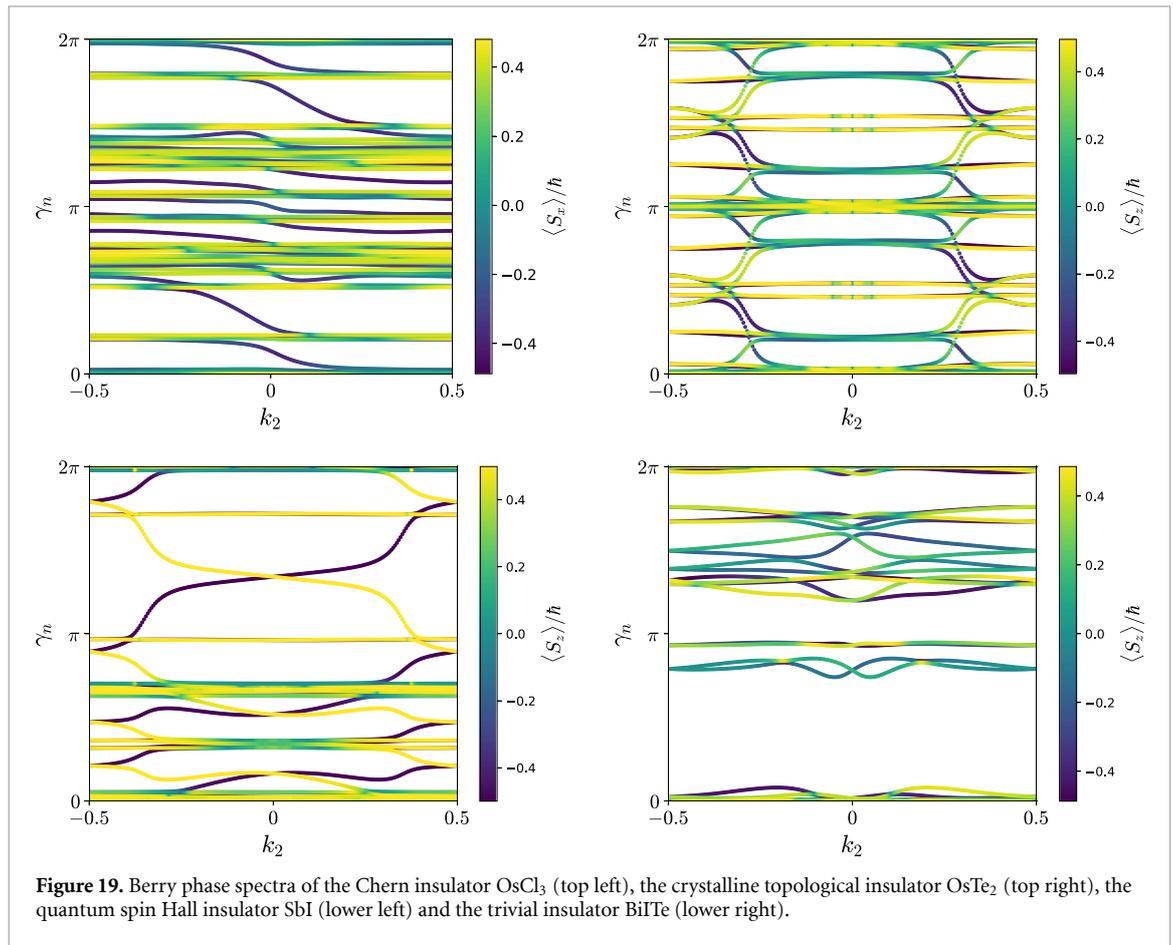
Figure 19. Berry phase spectra of the Chern insulator OsCl$_3$ (top left), the crystalline topological insulator OsTe$_2$ (top right), the quantum spin Hall insulator SbI (lower left) and the trivial insulator BiITe (lower right).

When investigating magnetic materials the thermodynamical properties (for example the critical temperatures for ordering) are of crucial interest. In two dimensions the Mermin–Wagner theorem [108] comprises an extreme example of the importance of thermal effects since it implies that magnetic order is only possible at $T = 0$ unless the spin-rotational symmetry is explicitly broken. The thermodynamic properties cannot be accessed directly by DFT. Consequently, magnetic models that capture the crucial features of magnetic interactions must be employed. For insulators, the Heisenberg model has proven highly successful in describing magnetic properties of solids in 3D as well as 2D [109]. It represents the magnetic degrees of freedom as a lattice of localised spins that interact through a set of exchange coupling constants. If the model is restricted to include only nearest neighbour exchange and assume magnetic isotropy in the plane, it reads:

$$H = -\frac{J}{2}\sum_{\langle ij \rangle} \mathbf{S}_i \cdot \mathbf{S}_j - \frac{\lambda}{2}\sum_{\langle ij \rangle} S_i^z S_j^z - A\sum_i \left(S_i^z\right)^2, \quad (12)$$

where J is the nearest neighbour exchange constant, λ is the nearest neighbour anisotropic exchange constant and A measures the strength of single-ion anisotropy. We also neglect off-diagonal exchange coupling constants that give rise to terms proportional to $S_i^x S_j^y$, $S_i^y S_j^z$ and $S_i^z S_j^x$. The out-of-plane direction has

been chosen as z and $\langle ij \rangle$ implies that for each site i we sum over all nearest neighbour sites j. The parameters J, λ and A may be obtained from an energy mapping analysis involving four DFT calculations with different spin configurations [70, 110, 111]. The thermodynamic properties of the resulting 'first principles Heisenberg model' may subsequently be analysed with classical Monte Carlo simulations or renormalised spin wave theory [36, 112].

The C2DB provides the values of J, λ, and A as well as the number of nearest neighbours N_{nn} and the maximum eigenvalue of S_z (S), which is obtained from the total magnetic moment per atom in the FM ground state (rounded to nearest half-integer for metals). These key parameters facilitate easy post-processing analysis of thermal effects on the magnetic structure. In [113] such an analysis was applied to estimate the critical temperature of all FM materials in the C2DB based on a model expression for T_C and the parameters from equation (12).

For metals, the Heisenberg parameters available in C2DB should be used with care because the Heisenberg model is not expected to provide an accurate description of magnetic interactions in this case. Nevertheless, even for metals the sign and magnitude of the parameters provide an important qualitative measure of the magnetic interactions that may be used to screen and select materials for more detailed investigations of magnetic properties.
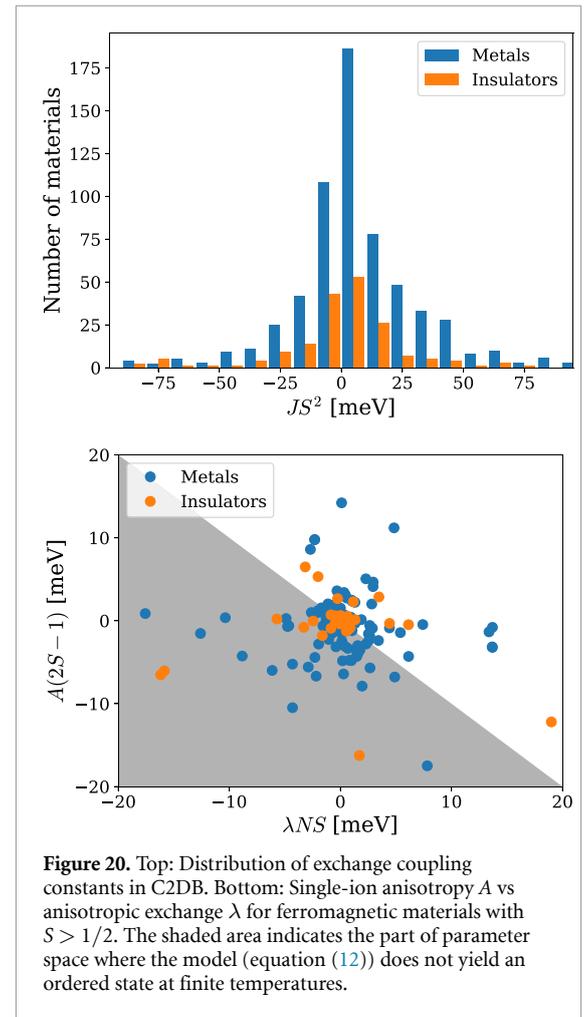
A negative value of J implies the existence of an AFM state with lower energy than the FM state used in C2DB. This parameter is thus crucial to consider when judging the stability and relevance of a material classified as magnetic in C2DB (see section 2.5). Figure 20 shows the distribution of exchange coupling constants (weighted by S^2) of the magnetic materials in the C2DB. The distribution is slightly skewed to the positive side indicating that FM order is more common than AFM order.

The origin of magnetic anisotropy may stem from either single-ion anisotropy or anisotropic exchange and it is in general difficult *a priori* to determine, which mechanism is most important. There is, however, a tendency in the literature to neglect anisotropic exchange terms in a Heisenberg model description of magnetism and focus solely on the single-ion anisotropy. In figure 20 we show a scatter plot of the anisotropy parameters A and λ for the FM materials ($J > 0$). The spread of the parameters indicate that the magnetic anisotropy is in general equally likely to originate from both mechanisms and neglecting anisotropic exchange is not advisable. For ferromagnets, the model (equation (12)) only exhibits magnetic order at finite temperatures if $A(2S-1) + \lambda N_{nn} > 0$ [113]. Neglecting anisotropic exchange thus excludes materials with $A < 0$ that satisfies $A(2S-1) + \lambda N_{nn} > 0$. This is in fact the case for 11 FM insulators and 31 FM metals in the C2DB.

5.9. Raman spectrum

Raman spectroscopy is an important technique used to probe the vibrational modes of a solid (or molecule) by means of inelastic scattering of light [114]. In fact, Raman spectroscopy is the dominant method for characterising 2D materials and can yield detailed information about chemical composition, crystal structure and layer thickness. There exist several different types of Raman spectroscopies that differ mainly by the number of photons and phonons involved in the scattering process [114]. The first-order Raman process, in which only a single phonon is involved, is the dominant scattering process in samples with low defect concentrations.
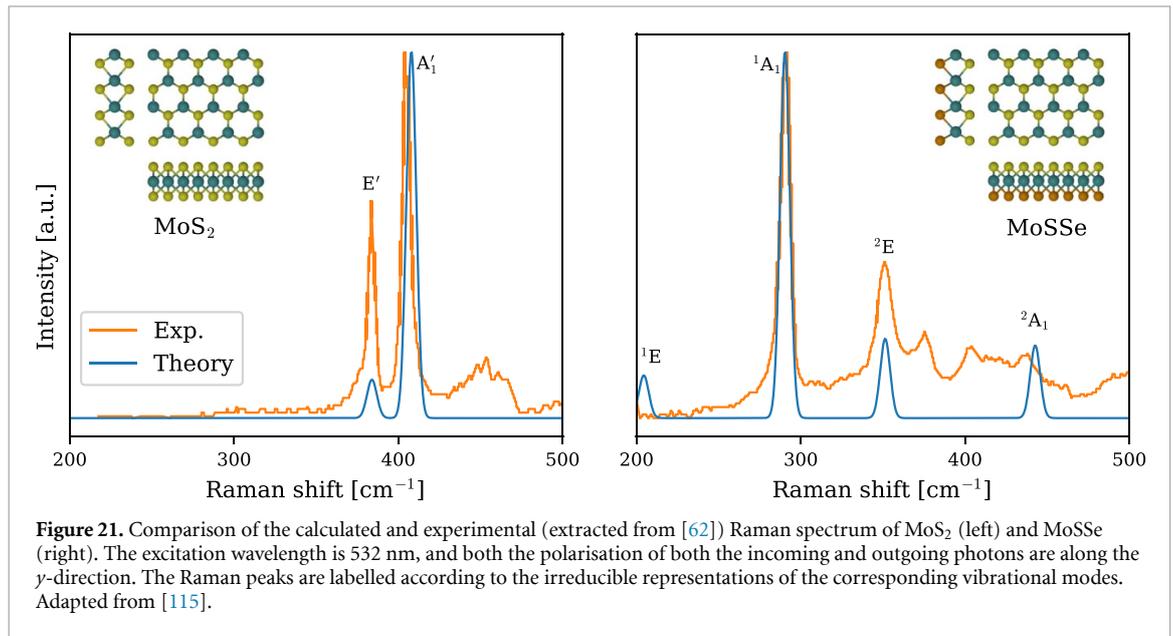
In a recent work, the first-order Raman spectra of 733 monolayer materials from the C2DB were calculated, and used as the basis for an automatic procedure for identifying a 2D material entirely from its experimental Raman spectrum [115]. The Raman spectrum is calculated using third-order perturbation theory to obtain the rate of scattering processes involving creation/annihilation of one phonon and two photons, see reference [115] for details. The light field is written as $\mathcal{F}(t) = \mathcal{F}_{in}\mathbf{u}_{in}\exp(-i\omega_{in}t) + \mathcal{F}_{out}\mathbf{u}_{out}\exp(-i\omega_{out}t)$+c.c. where $\mathcal{F}_{in/out}$ and $\omega_{in/out}$ denote the amplitudes and frequencies of the input/output electromagnetic fields, respectively. In addition, $\mathbf{u}_{in/out} = \sum_i u^i_{in/out}\mathbf{e}_i$ are the corresponding polarisation vectors, where $\mathbf{e}_i$ denotes the unit

Figure 20. Top: Distribution of exchange coupling constants in C2DB. Bottom: Single-ion anisotropy A vs anisotropic exchange λ for ferromagnetic materials with $S > 1/2$. The shaded area indicates the part of parameter space where the model (equation (12)) does not yield an ordered state at finite temperatures.

vector along the i-direction with $i \in \{x, y, z\}$. Using this light field, the final expression for the Stokes Raman intensity involving scattering events by only one phonon reads [115]:

$$I(\omega) = I_0 \sum_\nu \frac{n_\nu + 1}{\omega_\nu} \left| \sum_{ij} u^i_{in} R^\nu_{ij} u^j_{out} \right|^2 \delta(\omega - \omega_\nu). \quad (13)$$

Here, I_0 is an unimportant constant (since Raman spectra are always reported normalised), and n_ν is obtained from the Bose–Einstein distribution, i.e. $n_\nu \equiv (\exp[\hbar\omega_\nu/k_B T] - 1)^{-1}$ at temperature T for a Raman mode with energy $\hbar\omega_\nu$. Note that only phonons at the Brillouin zone center (with zero momentum) contribute to the one-phonon Raman processes due to momentum conservation. In equation (13), R^ν_{ij} is the Raman tensor for phonon mode ν, which involves electron–phonon and dipole matrix elements as well as the electronic transition energies and the incident excitation frequency. Equation (13) has been used to compute the Raman spectra of the 733 most stable, non-magnetic monolayers in C2DB for a range of excitation frequencies and polarisation configurations. Note that the Raman shift $\hbar\omega$ is typically expressed in cm^{-1} with

Figure 21. Comparison of the calculated and experimental (extracted from [62]) Raman spectrum of MoS$_2$ (left) and MoSSe (right). The excitation wavelength is 532 nm, and both the polarisation of both the incoming and outgoing photons are along the y-direction. The Raman peaks are labelled according to the irreducible representations of the corresponding vibrational modes. Adapted from [115].

1 meV equivalent to 8.0655 cm^{-1}. In addition, for generating the Raman spectra, we have used a Gaussian $[G(\omega) = (\sigma\sqrt{2\pi})^{-1}\exp(-\omega^2/2\sigma^2)]$ with a variance $\sigma = 3$ cm^{-1} to replace the Dirac delta function, which accounts for the inhomogeneous broadening of phonon modes.

As an example, figure 21 shows the calculated Raman spectrum of monolayer MoS$_2$ and the Janus monolayer MoSSe (see section 4.1). Experimental Raman spectra extracted from reference [62] are shown for comparison. For both materials, good agreement between theory and experiment is observed for the peak positions and relative amplitudes of the main peaks. The small deviations can presumably be attributed to substrate interactions and defects in the experimental samples as well as the neglect of excitonic effects in the calculations. The qualitative differences between the Raman spectra can be explained by the different point groups of the materials (C_{3v} and D_{3h}, respectively), see reference [115]. In particular, the lower symmetry of MoSSe results in a lower degeneracy of its vibrational modes leading to more peaks in the Raman spectrum.

Very recently, the Raman spectra computed from third order perturbation theory as described above, were supplemented by spectra obtained from the more conventional Kramers–Heisenberg–Dirac (KHD) approach. Within the KHD method, the Raman tensor is obtained as the derivative of the static electric polarisability (or equivalently, the susceptibility) along the vibrational normal modes [116, 117]:

$$R_{ij}^{\nu} = \sum_{\alpha l} \frac{\partial \chi_{ij}^{(1)}}{\partial r_{\alpha l}} \frac{v_{\alpha l}^{\nu}}{\sqrt{M_{\alpha}}}. \qquad (14)$$

Here, $\chi_{ij}^{(1)}$ is the (first-order) susceptibility tensor, r_{α} and M_{α} are the position and atomic mass of atom

α, respectively, and $v_{\alpha l}^{\nu}$ is the eigenmode of phonon ν. The two approaches, i.e. the KHD and third-order perturbation approach, can be shown to be equivalent [114], at least when local field effects can be ignored as is typically the case for 2D materials [35]. We have also confirmed this equivalence from our calculations. Furthermore, the computational cost of both methods is also similar [115]. However, the KHD approach typically converge faster with respect to both the number of bands and k-grid compared to the third-order perturbation method. This stems from the general fact that higher-order perturbation calculations converge slower with respect to k-grid and they require additional summations over a complete basis set (virtual states) and hence a larger number of bands [118]. Currently, Raman spectra from both approaches can be found at the C2DB website.

5.10. Second harmonics generation

Nonlinear optical (NLO) phenomena such as harmonic generation, Kerr, and Pockels effects are of great technological importance for lasers, frequency converters, modulators, etc. In addition, NLO spectroscopy has been extensively employed to obtain insight into materials properties [119] that are not accessible by e.g. linear optical spectroscopy. Among numerous nonlinear processes, second-harmonic generation (SHG) has been widely used for generating new frequencies in lasers as well as identifying crystal orientations and symmetries.

Recently, the SHG spectrum was calculated for 375 non-magnetic, non-centrosymmetric semiconducting monolayers of the C2DB and multiple 2D materials with giant optical nonlinearities were identified [120]. In the SHG process, two incident photons at frequency ω generate an emitted photon at frequency of 2ω. Assume that a mono-harmonic electric
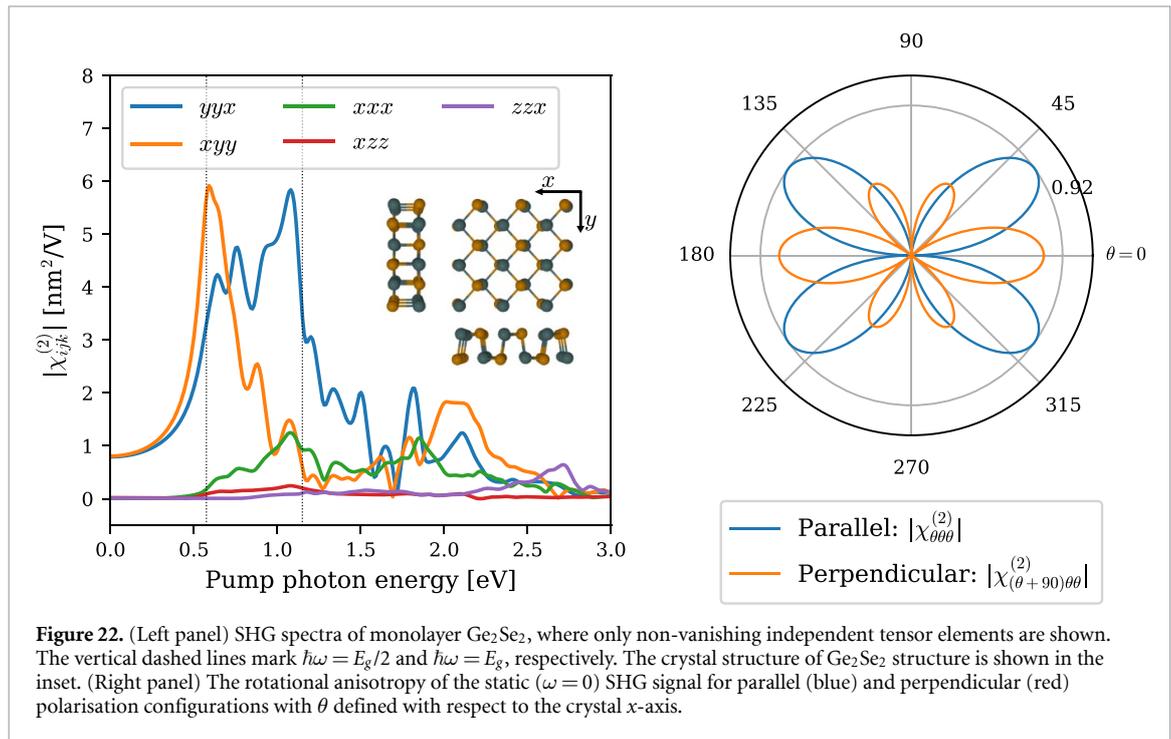
Figure 22. (Left panel) SHG spectra of monolayer Ge_2Se_2, where only non-vanishing independent tensor elements are shown. The vertical dashed lines mark $\hbar\omega = E_g/2$ and $\hbar\omega = E_g$, respectively. The crystal structure of Ge_2Se_2 structure is shown in the inset. (Right panel) The rotational anisotropy of the static ($\omega = 0$) SHG signal for parallel (blue) and perpendicular (red) polarisation configurations with θ defined with respect to the crystal x-axis.

field written $\mathcal{F}(t) = \sum_i \mathcal{F}_i \mathbf{e}_i e^{-i\omega t}$+c.c. is incident on the material, where $\mathbf{e}_i$ denotes the unit vector along direction $i \in \{x, y, z\}$. The electric field induces a SHG polarisation density $\mathbf{P}^{(2)}$, which can be obtained from the quadratic susceptibility tensor $\chi_{ijk}^{(2)}$,

$$P_i^{(2)}(t) = \epsilon_0 \sum_{jk} \chi_{ijk}^{(2)}(\omega, \omega)\mathcal{F}_i \mathcal{F}_j e^{-2i\omega t} + \text{c.c.}, \quad (15)$$
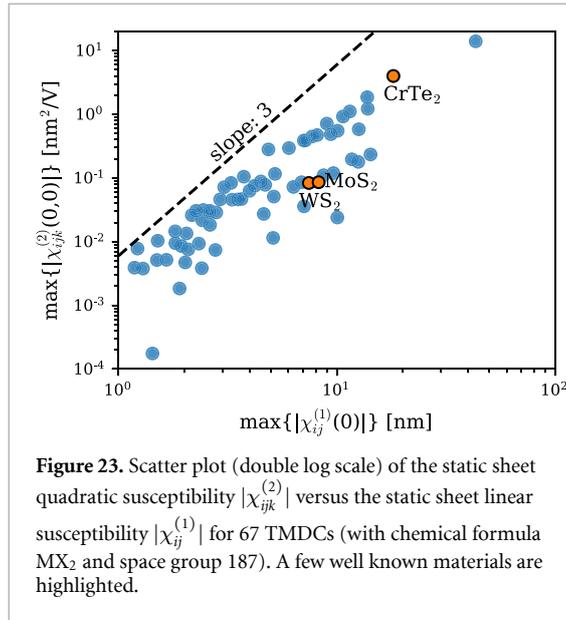
where ε_0 denotes the vacuum permittivity. $\chi_{ijk}^{(2)}$ is a symmetric (due to intrinsic permutation symmetry i.e. $\chi_{ijk}^{(2)} = \chi_{ijk}^{(2)}$) rank-3 tensor with at most 18 independent elements. Furthermore, similar to the piezoelectric tensor, the point group symmetry reduces the number of independent tensor elements.

In the C2DB, the quadratic susceptibility is calculated using density matrices and perturbation theory [118, 121] with the involved transition dipole matrix elements and band energies obtained from DFT. The use of DFT single-particle orbitals implies that excitonic effects are not accounted for. The number of empty bands included in the sum over bands was set to three times the number of occupied bands. The width of the Fermi–Dirac occupation factor was set to $k_BT = 50$ meV, and a line-shape broadening of $\eta = 50$ meV was used in all spectra. Furthermore, time-reversal symmetry was imposed in order to reduce the **k**-integrals to half the BZ. For various 2D crystal classes, it was verified by explicit calculation that the quadratic tensor elements fulfil the expected symmetries, e.g. that they all vanish identically for centrosymmetric crystals.

As an example, the calculated SHG spectra for monolayer Ge_2Se_2 is shown in figure 22 (left panel).

Monolayer Ge_2Se_2 has five independent tensor elements, $\chi_{xxx}^{(2)}$, $\chi_{xyy}^{(2)}$, $\chi_{xzz}^{(2)}$, $\chi_{yyx}^{(2)} = \chi_{yxy}^{(2)}$, and $\chi_{zzx}^{(2)} = \chi_{zxz}^{(2)}$, since it is a group-IV dichalcogenide with an orthorhombic crystal structure (space group 31 and point group C_{2v}). Note that, similar to the linear susceptibility, the bulk quadratic susceptibility (with SI units of $m\,V^{-1}$) is ill-defined for 2D materials (since the volume is ambiguous) [120]. Instead, the unambiguous *sheet* quadratic susceptibility (with SI units of $m^2\,V^{-1}$) is evaluated. In addition to the frequency-dependent SHG spectrum, the angular dependence of the static ($\omega = 0$) SHG intensity at normal incidence for parallel and perpendicular polarisations (relative to the incident electric field) is calculated, see figure 22 (right panel). Such angular resolved SHG spectroscopy has been widely used for determining the crystal orientation of 2D materials. The calculated SHG spectra for all non-vanishing inequivalent polarisation configurations and their angular dependence, are available in the C2DB.

Since C2DB has already gathered various material properties of numerous 2D materials, it provides a unique opportunity to investigate interrelations between different material properties. For example, the strong dependence of the quadratic optical response on the electronic band gap was demonstrated on basis of the C2DB data [120]. As another example of a useful correlation, the static quadratic susceptibility is plotted versus the static linear susceptibility for 67 TMDCs (with formula MX_2, space group 187) in figure 23. Note that for materials with several independent tensor elements, only the largest is shown. There is a very clear correlation between the two quantities. This is not unexpected as both

90

Figure 23. Scatter plot (double log scale) of the static sheet quadratic susceptibility $|\chi_{ijk}^{(2)}|$ versus the static sheet linear susceptibility $|\chi_{ij}^{(1)}|$ for 67 TMDCs (with chemical formula MX_2 and space group 187). A few well known materials are highlighted.

the linear and quadratic optical responses are functions of the transition dipole moments and transition energies. More interestingly, the strength of the quadratic response seems to a very good approximation to be given by a universal constant times the linear susceptibility to the power of three (ignoring polarisation indices), i.e.

$$\chi^{(2)}(0,0) \approx A\chi^{(1)}(0)^3, \qquad (16)$$

where A is only weakly material dependent. Note that this scaling law is also known in classical optics as semi-empirical Miller's rule for non-resonant quadratic responses [122], which states that the second order electric susceptibility is proportional to the product of the first-order susceptibilities at the three frequencies involved.

6. Machine learning properties

In recent years, material scientists have shown great interest in exploiting the use of machine learning (ML) techniques for predicting materials properties and guiding the search for new materials. ML is the scientific study of algorithms and statistical models that computer systems can use to perform a specific task without using explicit instructions but instead relying on patterns and inference. Within the domain of materials science, one of the most frequent problems is the mapping from atomic configuration to material property, which can be used e.g. to screen large material spaces in search of optimal candidates for specific applications [123, 124].

In the ML literature, the mathematical representation of the input observations is often referred to as a fingerprint. Any fingerprint must satisfy a number of general requirements [125]. In particular, a fingerprint must be:
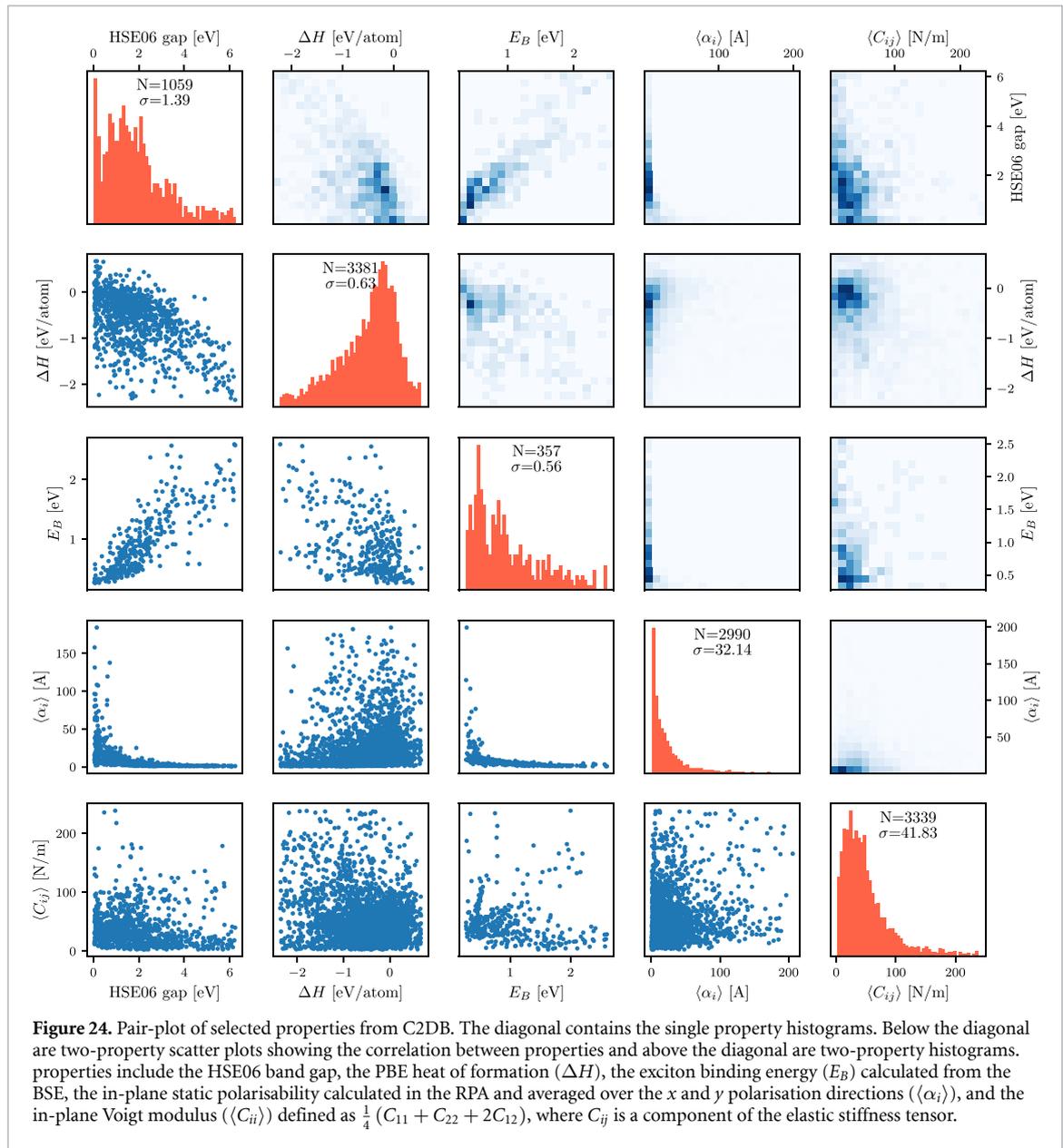
(a) *Complete:* The fingerprint should incorporate all the relevant input for the underlying problem, i.e. materials with different properties should have different fingerprints.

(b) *Compact:* The fingerprint should contain no or a minimal number of features redundant to the underlying problem. This includes being invariant to rotations, translations and other transformations that leave the properties of the system invariant.

(c) *Descriptive:* Materials with similar target values should have similar fingerprints.

(d) *Simple:* The fingerprint should be efficient to evaluate. In the present context, this means that calculating the fingerprint should be significantly faster than calculating the target property.

Several types of atomic-level materials fingerprints have been proposed in the literature, including general purpose fingerprints based on atomistic properties [126, 127] possibly encoding information about the atomic structure, i.e. atomic positions [125, 128, 129], and specialised fingerprints tailored for specific applications (materials/properties) [130, 131].

The aim of this section is to demonstrate how the C2DB may be utilised for ML-based prediction of general materials properties. Moreover, the study serves to illustrate the important role of the fingerprint for such problems. The 2D materials are represented using three different fingerprints: two popular structural fingerprints and a more advanced fingerprint that encodes information about the electronic structure via the PDOS. The target properties include the HSE06 band gap, the PBE heat of formation (ΔH), the exciton binding energy (E_B) obtained from the many-body BSE, the in-plane static polarisability calculated in the RPA averaged over the x and y polarisation directions ($\langle\alpha_i\rangle$), and the in-plane Voigt modulus ($\langle C_{ii}\rangle$) defined as $\frac{1}{4}(C_{11} + C_{22} + 2C_{12})$, where C_{ij} is a component of the elastic stiffness tensor in Mandel notation.

To introduce the data, figure 24 shows pair-plots of the dual-property relations of these properties. The plots in the diagonal show the single-property histograms, whereas the off-diagonals show dual-property scatter plots below the diagonal and histograms above the diagonal. Clearly, there are only weak correlations between most of the properties, with the largest degree of correlation observed between the HSE06 gap and exciton binding energy. The lack of strong correlations motivates the use of ML for predicting the properties.

The prediction models are build using the Ewald sum matrix and many-body tensor representation (MBTR) as structural fingerprints. The Ewald fingerprint is a version of the simple Coulomb matrix fingerprint [128] modified to periodic systems [125]. The MBTR encodes first, second and third order

Figure 24. Pair-plot of selected properties from C2DB. The diagonal contains the single property histograms. Below the diagonal are two-property scatter plots showing the correlation between properties and above the diagonal are two-property histograms. properties include the HSE06 band gap, the PBE heat of formation (ΔH), the exciton binding energy (E_B) calculated from the BSE, the in-plane static polarisability calculated in the RPA and averaged over the x and y polarisation directions ($\langle \alpha_i \rangle$), and the in-plane Voigt modulus ($\langle C_{ii} \rangle$) defined as $\frac{1}{4}(C_{11} + C_{22} + 2C_{12})$, where C_{ij} is a component of the elastic stiffness tensor.

terms like atomic numbers, distances and angles between atoms in the system [129]. As an alternative to the structural fingerprints, a representation based on the PBE PDOS is also tested. This fingerprint[6] encodes the coupling between the PDOS at different atomic orbitals in both energy and real space. It is defined as:

$$\rho_{\nu\nu'}(E,R) = \sum_{a \in \text{cell}} \sum_{a'} \rho_{a\nu}(E)\rho_{a'\nu'}(E)G$$
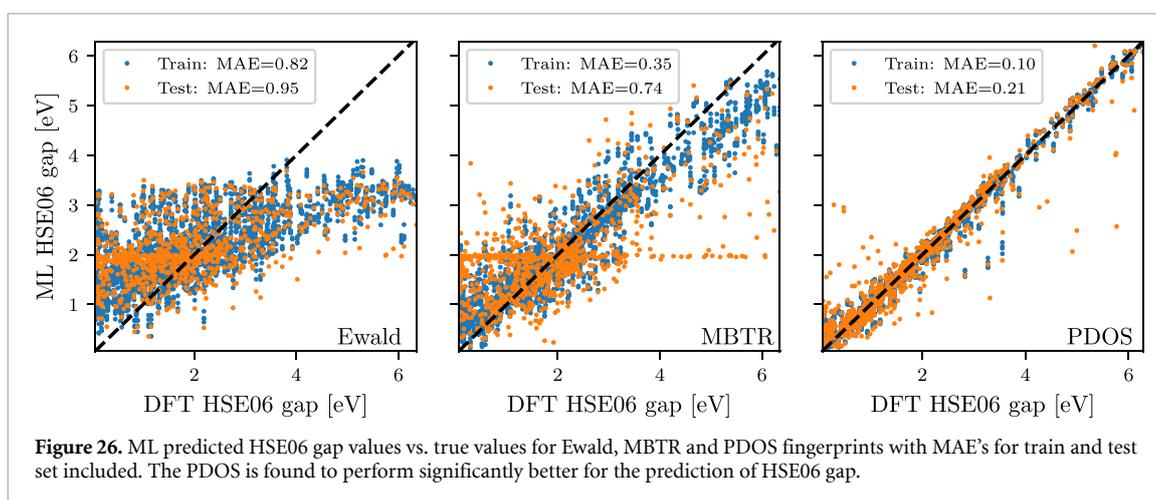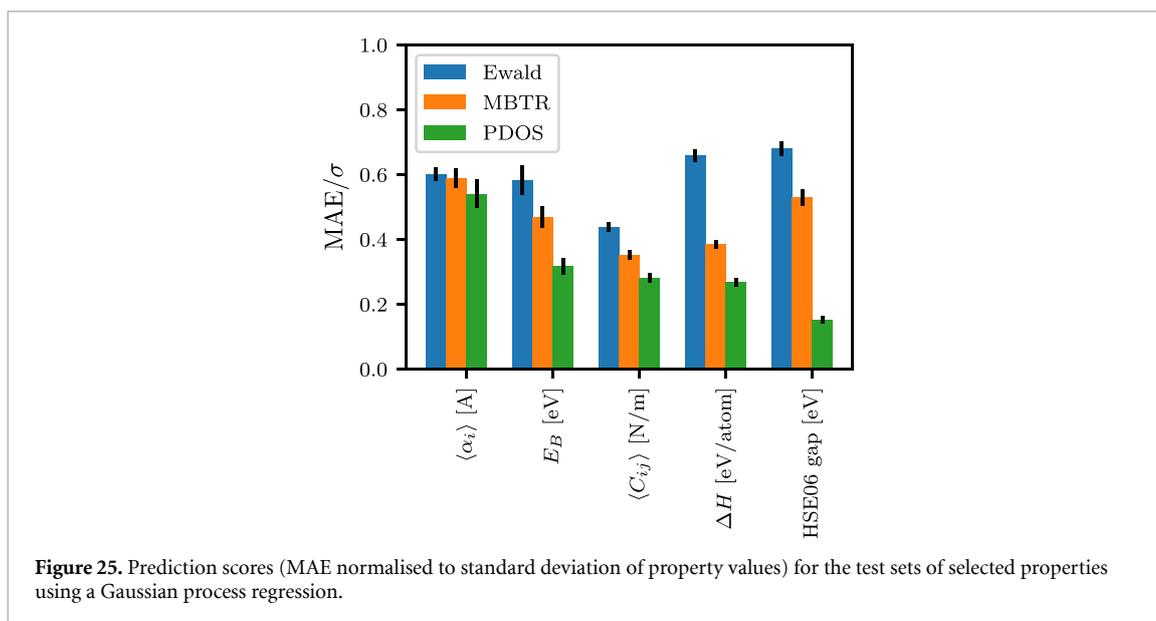$$\times (R - |R_a - R_{a'}|), \quad (17)$$

where G is a Gaussian smearing function, a denotes the atoms, ν denotes atomic orbitals, and the PDOS is given by:

$$\rho_{a\nu}(E) = \sum_n |\langle \psi_n | a\nu \rangle|^2 G(E - \epsilon_n), \quad (18)$$

where n runs over all eigenstates of the system. Since this fingerprint requires a DFT-PBE calculation to be performed, additional features derivable from the DFT calculation can be added to the fingerprint. In this study, the PDOS fingerprint is amended by the PBE band gap. The latter can in principle be extracted from the PDOS, but its explicit inclusion has been found to improve the performance of the model.

A Gaussian process regression using a simple Gaussian kernel with a noise component is used as learning algorithm. The models are trained using 5-fold cross validation on a training set consisting of 80% of the materials with the remaining 20% held aside as test data. Prior to training the model, the input space is reduced to 50 features using principal component analysis (PCA). This step is necessary to reduce the huge number of features in the MBTR fingerprint to a manageable size. Although this is not required for the Ewald and PDOS fingerprints,

92

---

6 Details will be published elsewhere.

Figure 25. Prediction scores (MAE normalised to standard deviation of property values) for the test sets of selected properties using a Gaussian process regression.

Figure 26. ML predicted HSE06 gap values vs. true values for Ewald, MBTR and PDOS fingerprints with MAE's for train and test set included. The PDOS is found to perform significantly better for the prediction of HSE06 gap.

we perform the same feature reduction in all cases. The optimal number of features depends on the choice of fingerprint, target property and learning algorithm, but for consistency 50 PCA components are used for all fingerprints and properties in this study.

Figure 25 shows the prediction scores obtained for the five properties using the three different fingerprints. The employed prediction score is the mean absolute error of the test set normalised by the standard deviation of the property values (standard deviations are annotated in the diagonal plots in figure 24). In general, the PDOS fingerprint outperforms the structural fingerprints. The difference between prediction scores is smallest for the static polarisability $\langle \alpha_i \rangle$ and largest for the HSE06 gap. It should be stressed that although the evaluation of the PBE-PDOS fingerprint is significantly more time consuming than the evaluation of the structural fingerprints, it is still much faster than the evaluation of all the target properties. Moreover, structural fingerprints require the atomic structure, which in turns

requires a DFT structure optimisation (unless the structure is available by other means).

The HSE06 band gap shows the largest sensitivity to the employed fingerprint. To elaborate on the HSE06 results, figure 26 shows the band gap predicted using each of the three different fingerprints plotted against the true band gap. The mean absolute errors on the test set is 0.95 and 0.74 eV for Ewald and MBTR fingerprints, respectively, while the PDOS significantly outperforms the other fingerprints with a test MAE of only 0.21 eV. This improvement in prediction accuracy is partly due to the presence of the PBE gap in the PDOS fingerprint. However, our analysis shows that the pure PDOS fingerprint without the PBE gap still outperforms the structural fingerprints. Using only the PBE gap as feature results in a test MAE of 0.28 eV.

The current results show that the precision of ML-based predictions are highly dependent on the type of target property and the chosen material representation. For some properties, the mapping between atomic structure and property is easier to learn while

93

others might require more/deeper information, e.g. in terms of electronic structure fingerprints. Our results clearly demonstrate the potential of encoding electronic structure information into the material fingerprint, and we anticipate more work on this relevant and exciting topic in the future.

7. Summary and outlook

We have documented a number of extensions and improvements of the C2DB made in the period 2018–2020. The new developments include: (1) A refined and more stringent workflow for filtering prospective 2D materials and classifying them according to their crystal structure, magnetic state and stability. (2) Improvements of the methodology used to compute certain challenging properties such as the full stiffness tensor, effective masses, G_0W_0 band structures, and optical absorption spectra. (3) New materials including 216 MXY Janus monolayers and 574 monolayers exfoliated from experimentally known bulk crystals. In addition, ongoing efforts to systematically obtain and characterise bilayers in all possible stacking configurations as well as point defects in the semiconducting monolayers, have been described. (4) New properties including exfoliation energies, spontaneous polarisations, Bader charges, piezoelectric tensors, IR polarisabilities, topological invariants, magnetic exchange couplings, Raman spectra, and SHG spectra. It should be stressed that the C2DB will continue to grow as new structures and properties are being added, and thus the present paper should not be seen as a final report on the C2DB but rather a snapshot of its current state.

In addition to the above mentioned improvements relating to data quantity and quality, the C2DB has been endowed with a comprehensive documentation layer. In particular, all data presented on the C2DB website are now accompanied by an information field that explains the meaning and representation (if applicable) of the data and details how it was calculated thus making the data easier to understand, reproduce, and deploy.

The C2DB has been produced using the ASR in combination with the GPAW electronic structure code and the MyQueue task and workflow scheduling system. The ASR is a newly developed Python-based framework designed for high-throughput materials computations. The highly flexible and modular nature of the ASR and its strong coupling to the well established community-driven ASE project, makes it a versatile framework for both high- and low-throughput materials simulation projects. The ASR and the C2DB-ASR workflow are distributed as open source code. A detailed documentation of the ASR will be published elsewhere.

While the C2DB itself is solely concerned with the properties of perfect monolayer crystals, ongoing efforts focus on the systematic characterisation of homo-bilayer structures as well as point defects in monolayers. The data resulting from these and other similar projects will be published as separate, independent databases, but will be directly interlinked with the C2DB making it possible to switch between them in a completely seamless fashion. These developments will significantly broaden the scope and usability of the C2DB+ (+ stands for associated databases) that will help theoreticians and experimentalists to navigate one of the most vibrant and rapidly expanding research fields at the crossroads of condensed matter physics, photonics, nanotechnology, and chemistry.

Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: https://doi.org/10.11583/DTU.14616660.

ORCID iDs

Morten Niklas Gjerding ⬤ https://orcid.org/0000-0002-5256-660X
Alireza Taghizadeh ⬤ https://orcid.org/0000-0003-0876-9538
Asbjørn Rasmussen ⬤ https://orcid.org/0000-0001-7110-9255
Sajid Ali ⬤ https://orcid.org/0000-0001-7865-2664
Fabian Bertoldo ⬤ https://orcid.org/0000-0002-1219-8689
Thorsten Deilmann ⬤ https://orcid.org/0000-0003-4165-2446
Nikolaj Rørbæk Knøsgaard ⬤ https://orcid.org/0000-0003-3709-5464
Mads Kruse ⬤ https://orcid.org/0000-0002-0599-5110
Ask Hjorth Larsen ⬤ https://orcid.org/0000-0001-5267-6852
Simone Manti ⬤ https://orcid.org/0000-0003-3770-0863
Thomas Garm Pedersen ⬤ https://orcid.org/0000-0002-9466-6190
Urko Petralanda ⬤ https://orcid.org/0000-0003-0226-0028

94

Thorbjørn Skovhus ⓘ https://orcid.org/0000-0001-5215-6419

Mark Kamper Svendsen ⓘ https://orcid.org/0000-0001-9718-849X

Jens Jørgen Mortensen ⓘ https://orcid.org/0000-0001-5090-6706

Thomas Olsen ⓘ https://orcid.org/0000-0001-6256-9284

Kristian Sommer Thygesen ⓘ https://orcid.org/0000-0001-5197-214X

References

[1] Kohn W and Sham L J 1965 *Phys. Rev.* **140** A1133
[2] Schwierz F 2010 *Nat. Nanotechnol.* **5** 487
[3] Novoselov K, Mishchenko A, Carvalho A and Castro Neto A 2016 *Science* **353** 6298
[4] Ferrari A C *et al* 2015 *Nanoscale* **7** 4598–810
[5] Bhimanapati G R *et al* 2015 *ACS Nano* **9** 11509–39
[6] Haastrup S *et al* 2018 *2D Mater.* **5** 042002
[7] Shivayogimath A *et al* 2019 *Nat. Commun.* **10** 1–7
[8] Zhou J *et al* 2018 *Nature* **556** 355–9
[9] Anasori B, Lukatskaya M R and Gogotsi Y 2017 *Nat. Rev. Mater.* **2** 1–17
[10] Dou L *et al* 2015 *Science* **349** 1518–21
[11] Mounet N *et al* 2018 *Nat. Nanotechnol.* **13** 246–52
[12] Ashton M, Paul J, Sinnott S B and Hennig R G 2017 *Phys. Rev. Lett.* **118** 106101
[13] Geim A K and Grigorieva I V 2013 *Nature* **499** 419–25
[14] Cao Y, Fatemi V, Fang S, Watanabe K, Taniguchi T, Kaxiras E and Jarillo-Herrero P 2018 *Nature* **556** 43–50
[15] Bistritzer R and MacDonald A H 2011 *Proc. Natl Acad. Sci.* **108** 12233–7
[16] Zhao X *et al* 2020 *Nature* **581** 171–7
[17] Wan J, Lacey S D, Dai J, Bao W, Fuhrer M S and Hu L 2016 *Chem. Soc. Rev.* **45** 6742–65
[18] Wilkinson M D *et al* 2016 *Sci. Data* **3** 1–9
[19] Wirtz L, Marini A and Rubio A 2006 *Phys. Rev. Lett.* **96** 126104
[20] Cudazzo P, Tokatly I V and Rubio A 2011 *Phys. Rev. B* **84** 085406
[21] Klots A *et al* 2014 *Sci. Rep.* **4** 6608
[22] Chernikov A, Berkelbach T C, Hill H M, Rigosi A, Li Y, Aslan O B, Reichman D R, Hybertsen M S and Heinz T F 2014 *Phys. Rev. Lett.* **113** 076802
[23] Olsen T, Latini S, Rasmussen F and Thygesen K S 2016 *Phys. Rev. Lett.* **116** 056401
[24] Riis-Jensen A C, Gjerding M N, Russo S and Thygesen K S 2020 *Phys. Rev. B* **102** 201402
[25] Felipe H, Xian L, Rubio A and Louie S G 2020 *Nat. Commun.* **11** 1–10
[26] Sohier T, Gibertini M, Calandra M, Mauri F and Marzari N 2017 *Nano Lett.* **17** 3758–63
[27] Ugeda M M *et al* 2014 *Nat. Mater.* **13** 1091–5
[28] Winther K T and Thygesen K S 2017 *2D Mater.* **4** 025059
[29] Wang Z, Rhodes D A, Watanabe K, Taniguchi T, Hone J C, Shan J and Mak K F 2019 *Nature* **574** 76–80
[30] Gong C *et al* 2017 *Nature* **546** 265–9
[31] Huang B *et al* 2017 *Nature* **546** 270–3
[32] Chang K *et al* 2016 *Science* **353** 274–8
[33] Olsen T, Andersen E, Okugawa T, Torelli D, Deilmann T and Thygesen K S 2019 *Phys. Rev. Mater.* **3** 024005
[34] Marrazzo A, Gibertini M, Campi D, Mounet N and Marzari N 2019 *Nano Lett.* **19** 8431–40
[35] Thygesen K S 2017 *2D Mater.* **4** 022004
[36] Torelli D and Olsen T 2018 *2D Mater.* **6** 015028
[37] Rasmussen F A and Thygesen K S 2015 *J. Phys. Chem. C* **119** 13169–83
[38] Enkovaara J *et al* 2010 *J. Phys.: Condens. Matter.* **22** 253202
[39] Mortensen J J, Hansen L B and Jacobsen K W 2005 *Phys. Rev. B* **71** 035109
[40] Gjerding M, Skovhus T, Rasmussen A, Bertoldo F, Larsen A H, Mortensen J J and Thygesen K S 2021 Atomic simulation recipes—a python framework and library for automated workflows (arXiv:2104.13431)
[41] Larsen A H *et al* 2017 *J. Phys.: Condens. Matter.* **29** 273002
[42] Mortensen J J, Gjerding M and Thygesen K S 2020 *J. Open Source Softw.* **5** 1844
[43] Saal J E, Kirklin S, Aykol M, Meredig B and Wolverton C 2013 *JOM* **65** 1501–9
[44] Jain A *et al* 2013 *APL Mater.* **1** 011002
[45] Curtarolo S *et al* 2012 *Comput. Mater. Sci.* **58** 218–26
[46] Ataca C, Sahin H and Ciraci S 2012 *J. Phys. Chem. C* **116** 8983–99
[47] Lebègue S, Björkman T, Klintenberg M, Nieminen R M and Eriksson O 2013 *Phys. Rev. X* **3** 031002
[48] Kormányos A, Burkard G, Gmitra M, Fabian J, Zólyomi V, Drummond N D and Fal'ko V 2015 *2D Mater.* **2** 022001
[49] Zhou J *et al* 2019 *Sci. Data* **6** 1–10
[50] Choudhary K, Kalish I, Beams R and Tavazza F 2017 *Sci. Rep.* **7** 1–16
[51] Larsen A H, Vanin M, Mortensen J J, Thygesen K S and Jacobsen K W 2009 *Phys. Rev. B* **80** 195112
[52] Larsen P M, Pandey M, Strange M and Jacobsen K W 2019 *Phys. Rev. Mater.* **3** 034003
[53] Ong S P *et al* 2013 *Comput. Mater. Sci.* **68** 314–19
[54] Patrick C E, Jacobsen K W and Thygesen K S 2015 *Phys. Rev. B* **92** 201205
[55] Kirklin S, Saal J E, Meredig B, Thompson A, Doak J W, Aykol M, Rühl S and Wolverton C 2015 *npj Computat. Mater.* **1** 1–15
[56] Perdew J P, Burke K and Ernzerhof M 1996 *Phys. Rev. Lett.* **77** 3865–8
[57] Maździarz M 2019 *2D Mater.* **6** 048001
[58] Li Y and Heinz T F 2018 *2D Mater.* **5** 025021
[59] Hadley L N and Dennison D 1947 *J. Opt. Soc. Am.* **37** 451–65
[60] Wang G, Chernikov A, Glazov M M, Heinz T F, Marie X, Amand T and Urbaszek B 2018 *Rev. Mod. Phys.* **90** 021001
[61] Lu A Y *et al* 2017 *Nat. Nanotechnol.* **12** 744–9
[62] Zhang J *et al* 2017 *ACS Nano* **11** 8192–8
[63] Fülöp B *et al* 2018 *2D Mater.* **5** 031013
[64] Riis-Jensen A C, Deilmann T, Olsen T and Thygesen K S 2019 *ACS Nano* **13** 13354
[65] Bychkov Y A and Rashba E I 1984 *J. Phys. C: Solid State Phys.* **17** 6039
[66] Petersen L and Hedegård P 2000 *Surf. Sci.* **459** 49–56
[67] Bergerhoff G, Brown I and Allen F *et al* 1987 *Int. Union Crystallogr., Chester* **360** 77–95
[68] Gražulis S *et al* 2012 *Nucleic Acids Res.* **40** D420–7
[69] Qian X, Liu J, Fu L and Li J 2014 *Science* **346** 1344–7
[70] Torelli D, Moustafa H, Jacobsen K W and Olsen T 2020 *npj Comput. Mater.* **6** 158
[71] Mak K F, Lee C, Hone J, Shan J and Heinz T F 2010 *Phys. Rev. Lett.* **105** 136805
[72] Splendiani A, Sun L, Zhang Y, Li T, Kim J, Chim C Y, Galli G and Wang F 2010 *Nano Lett.* **10** 1271–5
[73] Xiao J *et al* 2020 *Nat. Phys.* **16** 1028–34
[74] Sivadas N, Okamoto S, Xu X, Fennie C J and Xiao D 2018 *Nano Lett.* **18** 7658–64
[75] Liu Y, Wu L, Tong X, Li J, Tao J, Zhu Y and Petrovic C 2019 *Sci. Rep.* **9** 1–8
[76] Yasuda K, Wang X, Watanabe K, Taniguchi T and Jarillo-Herrero P 2020 (arXiv:2010.06600)
[77] Grimme S, Antony J, Ehrlich S and Krieg H 2010 *J. Chem. Phys.* **132** 154104
[78] Northup T and Blatt R 2014 *Nat. Photon.* **8** 356–63
[79] O'Brien J, Furusawa A and Vuckovic J 2009 Photonic quantum technologies nat *Photonics* **3** 687
[80] Zhang S and Northrup J E 1991 *Phys. Rev. Lett.* **67** 2339

95

[81] Van de Walle C G, Laks D, Neumark G and Pantelides S 1993 *Phys. Rev.* B **47** 9425

[82] Janak J F 1978 *Phys. Rev.* B **18** 7165

[83] Pandey M, Rasmussen F A, Kuhar K, Olsen T, Jacobsen K W and Thygesen K S 2016 *Nano Lett.* **16** 2234–9

[84] Kaappa S, Malola S and Häkkinen H 2018 *J. Phys. Chem.* A **122** 8576–84

[85] Levi G, Ivanov A V and Jonsson H 2020 *Faraday Discuss.* **224** 448-66

[86] Grimme S, Antony J, Ehrlich S and Krieg H 2010 *J. Chem. Phys.* **132** 154104

[87] Bader R F W 1990 *Atoms in Molecules: A Quantum Theory (The Int. Series of Monographs on Chemistry* vol 22) (Oxford: Clarendon)

[88] Tang W, Sanville E and Henkelman G 2009 *J. Phys.: Condens. Matter.* **21** 084204

[89] Resta R 1992 *Ferroelectrics* **136** 51–5

[90] King-Smith R D and Vanderbilt D 1993 *Phys. Rev.* B **47** 3

[91] Zhang S and Yu F 2011 *J. Am. Ceram. Soc.* **94** 3153–70

[92] Maeder M D, Damjanovic D and Setter N 2004 *J. Electroceram.* **13** 385–92

[93] Scott J F 2000 *Ferroelectric Memories* vol 3 (Berlin: Springer)

[94] Rangel T, Fregoso B M, Mendoza B S, Morimoto T, Moore J E and Neaton J B 2017 *Phys. Rev. Lett.* **119** 067402

[95] Resta R and Vanderbilt D 2007 Theory of Polarization: A Modern Approach *Phys. Ferroelectr.* vol 105 (Berlin: Springer) pp 31–68

[96] Gjerding M N, Cavalcante L S R, Chaves A and Thygesen K S 2020 *J. Phys. Chem.* C **124** 11609–16

[97] Ye Z G 2008 *Handbook of Advanced Dielectric, Piezoelectric and Ferroelectric Materials: Synthesis, Properties and Applications* (Amsterdam: Elsevier)

[98] Ogawa T 2016 *Piezoelectric Materials* (Croatia: InTech)

[99] Vanderbild D 1999 *J. Phys. Chem. Solids* **61** 147–51

[100] Authier A 2003 *Int. Tables for Crystallography: Volume D: Physical Properties of Crystals* (Dordrecht: Springer)

[101] Duerloo K A N, Ong M T and Reed E J 2012 *J. Phys. Chem. Lett.* **3** 2871–6

[102] Zhu H *et al* 2015 *Nat. Nanotechnol.* **10** 151–5

[103] Taherinejad M, Garrity K F and Vanderbilt D 2014 *Phys. Rev.* B **89** 115102

[104] Olsen T 2016 *Phys. Rev.* B **94** 235106

[105] Fu L 2011 *Phys. Rev. Lett.* **106** 106802

[106] Olsen T, Andersen E, Okugawa T, Torelli D, Deilmann T and Thygesen K S 2019 *Phys. Rev. Mater.* **3** 024005

[107] Benalcazar W A, Bernevig B A and Hughes T L 2017 *Phys. Rev.* B **96** 245115

[108] Mermin N D and Wagner H 1966 *Phys. Rev. Lett.* **17** 1133–6

[109] Olsen T 2019 *MRS Commun.* **9** 1142–50

[110] Olsen T 2017 *Phys. Rev.* B **96** 125143

[111] Torelli D and Olsen T 2020 *J. Phys.: Condens. Matter.* **32** 335802

[112] Lado J L and Fernández-Rossier J 2017 *2D Mater.* **4** 035002

[113] Torelli D, Thygesen K S and Olsen T 2019 *2D Mater.* **6** 045018

[114] Long D A 2002 *The Raman Effect: A Unified Treatment of the Theory of Raman Scattering by Molecules* (Chichester: Wiley)

[115] Taghizadeh A, Leffers U, Pedersen T G and Thygesen K S 2020 *Nat. Commun.* **11** 3011

[116] Lee S Y and Heller E J 1979 *J. Chem. Phys.* **71** 4777

[117] Umari P and Pasquarello A 2003 *J. Phys. Condens. Matter* **15** S1547–52

[118] Taghizadeh A, Hipolito F and Pedersen T G 2017 *Phys. Rev.* B **96** 195413

[119] Prylepa A *et al* 2018 *J. Phys. D: Appl. Phys* **51** 043001

[120] Taghizadeh A, Thygesen K S and Pedersen T G 2021 *ACS Nano* **15** 7155

[121] Aversa C and Sipe J E 1995 *Phys. Rev.* B **52** 14636–45

[122] Miller R C 1964 *Appl. Phys. Lett.* **5** 17–19

[123] Schmidt J, Marques M R G, Botti S and Marques M A L 2019 *Computat. Mater.* **5** 83

[124] Zhuo Y, Mansouri Tehrani A and Brgoch J 2018 *J. Phys. Chem. Lett.* **9** 1668–73

[125] Faber F, Lindmaa A, von Lilienfeld O A and Armiento R 2015 *Int. J. Quantum Chem.* **115** 1094–101

[126] Ward L, Agrawal A, Choudhary A and Wolverton C 2016 *Computat. Mater.* **2** 1–7

[127] Ghiringhelli L M, Vybiral J, Levchenko S V, Draxl C and Scheffler M 2015 *Phys. Rev. Lett.* **114** 105503

[128] Rupp M, Tkatchenko A, Müller K R and Von Lilienfeld A 2012 *Phys. Rev. Lett.* **108** 058301

[129] Huo H and Rupp M 2018 Unified representation of molecules and crystals for machine learning (arXiv:1704.06439)

[130] Jorgensen P B, Mesta M, Shil S, García Lastra J M, Jacobsen K W, Thygesen K S and Schmidt M N 2018 *J. Chem. Phys.* **148** 241735

[131] Rajan A C, Mishra A, Satsangi S, Vaish R, Mizuseki H, Lee K R and Singh A K 2018 *Chem. Mater.* **30** 4031–8

96

Paper II: High-throughput assessment of structural instabilities in 2D materials

Simone Manti, Kristian S. Thygesen
   Submitted

High-throughput assessment of structural instabilities in 2D materials

Simone Manti, Kristian S. Thygesen

CAMD, Department of Physics, Technical University of Denmark, DK - 2800 Kongens Lyngby, Denmark Center for Nanostructured Graphene (CNG), Department of Physics, Technical University of Denmark, DK - 2800 Kongens Lyngby, Denmark

Abstract. The dynamical stability of a crystal structure at zero temperature is usually established by asserting that all phonon frequencies are real throughout the Brillouin zone (BZ). Here we present empirical evidence that it is often sufficient to check the phonon frequencies at the center and boundary of the BZ. For reasonably simple crystals, this is a much more manageable test that is amenable to high-throughput implementation. Using two-dimensional (2D) crystals as a test ground, we find that the simple test correctly predicts the dynamical stability of 2D materials with above 90% success rate. For 137 dynamically unstable 2D materials, we displace the atoms along an unstable mode, relax the structure, and repeat the test for dynamical stability. This procedure yields a dynamically stable crystal in 47% or 13% of the cases, depending on whether the Hessian is isotropic or not. The properties of the distorted, dynamically stable materials are unraveled using the workflow behind the Computational 2D Materials Database (C2DB), and it is found that they can differ significantly from those of the original high-symmetry crystals, e.g. band gaps are opened by 0.3 eV on average. All the crystal structures and their calculated properties are available in the C2DB.

1. Introduction

Computational materials discovery aims at identifying novel types of materials for specific applications often employing first principles methods such as density functional theory (DFT) [1]. The potential of a given material for the targeted application is usually evaluated based on elementary properties calculated for the perfect crystal, e.g. the electronic band gap, the optical absorption spectrum, or the magnetic order. Such properties can be highly sensitive to even small distortions of the lattice or the atomic positions, if they reduce the symmetry of the crystal, and it is therefore important to develop efficient methods for identifying and accounting for such distortions.

Lattice distortions can be classified according to their periodicity relative to the primitive cell of the crystal. Local instabilities conserve the periodicity of the crystal, i.e. they do not enlarge the number of atoms in the primitive cell. Other distortions, known as charge density wave (CDW) [2], can extend over many repetitions of the primitive cell, i.e. they lead to an enlargement of the period of the crystal. A general microscopic description of the CDW is still missing due to the many possible driving mechanisms behind the CDW state, e.g. electron-phonon interaction [3], doping [4], or phonon-phonon interactions [5]. As a testimony to the complexity of the problem, different models and concepts are used to describe the CDW phase depending on the dimensionality of the material [6, 7, 8, 9].

The last few years have witnessed an increased interest in CDW states of two-dimensional (2D) materials. For example, CDW physics is believed to govern the transition from the trigonal prismatic T-phase to the lower symmetry T'-phase in monolayer MoS_2 [10] as well as the plethora of temperature dependent phases in monolayers of $NbSe_2$ [11, 12], TaS_2 [13, 14], $TaSe_2$ [15, 16], and $TiSe_2$ [17, 18]. In addition, a number of recent studies have investigated the possibility to control CDW phase transitions. For instance, the T-phase of monolayer MoS_2 can be stabilized by argon bombardment [19], exposure to electron beams [10], or Li-ion intercalation [20]. Similar results have been reported for $MoTe_2$ [21].

Regardless of the fundamental microscopic origin of lattice distortions it remains of great practical importance to devise efficient schemes that makes it possible to verify whether or not a given structure is dynamically stable, i.e. whether it represents a local minimum of the potential energy surface. Structures that are not dynamically stable are frequently generated in computational studies, e.g. when a structure is relaxed under symmetry constraints or the chosen unit cell is too small. Tests for dynamical stability are rarely performed in large-scale discovery studies, because there is no established way of doing it apart from calculating the phonon band structure[22], which is a time-consuming task. At the same time, the importance of incorporating such tests is in fact unclear; that is, it is not known how much symmetry-breaking distortions generally influence the properties of a material close to the convex hull.

A straightforward strategy to generate potentially stable structures from dynamically unstable ones, is to displace the atoms along an unstable phonon mode using a supercell that can accommodate the distortion. This approach has previously been adopted to explore structural distortions in bulk perovskites [23, 24] and one-dimensional organometallic chains [25]. However, systematic studies of structural instabilities in 2D materials, have so far been lacking.

In this work, we perform a systematic study of structural distortions in thermodynamically stable two-dimensional (2D) materials from the Computational 2D Materials Database (C2DB) [26, 27]. We focus on small-period distortions that can be accommodated in a 2×2 supercell, and refer to the test for the occurrence of such distortions as the Center and Boundary Phonon (CBP) protocol. The motivation behind the present work is threefold: (i) To assess the reliability of the CBP protocol for 2D materials. (ii) To elucidate the effect of symmetry-breaking distortions on the basic electronic properties. (iii) To obtain the dynamically stable phases of a set of dynamically unstable 2D materials that were originally generated by combinatorial lattice decoration, and make them available to the community via the C2DB.

The paper is structured as follows. In Section 2 we describe the CBP protocol. In Section 3 we first benchmark the CBP protocol against full phonon band structure calculations and evaluate its statistical success rate. For 137 dynamically unstable 2D materials, we further analyse how the small-period distortions that stabilise the materials influence their electronic properties. Section 4 concludes the paper.

2. Methodology

In this Section we briefly discuss the CBP protocol for testing the dynamical stability of a crystal and for generating distorted and potentially stable crystal structures. We also describe the methodology and computational details of the phonon calculations.

2.1. The CBP protocol: Stability test

Given a material that has been relaxed in some unit cell (from hereon referred to as the primitive unit cell), the CBP protocol proceeds by evaluating the stiffness tensor of the material and the Hessian matrix of a supercell obtained by repeating the primitive cell 2×2
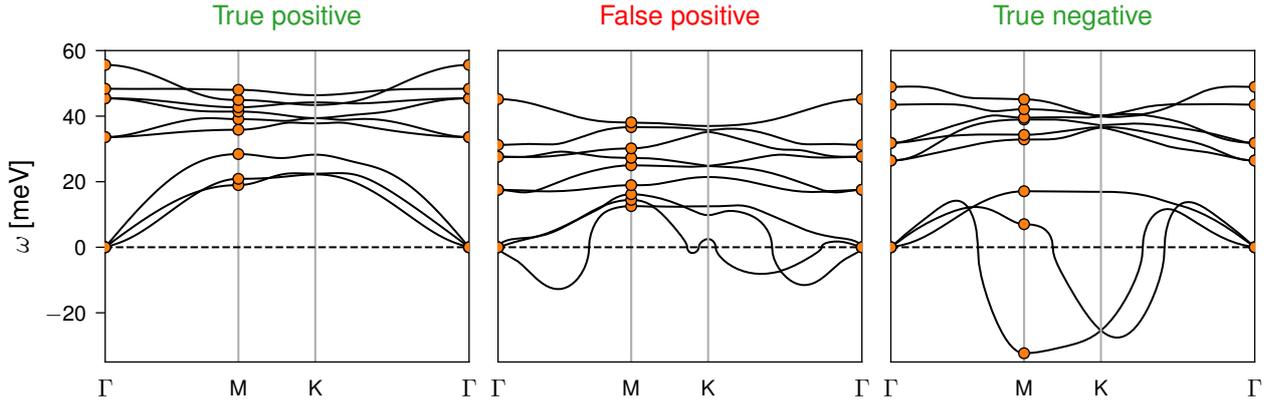
Figure 1. Phonon band structure for monolayer MoS_2 in the H-phase (left), NbSSe (middle), and MoS_2 in the T-phase (right). Note that imaginary phonon frequencies are represented by negative values. The CBP protocol (orange dots) is sufficient to conclude that a material is dynamically stable (unstable) in the situations depicted in the left (right) panels. In contrast, when the relevant distortion requires a supercell larger than a 2×2, and the phonon frequencies are real at the center and boundary of the BZ, the CBP protocol will result in a false positive result.

times. In the current work, the stiffness tensor is calculated as a finite difference of the stress under an applied strain, while the Hessian matrix is calculated as a finite difference of the forces on all the atoms of the 2×2 supercell under displacement of the atoms in one primitive unit cell (this is equivalent to calculating the phonons at the center and specific high symmetry points at boundary of the BZ of the primitive cell, see Fig. (3). Next, the stiffness tensor and the Hessian matrix are diagonalised, and the eigenvalues are used to infer a structural stability. A negative eigenvalue of the stiffness tensor indicates an instability of the lattice (the shape of the unit cell) while a negative eigenvalue of the 2×2 Hessian signals an instability of the atomic structure. The obvious question here, is whether it suffices to consider the Hessian of the 2×2 supercell, or equivalently consider the phonons at the BZ center and boundaries.

All phonon calculations were performed using the `asr.phonopy` recipe of the Atomic Simulation Recipes (ASR) [28], which makes use of the Atomic Simulation Environment (ASE)[29] and PHONOPY [30]. The DFT calculations were performed with the GPAW[31] code and the Perdew-Burke-Ernzerhof (PBE) exchange-correlation functional [32]. The BZ was sampled on a uniform k-point mesh of density of 6.0 $Å^2$ and the plane wave cutoff was set to 800 eV. To evaluate the Hessian matrix, the small displacement method was used with a displacement size of 0.01 Å and forces were converged up to 10^{-4} eV/Å. To benchmark the CBP protocol, we compare to full phonon band structures. In these calculations, the size of the supercell is chosen such that the Hessian matrix includes interactions between pairs of atoms within a radius of at least 12 Å. (This implies that the supercell

must contain a sphere of radius 12 Å).

We can distinguish three possible outcomes when comparing the CBP protocol against full phonon calculations (see Figure (1)), namely a true positive result, a true negative result, and a false positive result. We note that the case of a false negative is not possible, because a material that is unstable in a 2×2 cell is de facto unstable. The false positive case occurs when a material is stable in a 2×2 supercell, but unstable if allowed to distort in a larger cell. Our results show that such large-period distortions that do not show as distortions in a 2×2 cell, are relatively rare (see Section 3.1).

2.2. The CBP protocol: Structure generation

Here we outline a simple procedure to generate distorted and potentially stable structures from an initial dynamically unstable structure. The basic idea is to displace the atoms along an unstable phonon mode followed by a relaxation. In practice, the unstable mode is obtained as the eigen function corresponding to a negative eigenvalue of the Hessian matrix of the 2×2 supercell. The procedure is illustrated in Figure (2) for the well known T-T' phase transition of MoS_2 [10]. The left panel shows the atomic structure and phonon band structure of monolayer MoS_2 in the T-phase. Both the primitive unit cell (black) and the 2×2 supercell (orange) are indicated. The CBP method identifies an unstable mode at the BZ boundary (M point). After displacing the atoms along the unstable mode, a distorted structure is obtained, which after relaxation leads to the dynamically stable T'-phase of MoS_2 shown in the right panel.

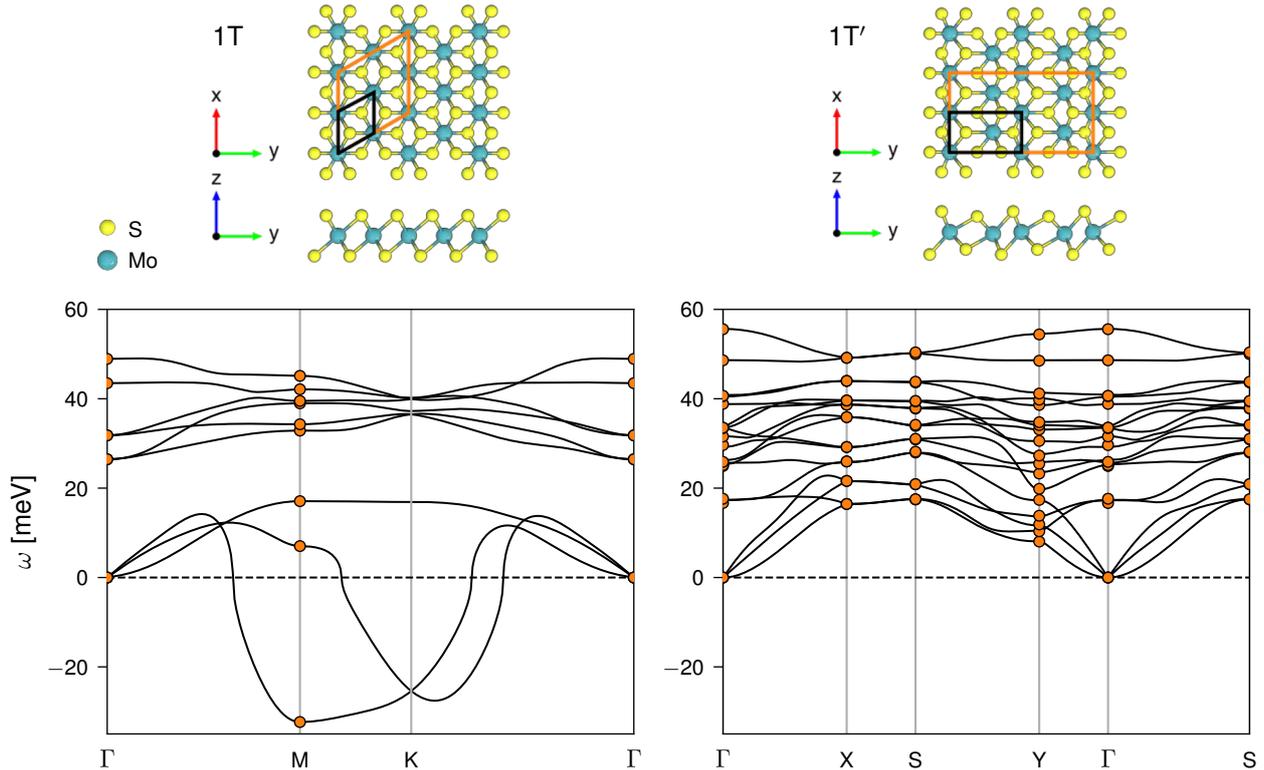In this work, we have applied the method system-

Figure 2. The CBP protocol captures the instability of the T-phase (left) of MoS_2. Both the primitive unit cell (black) and the 2×2 supercell (orange) are shown. Displacing the atoms along the unstable TA mode at the M-point ($\mathbf{q} = (\frac{1}{2}, 0)$), which can be accommodated in the 2×1 supercell, and subsequently relaxing the structure results in the dynamically stable T'-phase (right).

atically to 137 dynamically unstable 2D materials. The 137 monolayers were selected from the C2DB according to the following two criteria: First, to ensure that all materials are chemically "reasonable", only materials with a low formation energy were selected. Specifically, we require that $\Delta H_{\text{hull}} < 0.2$ eV/atom, where ΔH_{hull} is the energy above the convex hull defined by the most stable (possibly mixed) bulk phases of the relevant composition[33, 27]. Secondly, we consider only materials with exactly one unstable mode, i.e. one negative eigenvalue of the Hessian matrix at a given q-point.

The 137 dynamically unstable materials were displaced along the only unstable mode. The size of the displacement was chosen such that the maximum atomic displacement was exactly 0.1 Å. This displacement size was chosen based on the MoS_2 example discussed above, where it results in a minimal number of subsequent relaxation steps. A smaller value does not guarantee that the system leaves the saddle point, while a larger value creates a too large distortion resulting in additional relaxation steps. During relaxation the unit cell was allowed to change with no symmetry constraints and the relaxation was

stopped when the forces on all atoms were below 0.01 eV / Å.

3. Results

3.1. Assessment of the CBP protocol

To test the validity the CBP protocol, we have performed full phonon calculations for a set of 20 monolayers predicted as dynamically stable by the CBP protocol. The 20 materials were obtained from the C2DB and cover 7 different crystal structures. Out of the 20 materials 10 are metals and 10 are insulators/semiconductors. The calculated phonon band structures are reported in the supplementary material. For all materials, the phonon frequencies obtained with the CBP protocol equal the frequencies of the full phonon band structure at the q-points $\mathbf{q} \in \{(0,0), (\frac{1}{2}, 0), (0, \frac{1}{2}), (\frac{1}{2}, \frac{1}{2})\}$. This is expected as the phonons at these q-points can be accommodated by the 2×2 supercell.

Within the set of 20 materials, we find three False-positive cases, namely $CoTe_2$, NbSSe, and $TaTe_2$. This small percentage in our representative samples
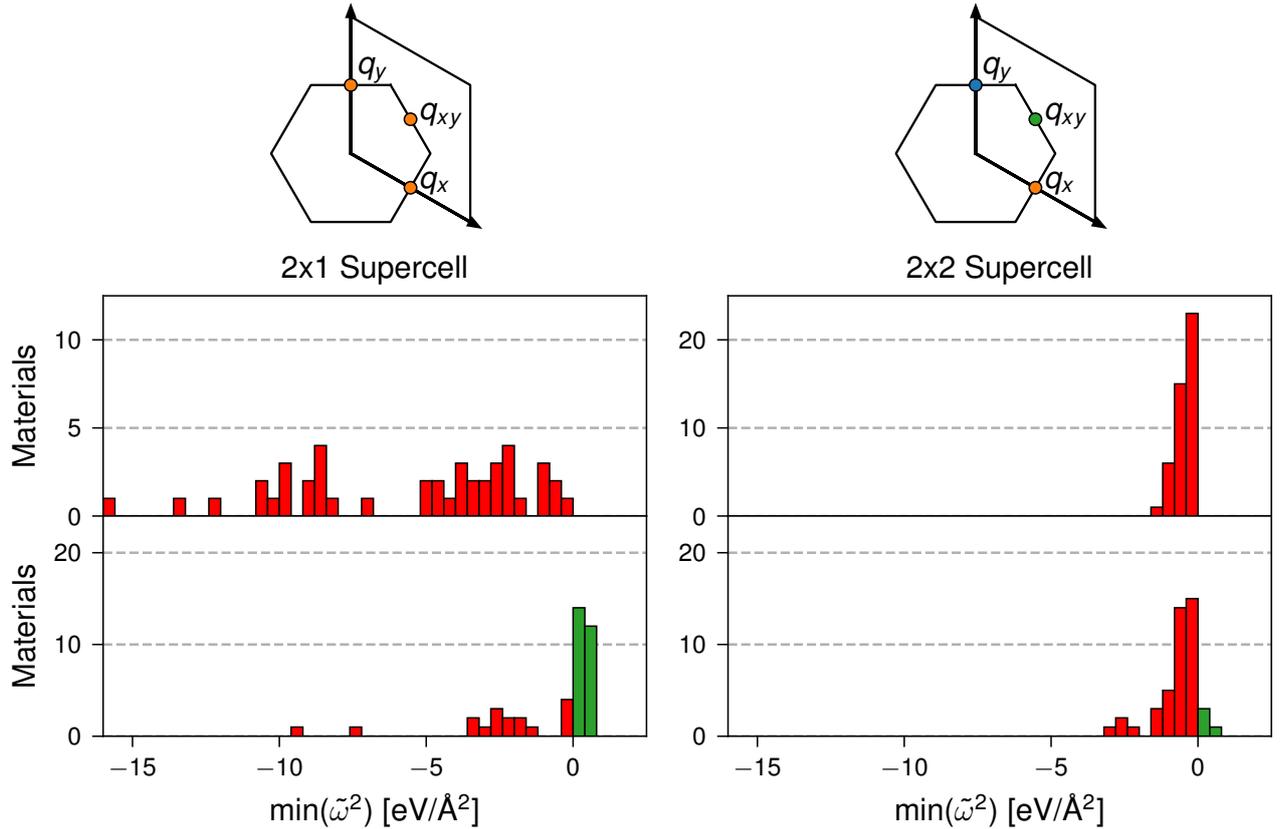
Figure 3. The 137 dynamically unstable 2D materials studied in this work can be divided into two groups depending on whether the negative eigenvalues of the Hessian matrix at $q = \{(\frac{1}{2}, 0), (0, \frac{1}{2}), (\frac{1}{2}, \frac{1}{2})\}$ are equal (left panel) or different (right panel). For the first group of materials, displacing the atoms along the mode at $q = (\frac{1}{2}, 0)$ and relaxing in a 2×1 supercell, yields a dynamically stable structure in 91/43 cases. For the second group, displacing the atoms along $q = (\frac{1}{2}, \frac{1}{2})$ and relaxing in a 2×2 supercell, yields a dynamically stable structure in 6/46 cases.

is in agreement with previous studies [22]. In that case out of 250 monolayers, there are 14 False-positive cases and half of them are well known materials in the H or T phase with metals like Co, Nb and Ta. In general, these materials exhibit unstable modes (imaginary frequencies or equivalently negative force constant eigenvalues) in the interior of the BZ (NbSSe and TaTe$_2$) or at the K-point (CoTe$_2$), while all phonon frequencies at the q-points covered by the CBP protocol, are real.

We note that the small imaginary frequencies in the out of plane modes around the Γ-point seen in some of the phonon band structures are not distortions, but are rather due to the interpolation of the dynamical matrix. In particularly, these artifacts occur because of the broken crystal point-group symmetry in the force constant matrix and they will vanish if a larger supercell is used or the rotational sum rule is imposed [34].

3.2. Stable distorted monolayers

The 137 dynamically unstable materials, which were selected from the C2DB according to the criteria described in Section 2.2, can be divided into two groups depending on whether the eigenvalues of the Hessian at the wave vectors $q_x = (\frac{1}{2}, 0)$, $q_y = (0, \frac{1}{2})$ and $q_{xy} = (\frac{1}{2}, \frac{1}{2})$, are equal or not. Equality of the eigenvalues implies an isotropic Hessian. For such materials, we generate distorted structures by displacing the atoms along the unstable mode at $q_x = (\frac{1}{2}, 0)$, followed by relaxation in a 2×1 supercell. In the case of an anisotropic Hessian, the atoms were displaced along $q_{xy} = (\frac{1}{2}, \frac{1}{2})$ and relaxed in a 2×2 supercell.

After atomic displacement and subsequent relaxation, the CBP protocol was applied again to test for dynamical stability of the distorted structures. Histograms of the minimum eigenvalue of the Hessian matrix are shown in Figure (3) with the materials before and after atomic displacement shown in the upper and lower panels, respectively. Negative eigenvalues, corre-

Material	Space group - Wyckoff		ΔH_{hull} [eV/atom]		$\min(\tilde{\omega}^2)$ [eV/Å^2]		$\varepsilon_{\text{gap}}^{\text{PBE}}$ [eV]	
	before	after	before	after	before	after	before	after
AgBr$_2$	187-bi	1-a	0.05	0.00	-1.01	0.00	0.00	0.00
AgCl$_2$	164-bd	1-a	0.05	0.00	-2.60	0.06	0.00	0.00
AsClTe	156-ac	1-a	0.19	0.02	-0.51	0.25	1.29	1.48
CdBr	2-i	1-a	0.18	0.07	-1.00	0.03	0.00	1.28
CdCl	156-ab	1-a	0.17	0.04	-0.74	0.17	0.00	1.67
CoSe	164-bd	2-i	0.05	0.03	-4.82	0.09	0.00	0.00
CrBrCl	156-abc	1-a	0.11	0.06	-0.98	0.15	0.00	0.64
CrBr$_2$	164-bd	1-a	0.10	0.06	-0.59	0.07	0.00	0.49
CrCl$_2$	164-bd	1-a	0.11	0.05	-1.87	0.15	0.00	0.76
CrSSe	156-abc	1-a	0.15	0.09	-9.75	0.71	0.00	0.00
CrS$_2$	156-abc	2-i	0.18	0.05	-15.62	0.74	0.00	0.00
CrSe$_2$	156-abc	2-i	0.14	0.05	-12.17	0.77	0.00	0.00
CrTe$_2$	156-abc	2-i	0.02	0.01	-1.71	0.08	0.00	0.00
CrPS$_3$	3-1-a	1-a	0.09	0.03	-3.13	0.00	0.00	0.34
FePSe$_3$	3-1-a	1-a	0.13	0.12	-0.42	0.04	0.13	0.13
FeSe$_2$	187-bi	1-a	0.15	0.00	-2.04	0.00	0.00	0.00
HfBrCl	156-abc	1-a	0.14	0.03	-9.61	0.40	0.00	0.82
HfBrI	156-abc	1-a	0.22	0.05	-10.41	0.39	0.00	0.73
HfBr$_2$	164-bd	2-i	0.14	0.04	-10.21	0.41	0.00	0.8
HfCl$_2$	164-bd	2-i	0.14	0.04	-8.51	0.38	0.00	0.85
HgSe	156-ab	1-a	0.11	0.04	-0.83	0.04	0.08	0.37
HgTe	156-ab	1-a	0.11	0.04	-0.61	0.04	0.08	0.37
InTe	156-ab	1-a	0.18	0.10	-0.41	0.23	0.00	0.00
InBrSe	59-ab	2-i	0.03	0.02	-0.50	0.04	1.23	1.23
InSe	187-hi	6-ab	0.00	0.00	-0.28	0.01	1.39	1.39
MoSeTe	156-abc	1-a	0.20	0.08	-10.67	0.69	0.00	0.00
MoTe$_2$	164-bd	2-i	0.17	0.01	-13.23	0.67	0.00	0.00
NbS$_2$	187-bi	6-ab	0.00	0.00	-1.08	0.06	0.00	0.00
NbTe$_2$	187-bi	1-a	0.00	0.00	-0.37	0.51	0.00	0.00
PdI$_2$	164-bd	1-a	0.17	0.03	-0.56	0.12	0.00	0.59
RhI$_2$	164-bd	1-a	0.17	0.17	-0.64	0.04	0.00	0.00
RhO$_2$	164-bd	2-i	0.16	0.15	-2.19	0.65	0.00	0.00
RhTe$_2$	164-bd	1-a	0.11	0.07	-1.67	0.36	0.00	0.13
ScI$_3$	162-dk	1-a	0.00	0.00	-0.25	0.02	1.85	1.85
TiBrCl	156-abc	1-a	0.05	0.00	-9.15	0.52	0.00	0.29
TiBr$_2$	164-bd	1-a	0.06	0.04	-0.54	0.30	0.00	0.12
TiCl$_2$	164-bd	2-i	0.11	0.00	-9.99	0.53	0.00	0.32
TiO$_2$	164-bd	6-ab	0.14	0.12	-1.96	0.47	2.70	2.85
TiS$_2$	187-bi	4-a	0.14	0.14	-0.49	0.00	0.73	0.79
TiPSe$_3$	1-a	4-a	0.16	0.00	-1.24	0.00	0.00	0.00
VTe$_2$	164-bd	6-ab	0.02	0.00	-1.05	0.38	0.00	0.00
ZrBrCl	156-abc	1-a	0.12	0.02	-8.05	0.40	0.00	0.59
ZrBrI	156-abc	1-a	0.15	0.02	-8.89	0.38	0.00	0.48
ZrBr$_2$	164-bd	2-i	0.11	0.00	-8.65	0.47	0.00	0.59
ZrClI	156-abc	1-a	0.19	0.07	-8.69	0.26	0.00	0.48
ZrCl$_2$	164-bd	1-a	0.11	0.03	-7.13	0.46	0.00	0.60
ZrI$_2$	164-bd	2-i	0.14	0.00	-8.59	0.48	0.00	0.43
ZrS$_2$	187-bi	2-i	0.19	0.18	-0.80	0.15	0.96	1.13

Table 1. Some of the calculated properties of the subset of the 137 materials that became dynamically stable after displacing the atoms along an unstable phonon mode. The properties are shown before and after the distortion, i.e. for the original dynamically unstable structures and the final dynamically stable structures, respectively.

sponding to unstable materials, are shown in red while positive eigenvalues are shown in green. Out of the 137 unstable materials, 49 become dynamically stable (according to the CBP protocol). By far the highest success rate for generating stable crystals was found for the isotropic materials (left panel), where 43 out of 91 materials became stable while only 6 out of the 43 anisotropic materials became stable.

A wide range of elementary properties of the 49 distorted, dynamically stable materials were computed using the C2DB workflow (see Table (1) in [27] for a complete list of the properties). The atomic structures together with the calculated properties are available in the C2DB. Table (1) provides an overview of the symmetries, minimal Hessian eigenvalues, total energies, and electronic band gap of the 49 materials before and after the distortion.

Apart from the reduction in symmetry, the distortion also lowers the total energy of the materials. An important descriptor for the thermodynamic
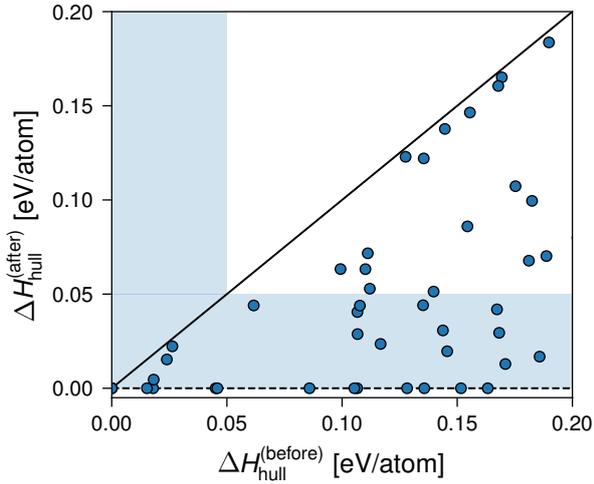
Figure 4. The energy above the convex hull for the 49 monolayers before and after distortion. Materials with a $\Delta H_{\mathrm{hull}}^{\mathrm{(after)}}$ close to zero are expected to be thermodynamically stable. The range up to 0.05 eV/atom above the convex hull has been indicated by a shaded blue region to visualise the importance of structural distortions for assessing the thermodynamic stability.

Figure 5. The average energy gain for the new stable materials is -0.067 eV / atom and the average gap opening is 0.29 eV.

stability of a material is the energy above the convex hull, ΔH_{hull}. Figure (4) shows a plot of ΔH_{hull} before and after the distortion of the 49 materials. The reduction in energy upon distortion ranges from 0 to 0.2 eV/atom. In fact, several of the materials come very close to the convex hull and some even fall onto the hull, indicating their global thermodynamic stability (at $T = 0$ K) with respect to the reference bulk phases. We note that all DFT energies, including the reference bulk phases, were calculated using the PBE xc-functional, which does not account for van der Waals interactions. Accounting for the vdW interactions will downshift the energies of layered bulk phases and thus increase ΔH_{hull} for the monolayers slightly. This effect will, however, not influence the relative stability of the pristine and distorted monolayers, which is the main focus of the current work.

Another characteristic trend observed is the opening/increase of the electronic band gap. The increase of the single-particle band gap is expected to be related to the total energy gained by making the distortion. Figure (5) shows the relation between the two quantities. Simplified models, for low dimensional systems and weak electron-phonon coupling, predict a proportionality between these two quantities [35]. From our results it is clear that there is no universal relationship between the change in band gap and total energy. In particular, several of the metals show large gain in total energy while the gap remains zero.

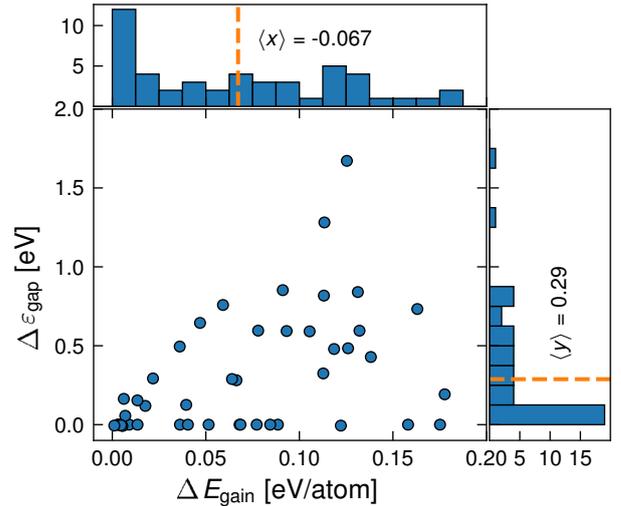It is interesting that within a threshold of 0.1 eV 21 of the distorted and dynamically stable materials

exhibit direct band gaps. Atomically thin direct band gap semiconductors are highly relevant as building blocks for opto-electronic or photonic devices, but only a hand full of such materials are known to date e.g. monolayers of transition metal dichalcogenides [36, 37] and black phosphorous [38]. As an example of a monolayer material that drastically changes from a metal to a direct band gap semiconductor upon distortion, we show the band structure of CdBr in Figure (6). The initial unstable metallic phase of the material becomes a dynamically stable upon distortion and opens a direct band gap of 1.28 eV at the C point.

4. Conclusions

In conclusion, we have performed a systematic study of structural instabilities in 2D materials. We have validated a simple protocol (here referred to as the CBP protocol) for identifying dynamical instabilities based on the frequency of phonons at the center and boundary of the BZ. The CBP protocol correctly classifies 2D materials as dynamically stable/unstable in 236 out of 250 cases[22] and is ideally suited for high-throughput studies where calculation of the full phonon band structure is not feasible. For 137 dynamically unstable monolayers with low formation energies from the C2DB database, we displaced the atoms along an unstable phonon mode and relaxed the structure in a 2×1 or 2×2 supercell. This resulted in 49 distorted, dynamically stable monolayers. The success rate of obtaining a dynamically stable structure from this protocol was found to be significantly higher for systems with only one (inequivalent) unstable
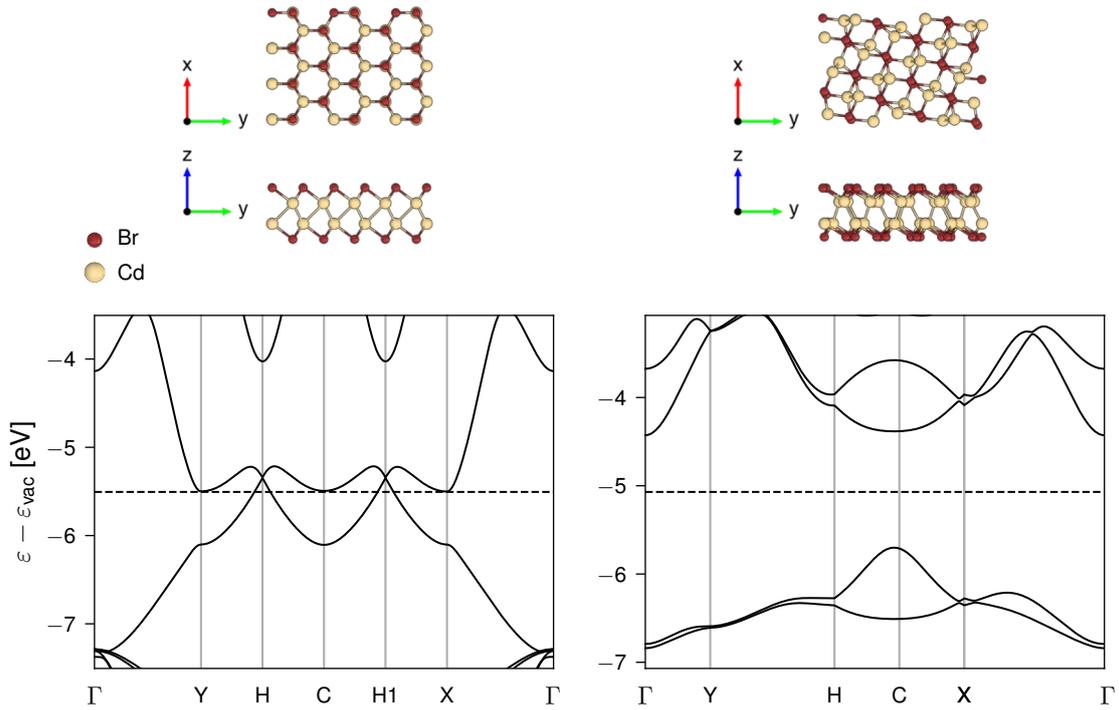
Figure 6. Extreme case of gap opening for the new stable material (CdBr) where the difference in the gap between the initial unstable metallic phase and the final structure is 1.28 eV.

phonon mode as compared to cases with several inequivalent modes where different displacements can lead to different, dynamically unstable structures. The 49 stable structures were fully characterised by an extensive computational property workflow and the results were made available via the C2DB database. The properties of the distorted structures can deviate significantly from the original high symmetry structures, and we found only a weak, qualitative relation between the gain in total energy and band gap opening upon distortion.

The current work has disclosed the zero temperature structure of 49 different 2D materials and examined the validity of the CBP protocol that should be useful for high-throughput assessment of dynamical stability. The work could form a useful starting point for future studies on structural instabilities in 2D materials focusing e.g. on the underlying driving mechanism or the effect of finite temperatures.

5. Acknowledgments

References

[1] Kohn W and Sham L J 1965 *Phys. Rev.* **140**(4A) A1133–A1138
[2] Grüner G 1988 *Rev. Mod. Phys.* **60**(4) 1129–1181
[3] Weber F, Rosenkranz S, Castellan J P, Osborn R, Karapetrov G, Hott R, Heid R, Bohnen K P and Alatas A 2011 *Phys. Rev. Lett.* **107**(26) 266401 URL https://link.aps.org/doi/10.1103/PhysRevLett.107.266401
[4] Zhou J S, Monacelli L, Bianco R, Errea I, Mauri F and Calandra M 2020 *Nano Letters* **20** 4809–4815 (*Preprint* https://doi.org/10.1021/acs.nanolett.0c00597)
[5] Bianco R, Errea I, Monacelli L, Calandra M and Mauri F 2019 *Nano Letters* **19** 3098–3103 (*Preprint* https://doi.org/10.1021/acs.nanolett.9b00504)
[6] Leroux M, Le Tacon M, Calandra M, Cario L, Méasson M A, Diener P, Borrissenko E, Bosak A and Rodière P 2012 *Phys. Rev. B* **86**(15) 155125
[7] Bianco R, Errea I, Monacelli L, Calandra M and Mauri F 2019 *Nano Letters* **19** 3098–3103
[8] Calandra M, Mazin I I and Mauri F 2009 *Phys. Rev. B* **80**(24) 241108
[9] Zhu X, Cao Y, Zhang J, Plummer E W and Guo J 2015

Proceedings of the National Academy of Sciences **112** 2367–2371

[10] Lin Y C, Dumcenco D O, Huang Y S and Suenaga K 2014 *Nature Nanotechnology* **9** 391–396 ISSN 1748-3395

[11] Xi X, Zhao L, Wang Z, Berger H, Forró L, Shan J and Mak K F 2015 *Nature Nanotechnology* **10** 765–769

[12] Ugeda M, Bradley A, Zhang Y, Onishi S, Chen Y, Ruan W, Ojeda-Aristizabal C, Ryu H, Edmonds M, Tsai H Z, Riss A, Mo S, Lee D, Zettl A, Hussain Z, Shen Z X and Crommie M 2016 *Nature Physics* **12** 92–97 ISSN 1745-2473

[13] Yang Y, Fang S, Fatemi V, Ruhman J, Navarro-Moratalla E, Watanabe K, Taniguchi T, Kaxiras E and Jarillo-Herrero P 2018 *Phys. Rev. B* **98**(3) 035203

[14] Yu Y, Yang F, Lu X F, Yan Y J, Cho Y H, Ma L, Niu X, Kim S, Son Y W, Feng D, Li S, Cheong S W, Chen X H and Zhang Y 2015 *Nature Nanotechnology* **10** 270–276 ISSN 1748-3395

[15] Ryu H, Chen Y, Kim H, Tsai H Z, Tang S, Jiang J, Liou F, Kahn S, Jia C, Omrani A A, Shim J H, Hussain Z, Shen Z X, Kim K, Min B I, Hwang C, Crommie M F and Mo S K 2018 *Nano Letters* **18** 689–694

[16] Ge Y and Liu A Y 2012 *Phys. Rev. B* **86**(10) 104101

[17] Sugawara K, Nakata Y, Shimizu R, Han P, Hitosugi T, Sato T and Takahashi T 2016 *ACS Nano* **10** 1341–1345

[18] Wang H, Chen Y, Duchamp M, Zeng Q, Wang X, Tsang S H, Li H, Jing L, Yu T, Teo E H T and Liu Z 2018 *Advanced Materials* **30** 1704382

[19] Zhu J, Wang Z, Yu H, Li N, Zhang J, Meng J, Liao M, Zhao J, Lu X, Du L, Yang R, Shi D, Jiang Y and Zhang G 2017 *Journal of the American Chemical Society* **139** 10216–10219

[20] Wang L, Xu Z, Wang W and Bai X 2014 *Journal of the American Chemical Society* **136** 6693–6697

[21] Krishnamoorthy A, Bassman L, Kalia R K, Nakano A, Shimojo F and Vashishta P 2018 *Nanoscale* **10**(6) 2742–2747

[22] Mounet N, Gibertini M, Schwaller P, Campi D, Merkys A, Marrazzo A, Sohier T, Castelli I E, Cepellotti A, Pizzi G and Marzari N 2018 *Nature Nanotechnology* **13** 246–252 ISSN 1748-3395

[23] Patrick C E, Jacobsen K W and Thygesen K S 2015 *Phys. Rev. B* **92**(20) 201205

[24] Yang R X, Skelton J M, da Silva E L, Frost J M and Walsh A 2020 *The Journal of Chemical Physics* **152** 024703

[25] Kayastha P and Ramakrishnan R 2021 *The Journal of Chemical Physics* **154** 061102

[26] Haastrup S, Strange M, Pandey M, Deilmann T, Schmidt P S, Hinsche N F, Gjerding M N, Torelli D, Larsen P M, Riis-Jensen A C, Gath J, Jacobsen K W, Mortensen J J, Olsen T and Thygesen K S 2018 *2D Materials* **5** 042002 URL

[27] Gjerding M N, Taghizadeh A, Rasmussen A, Ali S, Bertoldo F, Deilmann T, Holguin U P, Knøsgaard N R, Kruse M, Manti S, Pedersen T G, Skovhus T, Svendsen M K, Mortensen J J, Olsen T and Thygesen K S 2021 Recent progress of the computational 2d materials database (c2db) (*Preprint* 2102.03029)

[28] Gjerding M, Skovhus T, Rasmussen A, Bertoldo F, Larsen A H, Mortensen J J and Thygesen K S 2021 *Computational Materials Science* **199** 110731

[29] Larsen A H, Mortensen J J, Blomqvist J, Castelli I E, Christensen R, Dułak M, Friis J, Groves M N, Hammer B, Hargus C *et al.* 2017 *Journal of Physics: Condensed Matter* **29** 273002

[30] Togo A and Tanaka I 2015 *Scripta Materialia* **108** 1 – 5 ISSN 1359-6462

[31] Enkovaara J, Rostgaard C, Mortensen J J, Chen J, Dułak M, Ferrighi L, Gavnholt J, Glinsvad C, Haikola V, Hansen H A, Kristoffersen H H, Kuisma M, Larsen A H, Lehtovaara L, Ljungberg M, Lopez-Acevedo O, Moses P G, Ojanen J, Olsen T, Petzold V, Romero N A, Stausholm-Møller J, Strange M, Tritsaris G A, Vanin M, Walter M, Hammer B, Häkkinen H, Madsen G K H, Nieminen R M, Nørskov J K, Puska M, Rantala T T, Schiøtz J, Thygesen K S and Jacobsen K W 2010 *Journal of Physics: Condensed Matter* **22** 253202 URL

[32] Perdew J P, Burke K and Ernzerhof M 1996 *Phys. Rev. Lett.* **77**(18) 3865–3868

[33] https://cmr.fysik.dtu.dk/c2db/c2db.html

[34] Eriksson F, Fransson E and Erhart P *Advanced Theory and Simulations* **2** 1800184

[35] Rossnagel K 2011 *Journal of Physics: Condensed Matter* **23** 213001 URL https://doi.org/10.1088/0953-8984/23/21/213001

[36] Mak K F, Lee C, Hone J, Shan J and Heinz T F 2010 *Physical Review Letters* **105** 136805

[37] Manzeli S, Ovchinnikov D, Pasquier D, Yazyev O V and Kis A 2017 *Nature Reviews Materials* **2** 1–15

[38] Liu H, Neal A T, Zhu Z, Luo Z, Xu X, Tománek D and Ye P D 2014 *ACS Nano* **8** 4033–4041

Paper IV

Quantum point defects in 2D materials: The QPOD database

Fabian Bertoldo, Sajid Ali, <u>Simone Manti</u>, Kristian S. Thygesen
  Submitted

Quantum point defects in 2D materials: The QPOD database

Fabian Bertoldo, Sajid Ali, Simone Manti, Kristian S. Thygesen

[1] CAMD, Computational Atomic-Scale Materials Design, Department of Physics, Technical University of Denmark, 2800 Kgs. Lyngby Denmark

E-mail: `fafb@dtu.dk`

Abstract. Atomically thin two-dimensional (2D) materials are ideal hosts of quantum defects as they offer easier creation, manipulation, and read-out of defect states compared to bulk systems. Here we introduce the Quantum Point Defect (QPOD) database of 1500 intrinsic point defects (vacancies and antisites) in 70 2D insulators. The Atomic Simulation Recipes (ASR) workflow framework was used to perform density functional theory (DFT) calculations of defect formation energies, charge transition levels, Fermi level positions, equilibrium defect and carrier concentrations, transition dipole moments, hyperfine coupling, and zero-field splitting. Excited state calculations of photo luminescence spectra were performed for selected high-spin defects. In this paper we describe the calculations and workflow behind the database, present an overview of its content, and discuss some general trends and correlations in the data. We analyse the degree of defect tolerance of the host materials and identify promising defects for quantum technology applications. The database is freely available and can be browsed via a web-app interlinked with the Computational 2D Materials Database (C2DB).

Keywords: point defects, 2D materials, high-throughput

1. Introduction

Point defects are ubiquitous entities affecting the properties of any crystalline material. Under equilibrium conditions their concentration is given by the Boltzmann distribution, but strong deviations can occur in synthesised samples due to non-equilibrium growth conditions and significant energy barriers involved in the formation, transformation, or annihilation of defects. In many applications of semiconductor materials, in particular those relying on efficient carrier transport, the presence of defects has a detrimental impact on performance[1]. However, point defects in crystals can also be useful and form the basis for novel applications e.g. in spintronics [2], quantum computing[3], or quantum photonics[4, 5, 4, 5, 6, 7]. For such applications, defects may be introduced in a controlled manner e.g. by electron/ion beam irradiation, implantation, plasma treatment or high-temperature annealing in the presence of different gasses.

Over the past decade, atomically thin two-dimensional (2D) crystals have emerged as a promising class of materials with many attractive features including unique, easily tunable, and often superior physical properties.[8] This holds in particular for their defect-based properties and related applications. Compared to point defects buried deep inside a bulk structure, defects in 2D materials are inherently surface-near making them easier to create, manipulate, and characterise[9]. Recently, single photon emission (SPE) has been observed from point defects in 2D materials such as hexagonal boron-nitride (hBN)[10, 11, 12], MoS_2[13], and WSe_2[14, 15], and in a few cases optically detected magnetic resonance (ODMR) has been demonstrated[16, 17]. In the realm of catalysis, defects can act as active sites on otherwise chemically inert 2D materials [18, 19].

First-principles calculations based on density functional theory (DFT) can provide detailed insight into the physics and chemistry of point defects and how they influence materials properties at the atomic and electronic scales.[20, 21, 22, 23] In particular, such calculations have become an indispensable tool for interpreting experiments on defects, e.g. (magneto)optical experiments, which in themselves only provide indirect information about the microscopic nature of the involved defect(s).[24, 12] In combination with recently developed tools

for high-throughput workflow management[25, 26, 27, 28], first-principles calculations have potential to play a more proactive role in the search for new defect systems with promising properties. Here, a major challenge is the notorious complexity of defect calculations (even when performed in low-throughput mode) that involves large supercells, local magnetic moments, electrostatic corrections, etc. Performing such calculations for general defects in general host materials, requires a carefully designed workflow with optimised computational settings and a substantial amount of benchmarking.

In this work, we present a systematic study of 500 intrinsic point defects (vacancies and antisite defects) of 82 insulating 2D host materials. The host materials were selected from the Computational 2D Materials Database (C2DB)[29, 30] after applying a series of filtering criteria. Our computational workflow incorporates the calculation and analysis of thermodynamic properties such as defect formation energies, charge transition levels, equilibrium carrier concentrations and Fermi level position, as well as symmetry analysis of the defect atomic structures and wave functions, magnetic properties such as hyperfine coupling parameters and zero-field splittings, and optical transition dipoles. Defects with a high-spin ground state are particularly interesting for magneto-optical and quantum information technology applications. For such defects the excited state properties, including vertical excitation energies, reorganisation energy, radiative lifetime, and photoluminescence (PL) lineshapes, were also calculated.

The computational defect workflow was constructed with the Atomic Simulation Recipes (ASR)[27] and executed using the MyQueue [28] task scheduler frontend. The ASR provides a simple and modular framework for developing Python workflow scripts, and its automatic caching system keeps track of job status and logs data provenance. Our ASR defect workflow adds to other ongoing efforts to automate the computational characterisation of point defects.[31, 32, 33, 34] However, to the best of our knowledge, the present work represents the first actual high-throughput study of point defects. All of the generated data is collected in the Quantum Point Defect (QPOD) database and will be publicly available and accessible via a browsable web-service. The QPOD webpages are interlinked with the C2DB providing a seamless interface between the properties of the pristine host materials and their

intrinsic point defects.

The theoretical framework is based entirely on DFT with the Perdew Burke Ernzerhof (PBE) functional[35]. Charge transition levels are obtained using Slater-Janak transition state theory[36] while excited states are calculated by DO-MOM method[37]. We note that PBE suffers from delocalisation errors[38], which may introduce quantitative inaccuracies in the description of some localised defect states. While range-separated hybrid functionals represent the state-of-the art methodology for point defect calculations, such a description is currently too demanding for large-scale studies like the current. Moreover, thermodynamic properties of defects are generally well described by PBE[39].

In Section 2 we describe the theory and methodology employed at the various computational steps of the workflow. Section 3 gives a general overview of the workflow, introduces the set of host materials and the considered point defects, and outlines the structure and content of the QPOD database web-interface. In Section 4 we present our main results. These include statistical overviews of host crystal and defect system properties (formation trends, symmetries, *etc.*), an analysis of the effect of structural relaxations on defect formation energies and charge transition levels, an evaluation of the intrinsic (equilibrium) doping level in 59 host materials and identification a small subset of the host materials that are particularly defect tolerant. We also identify a few defect systems with promising properties for spin qubit applications or magneto-optical sensing. Section 5 summarises the work and looks ahead.

2. Theory and Methodology

2.1. Density functional theory calculations

All the DFT calculations (spin-polarized) are performed by the GPAW electronic structure code [40] using plane wave basis set with 800 eV plane wave cut off, k-point density of 6 Å^{-1} (12 Å^{-1}) for structural relaxations (for ground state calculations) and the PBE xc-functional [35]. The supercell is kept fixed and atoms are fully relaxed until forces are below 0.01 eV/Å. We apply a Fermi smearing of 0.02 eV (0.2 eV for relaxations) for all systems and increase that parameter slightly (0.05, 0.08 or 0.1) for systems whose ground state proves difficult to converge. We use the Pulay mixing scheme [41] where total density and magnetisation densities are treated separately. The excited states are calculated using the same computational parameters as the ground state, and using the DO-MOM method[37], where the maximum step length, $\rho_{\max}$, for the quasi-newton search direction is chosen to be 0.20. The parameters for the Slater-Janak calculations
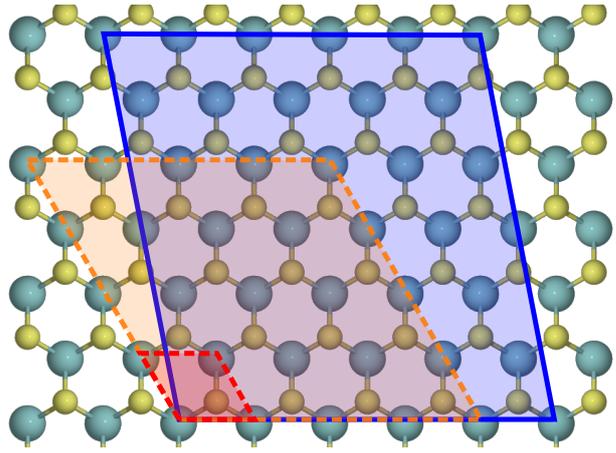
Figure 1. Creation of defect supercells at the example of MoS$_2$. Top view of a primitive unit cell of MoS$_2$ (red, dashed), an example of a conventional $4 \times 4 \times 1$ supercell (orange, dashed), as well as our different approach of a symmetry broken supercell (blue, solid).

are the same as for the ground state calculations.

2.2. Symmetry-broken supercell

A supercell in 2D can be created using linear combinations of the primitive unit cell vectors ($\mathbf{a_1}$, $\mathbf{a_2}$). The corresponding supercell lattice vectors ($\mathbf{b_1}$, $\mathbf{b_2}$) are written as:

$$\mathbf{b_1} = n_1\mathbf{a_1} + n_2\mathbf{a_2}, \tag{1}$$

$$\mathbf{b_2} = m_1\mathbf{a_1} + m_2\mathbf{a_2}, \tag{2}$$

where n_1, n_2, m_1, m_2 are integers larger or equal to zero.

In this study, we apply an algorithm that finds the most suitable parameters n_1, n_2, m_1, m_2: (i) set $n_2 = 0$ and evaluate all possible linear combinations of n_1, m_1, m_2, (ii) discard combinations where $m_1 = 0$ and $n_1 = m_2$, (iii) only keep cells where the minimum distance between mirrored defects is larger than 15 Å, (iv) keep supercells consisting of the minimum number of atoms, (v) choose the most isotropic supercell. We note that step (ii) is conducted in order to break the symmetry of the initial Bravais lattice, and steps and (iv) minimizes computational cost for a viable high-throughput execution.

Defects are introduced by analysing Wyckoff positions of the atoms within the primitive structure. For each non-equivalent position in the structure a vacancy defect and substitutional defects are created (the latter by replacing a specific atom with another atom of a different species intrinsic to the host material). For the example of MoS$_2$ the procedure yields the following point defects: sulfur vacancy V$_S$, molybdenum vacancy V$_{Mo}$, as well as two substitutional defects Mo$_S$ and S$_{Mo}$ where Mo replaces

the S atom and *vice versa*. Each defect supercell created by this approach undergoes the workflow which is presented in chapter 3.

2.3. Defect formation energy

The formation energy of a defect X in charge state q is defined by[42, 43]:

$$E^f[X^q] = E_{\text{tot}}[X^q] - E_{\text{tot}}[\text{bulk}] - \sum_i n_i \mu_i + qE_{\text{F}} \quad (3)$$

where μ_i is the chemical potential of the atom species i and n_i is the number of such atoms that have been added ($n_i > 0$) or removed ($n_i < 0$) in order to create the defect. Here, in light of the high-throughput approach used, we choose the standard states of the respective elemental phases of i calculated with GPAW [40] to obtain the values for μ_i. We note, that μ_i does not have to be a fixed value but can also be represented by a range of chemical potentials in order to specify *i-rich* and *i-poor* conditions as it is well documented in existing literature [44]. For finite charge states, the defect formation energy becomes a function of the Fermi energy, E_{F}, which represents the chemical potential of electrons. In equilibrium, the concentration of a specific defect type is determined by its formation energy, which in turn depends on E_{F}. Imposing global charge neutrality leads to a self-consistency problem for E_{F}, which we discuss in Sec. 2.5.

In general, the lower the formation energy of a particular defect, the higher the probability for it to be present in the material. In equilibrium, the defect concentration is given by the Boltzmann distribution,

$$C^{\text{eq}}[X^q] = N_X g_{X^q} \exp\left[-E^f[X^q]/(k_B T)\right], \quad (4)$$

where N_X and g_{X^q} specifies the site and defect state degeneracy, respectively, k_B is Boltzmann's constant and T is the temperature.

As an example, Fig. 2 shows the formation energy (blue solid lines) of a sulfur vacancy V_S in MoS_2 as a function of the Fermi level position. It follows that this particular defect is most stable in its neutral charge state ($q = 0$) for low to mid gap Fermi level positions. The transition from $q = 0$ to $q' = -1$ occurs for E_{F} around 1.2 eV (blue dotted line). The formation energy of the neutral V_S and its CTLs reported in Ref. [45] (orange lines) are in good agreement with our values. The differences are below 0.2 eV and can be ascribed to the difference in the employed supercells (symmetric vs. symmetry-broken) and xc-functionals (PBE-D vs. PBE). In this work, the CTLs are obtained using the Slater-Janak transition state theory (see next section) while Ref. [45] used total energy differences with electrostatic corrections.
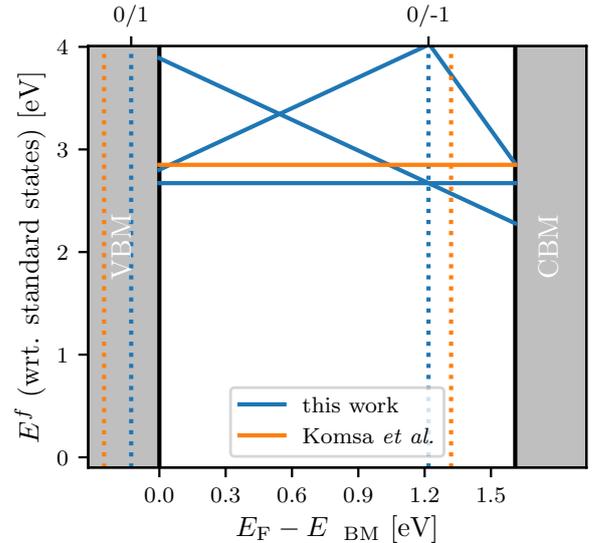

Figure 2. Formation energy of V_S in MoS_2 referenced to the standard states as a function of Fermi energy. Calculated formation energy (blue, solid) as a function of the Fermi energy plotted together with the CTL (blue, dotted). The orange solid line (orange dotted lines) highlight the PBE-D calculated neutral formation energy (CTL of (0/-1) and (0/1)) of V_S in MoS_2 taken from Komsa *et al.* [45].

2.4. Slater-Janak transition state theory

The prediction of charge transition levels (CTLs) requires the total energy of the defect in different charge state. In the standard approach, the extra electrons/holes are included in the self-consistent DFT calculation and a background charge distribution is added to make the supercell overall charge neutral. In a post process step, the spurious interactions between periodically repeated images is removed from the total energy using an electrostatic correction scheme that involves a Gaussian approximation to the localised charge distribution and a model for the dielectric function of the material[20]. While this approach is fairly straightforward and unambiguous for bulk materials, it becomes significantly more challenging for 2D materials due to the spatial confinement and non-local nature of the dielectric function and the dependence on detailed shape of the neutralising background charge[46, 47, 48, 49].

To avoid the difficulties associated with electrostatic corrections, we rely on the Slater-Janak (SJ) theorem[36], which relates the Kohn-Sham eigenvalue ε_i to the derivative of the total energy with respect to the orbital occupation number n_i,

$$\frac{\partial E}{\partial n_i} = \varepsilon_i(n_i). \quad (5)$$

The theorem can be used to express the difference in ground state energy between to charge states as an
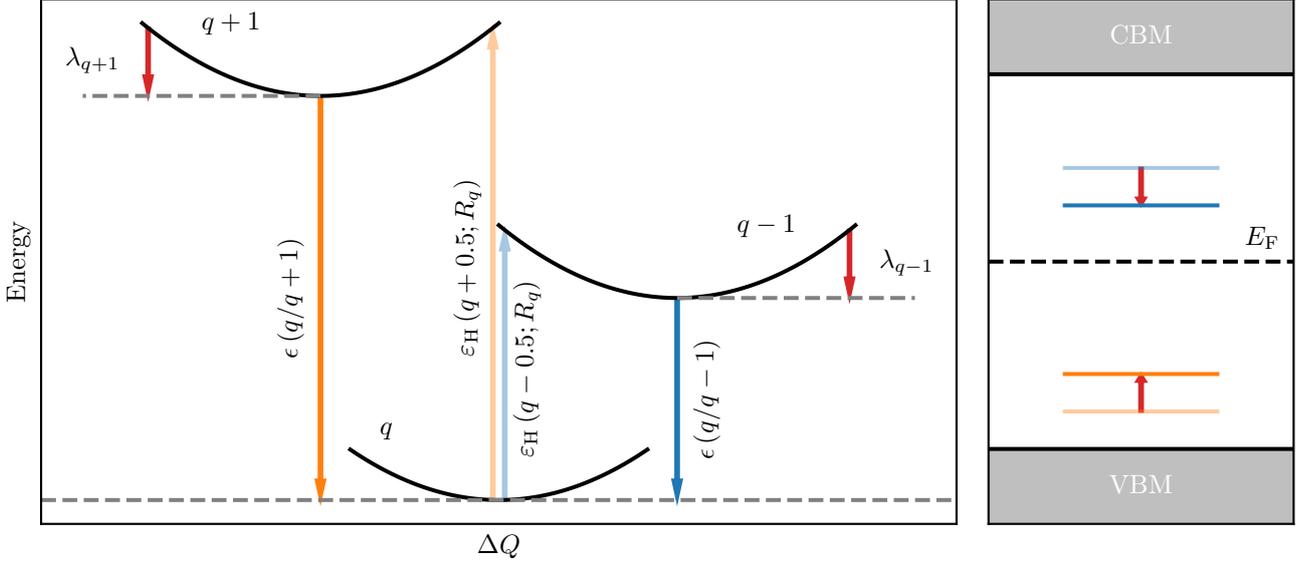
Figure 3. Schematic view of the SJ approach. Optical charge transition levels are calculated starting from a given charge state q by either adding (*i.e.* $q-1$) or removing (*i.e.* $q+1$) an electron and calculating the HOMO energies at frozen atomic configuration ($\varepsilon_{\mathrm{H}}\left(q-1; R_q\right)$, and $\varepsilon_{\mathrm{H}}\left(q+1; R_q\right)$, left panel). Due to ionic relaxations one has to include a relaxation correction that either needs to be added or subtracted from the optical transition level (red arrows, right panel) in order to compute the thermodynamic CTLs ($\epsilon\left(q/q-1\right)$ and $\epsilon\left(q/q+1\right)$).

integral over the eigenvalue as its occupation number is changed from 0 to 1. This approached, termed SJ transition state theory, has been used successfully used to evaluate core-level shifts in random alloys[50], and CTLs of impurities in GaN[51], native defects in LiNbO$_3$[52], and and chalcogen vacancies in monolayer TMDs [53]. Assuming a linear dependence of ε_i on n_i (which holds exactly for the true Kohn-Sham system), the transition energy between two localised states of charge q and $q' = q \pm 1$ can be written

$$\epsilon\left(q/q'\right) = \begin{cases} \varepsilon_{\mathrm{H}}\left(q + \frac{1}{2}; R_q\right) - \lambda_{q'}, & q' = q + 1 \\ \\ \varepsilon_{\mathrm{H}}\left(q - \frac{1}{2}; R_q\right) + \lambda_{q'}, & q' = q - 1 \end{cases} \quad (6)$$

Here ε_{H} represents highest eigenvalue with non-zero occupation, i.e. the half occupied state, and R_q refers to the configuration of charge state q. The reorganisation energy is obtained as a total energy difference between equal charge states

$$\lambda_{q'} = E_{\mathrm{tot}}\left(q'; R_{q'}\right) - E_{\mathrm{tot}}\left(q'; R_q\right). \quad (7)$$

Note that the reorganisation energy is always negative. The relevant quantities are illustrated graphically in Figure 3.

It has been shown for defects in bulk materials that the CTLs obtained from SJ theory are in good agreement with results obtained from total energy differences[50, 52]. For 2D materials, a major advantage of the SJ method is that it circumvents the issues related to the electrostatic correction, because it

completely avoids the comparison of energies between supercells with different number of electrons. The Kohn-Sham eigenvalues of a neutral or (partially) charged defect supercell are referenced relative to the electrostatic potential averaged over the PAW sphere of an atom located as far as possible away from the defect site (typically around 7 Ådepending on the exact size of the supercell). By performing the potential average over an equivalent atom of the pristine 2D layer, we can reference the Kohn-Sham eigenvalue of the defect supercell to the VBM of the pristine material. As an alternative to averaging the potential around an atom, the asymptotic vacuum potential can be used as reference. We have checked that the two procedures yield identical results (usually within 0.1 eV), but prefer the atom-averaging scheme as it can be applied to bulk materials as well.

2.5. Equilibrium defect concentrations

According to Eqs. (3) and (4), the formation energy of charged defects, and therefore their equilibrium concentration, is a function of the Fermi level. The Fermi level position of the system in thermal equilibrium is then determined self-consistently from a requirement of charge neutrality[54]

$$\sum_X \sum_q qC[X^q] = n_0 - p_0 \quad (8)$$

where the sum is over structural defects X in charge states q, and n_0 and p_0 are the electron and hole carrier
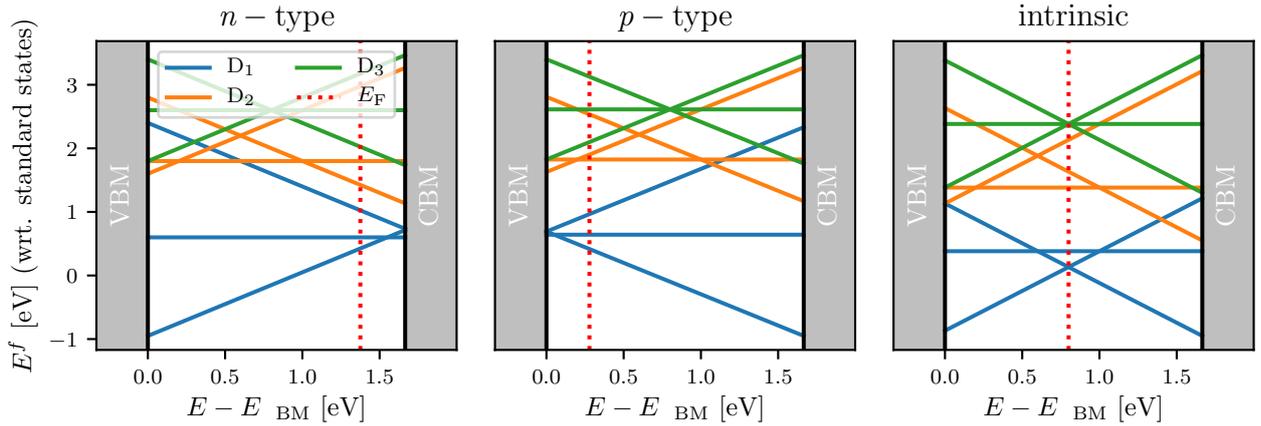
Figure 4. Intrinsic dopability of defect systems. Formation energies with respect to the standard states as a function of energy for three mock-sytems with three defect types present, respectively. Left: n-type dopable regime (E_F close to CBM) with dominating donor defect. Middle: p-type dopable regime (E_F close to VBM) with dominating acceptor defect. Right: intrinsic material (neither p- nor n-dopable due to the presence of competing acceptor and donor defects.)

concentrations, respectively. The latter are given by

$$n_0 = \int_{E_{\mathrm{gap}}}^{\infty} dE f(E) \rho(E), \qquad (9)$$

$$p_0 = \int_{-\infty}^{0} dE [1 - f(E)] \rho(E), \qquad (10)$$

where $\rho(E)$ is the local density of states, $f(E) = 1/\exp\left[(E - E_F)/k_B T\right]$ is the Fermi-Dirac distribution, and the energy scale is referenced to the valence band maximum.

Under the assumption that all the relevant defects, i.e. the intrinsic defects with the lowest formation energies, are accounted for, Eq. (8) will determine the Fermi level position of the material in thermal equilibrium. The equilibrium Fermi level position determines whether a material is intrinsically p-doped, n-doped, or intrinsic. The three different cases are illustrated schematically in Fig. 4. For the n-type case (left panel), the most stable defect is D_1 in charge state $+1$. Thus, charge carriers are transferred from the defect into the conduction band resulting in a Fermi level just below the CBM. Similarly, for the p-type case (middle panel) the defect D_1 in charge state -1 is the most stable. Consequently, charge carriers are promoted from the valence band into the defect resulting in E_F close to the VBM. In the right panel, donor and acceptor states are competing, which results in an effective cancelling of the p- and n-type behavior, pinning the Fermi level in the middle of the band gap. In Section 4.3 we analyse the intrinsic carrier types and concentrations of 59 host materials.

2.6. Symmetry analysis

All states with an energy in the band gap are classified according to the symmetry group of the defect using a generalization of the methodology previously implemented in GPAW for molecules [55]. In a first step, the point group, G, of the defect is determined. To determine G we first reintroduce the relaxed defect into a supercell that preserves the symmetry of the host material; precisely, a supercell with basis vectors defined by setting $n_1 = m_2 = n$ and $n_2 = m_1 = 0$ in Eqs. (1,2). We then use spglib[56] to obtain G as the point group of the new supercell. We stress that the high-symmetry supercell is only used to determine G while all actual calculations are performed for the low-symmetry supercell as described in Sec. 2.6.

The defect states in the band gap are labeled according to the irreducible representations (irrep) of G. To obtain the irrep of a given eigenstate, $\psi_n(\mathbf{r})$, we form the matrix elements

$$\Gamma(\mathrm{R}) = \Gamma_{nn}(\mathrm{R}) = \int d\mathbf{r}\, \psi_n(\mathbf{r})^* \mathrm{R}\, \psi_n(\mathbf{r}), \qquad (11)$$

where R is any symmetry transformation of G. It follows from the orthogonality theorem[57] that the vector $\Gamma(\mathrm{R})$ can be expanded in the character vectors, $\chi^{(\alpha)}(\mathrm{R})$,

$$\Gamma(\mathrm{R}) = \sum_\alpha c_\alpha \chi^{(\alpha)}(\mathrm{R}) \qquad (12)$$

where the quantity c_α represents the fraction of ψ_n that transforms according to the irrep α. For any well localised defect state, all the c_α will be zero except one, which is the irrep of the state. In general, less localised states will not transform according to an irrep of G. That is because the wave functions are calculated in the

symmetry-broken supercell, see Sec. 2.2, and therefore will have a lower symmetry than the defect. To exclude such low-symmetry tails on the wave functions, the integral in Eq. (11) is truncated beyond a cut-off radius measured from the center of symmetry of the defect.

As an example, Figure 5 shows the coefficients c_α for the in-gap states of the (neutral) sulfur vacancy in MoS_2. This is a well studied prototypical defect with C_{3v} symmetry. The absence of the chalcogen atom introduces three defect states in the gap: a totally symmetric a_1 state close to the VBM and two doubly degenerate mid-gap states (e_x, e_y). The symmetry coefficients c_α correctly captures the expected symmetry of the states[45], at least for small cut-off radii. The effect of employing a symmetry-broken supercell can be seen on the totally symmetric state a_1, which starts to be mixed with the antisymmetric a_2 irrep as a function of the radius, while there is no effect on the degenerate e_x state. Therefore a radius of 2 Å is used in the database to catch the expected local symmetry of the defect.

2.7. Transition dipole moment

The transition dipole moment is calculated between all single particle Kohn-Sham states inside the band gap,

$$\mu_{nm} = \langle \psi_n | \, \hat{\mathbf{r}} \, | \psi_m \rangle = \frac{i\hbar}{m} \frac{\langle \psi_n | \, \hat{\mathbf{p}} \, | \psi_m \rangle}{\varepsilon_n - \varepsilon_m}, \tag{13}$$

where $\hat{\mathbf{r}}$ is the dipole operator, $\hat{\mathbf{p}}$ is the momentum operator, and m is the electron mass. The transition dipole moment yields information on the possible polarization directions and oscillator strength of a given transition. In this work, the transition dipoles between localised defect states are calculated in real space, i.e. using the first expression in Eq. (13), after translating the defect to the center of the supercell.

2.8. Radiative recombination rate and life time

Radiative recombination refers to the spontaneous decay of an electron from an initial high energy state to a state of lower energy upon emitting a photon. The rate of a spin-preserving radiative transition between an initial state ψ_m and a final state ψ_n is given by [58]

$$\Gamma_{nm}^{\mathrm{rad}} = 1/\tau_{\mathrm{rad}} = \frac{E_{\mathrm{ZPL}}^3 \mu_{nm}^2}{3\pi\epsilon_0 c^3 \hbar^4} \tag{14}$$

here ϵ_0 is the vacuum permitivity, E_{ZPL} is the zero phonon line (ZPL) energy of the transition (see Sec. 2.11), and μ_{nm} is the transition dipole moment defined in Eq. (13). The ZPL energy includes the reorganisation energy due to structural differences between the initial and final states, which can be on the order of 1 eV. Consequently, an accurate estimate of the radiative lifetime requires a geometry optimisation in the excited state. Since this step is not part of our

general workflow, but is only performed for selected defects, radiative lifetimes are currently only available for a few transitions.

2.9. Hyperfine coupling

Hyperfine (HF) coupling refers to the interaction between the magnetic dipole associated with a nuclear spin, $\hat{\mathbf{I}}^N$, and the magnetic dipole of the electron-spin distribution, $\hat{\mathbf{S}}$. For a fixed atomic nuclei, N, the interaction is written

$$H_{\mathrm{HF}} = \sum_{i,j} \hat{S}_i A_{ij}^N \hat{I}_j^N \tag{15}$$

where the hyperfine tensor A^N is given by

$$\begin{aligned}
A_{ij}^N = &\frac{2\alpha^2 g_e m_e}{3m_p} \int \delta_T(\mathbf{r}) \rho_s(\mathbf{r}) d\mathbf{r} \\
&+ \frac{\alpha^2 g_e m_e}{4\pi m_p} \int \frac{3r_i r_j - \delta_{ij} r^2}{r^5} \rho_s(\mathbf{r}) d\mathbf{r},
\end{aligned} \tag{16}$$

The first time represents the isotropic part also called Fermi-contact term and results due to non-vanishing spin density $\rho_s(\mathbf{r})$ of electron at the centre of the nucleous. Where $\delta_T(\mathbf{r})$ is a smeared out δ function, α is fine structure constant, m_e and m_p are electronic and nuclear masses respectively and g_e is the gyromagnetic ratio for electrons. The second terms represents the aniotropic part of the hyperfine coupling tensor and results due to dipole-dipole interaction between nuclear and electronic magnetic moments. The isotropic component or Fermi-contact term $a = \mathrm{Tr}(\mathbf{A_{i,j}})/3$ is is sensitive only to the density of s electrons at the nucleus. The anisotropic term $\mathbf{A_{i,j}} - a$ is sensitive to the $l = 2$ component of the density near the nucleus which is determined by the p contribution of the wave function in the absence of any d orbitals. Therefore the hyperfine parameter provides direct insight into the electron distribution or wave function near the corresponding nucleus, and thus its hybridization. A direct comparison of calculated HF coupling constants with parameters of Electron paramagnetic resonance spectroscopy can help to identify the nature of defect centre[59].

In the present work one $A_{i,J}$ tensor per atom is calculated. We then perform diagonalization and calculate the principle values of the HF coupling constants and report them in the database only for the atoms showing significant HF interaction. Figure 6 shows the iso-surface of spin density on the atoms surrounding a typical defect site.

2.10. Zero field splitting

Zero field splitting (ZFS) refers to the splitting of magnetic sub levels of a triplet defect state due to the magnetic dipole-dipole interaction of unpaired electron
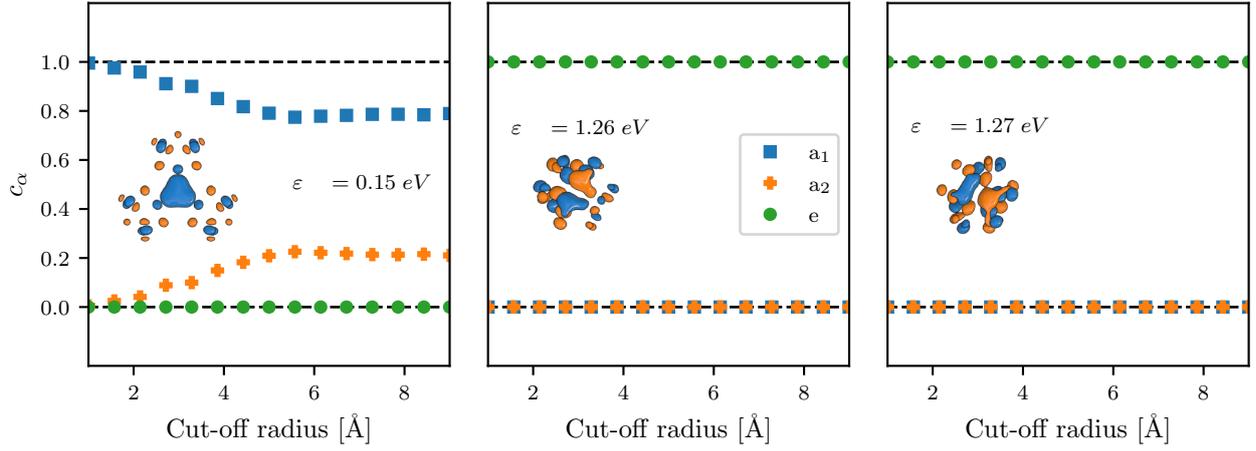
Figure 5. Orbital symmetry labels for different cutoff radii. All defect states with an energy inside the band gap of the host material are classified according to the irreps of the point group of the defect. The c_α coefficient is a measure of the degree to which a defect state transforms according to irrep α. Performing the symmetry analysis within a radius of a few Å of the center of symmetry leads to a correct classification of the totally symmetric a_1 state and the degenerate e_x and e_y states of the neutral sulfur vacancy in MoS_2. The isosurfaces of the orbitals are shown as insets and the energy eigenvalues reference to the VBM.
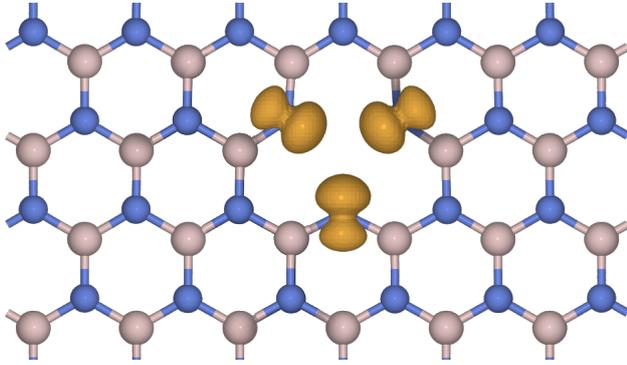

Figure 6. The Iso-surface of the calculated spin-density (isovalue of 0.002 $|e|/\overset{\circ}{A}^3$) for V_B^{-1} defect in hexagonal boron nitride.
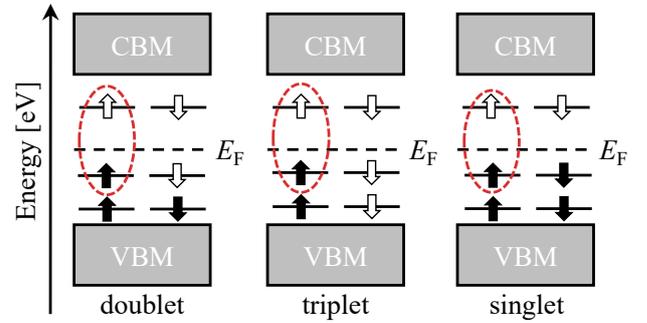

Figure 7. Possible occupancies resulting in different spin configurations of the defect systems for the excited state calculations. The states involved in the excitation are encircled red.

spins that takes place even in the absence of an external magnetic field. A triplet ($s = 1$) defect state can be described by the spin Hamiltonian of the following form

$$H_{\text{ZFS}} = \sum_{ij} S_i D_{ij} S_j \tag{17}$$

where $\mathbf{S}$ is the total spin and D is the ZFS tensor. This equation can also be written in the following form

$$\begin{aligned} H_{ZFS} &= D_{XX} S_x^2 + D_{YY} S_y^2 + D_{ZZ} S_z^2 \\ &= D(S_z^2 - S(S+1)/3) + E(S_x^2 - S_y^2) \end{aligned} \tag{18}$$

where, $D = 3/2 D_{zz}$ and $E = (D_{XX} - D_{YY})/2$ are called the axial and rhombic ZFS parameters, respectively. For high-spin defects i.e. triplets, D describes the splitting between the $m_s = 1$ and $m_s = 0$ magnetic sub-levels, while E describes the splitting of $m_s = 1$ sub-levels. D is generally zero for a spherically

symmetric wave function because there is no direction in which electrons of a triplet can move to minimize the repulsive dipole-dipole interaction. However, for a non-spherical distribution of molecular orbitals, D will be non-zero, due to lifting of the degeneracy of magnetic sub-levels $m_s = 1$ and $m_s = 0$. A positive value of D would symbolize an oblate spin-distribution, while a negative value would mean a prolate spin-distribution. The value of E will be zero for axially symmetric wave functions.

2.11. Excited states

In Kohn-Sham density functional theory (DFT), excited electronic states can be found by solving the Kohn-Sham equations for orbitals with non-Aufbau occupations.

The excited-state solutions are saddle points of

115

the Kohn-Sham energy functional and common self-consistent field (SCF) approaches often struggle to find such solutions, especially when nearly degenerate states are involved. SCF approaches fail specially for cases involving charge transfer or Rhydberg states. This is due to a significant re-arrangement of closely spaced orbitals upon excitation. Therefore, in the present work we will use an alternative to SCF method i.e. a method based on the direct optimization (DO) of orbital rotations, representing the electronic degrees of freedom, by means of efficient quasi-Newton methods for saddle points in combination with the maximum overlap method (MOM)[37]. Convergence on nth order saddle point is guided by an appropriate pre-conditioner based on an approximation to the Hessian, where the number of negative eigenvalues is consistent with the target nth-order saddle point. This method ensures the fast and robust convergence of the excited states, as compared to SCF [37]. DO-MOM method have been previously used for the calculation of excited state spectrum of molecules [37], however, it has never been applied to defect states. So, one novelty of the present work is to apply this method for the first time to defect states in semiconductors. We have first performed a broad benchmarking of the previously well known defect systems and established that this method generate accurate results. In the present work we study only one adiabatic excitation in each spin channel for defect systems for which the PBE kohn-sham gap between HOMO and LUMO is at least 0.5 eV. Possible spin configurations for different defect systems are skecthed in 7. For the doublet and triplet spin configurations, both ground and excited states can be expressed as single Slater-determinents. However, for the singlet spin configuration, the ground state is a closed shell singlet, while the excited state, as a result of single excitation in either spin channel, would result in open shell singlet state, which cannot be expressed as single Slater- determinent. This open shell singlet state is a mixed singlet-triplet state and can be expressed as sum of two Slater-determinents i.e.

$$E_{st} = \frac{1}{2}(E_s + E_t) \Rightarrow E_s = 2E_{st} - E_t \qquad (19)$$

Where E_{st} is the DFT energy calculated by setting the occupancy for open-shell singlet state, while E_s and E_t are the energies of the corresponding singlet and triplet states.

The line shape of the emission from the electronic excited state to the vibrationally excited electronic ground state can be computed with DFT using generating function approach[60]. First the mass weighted difference between atomic co-ordinates of the ground and excited states is computed as follows

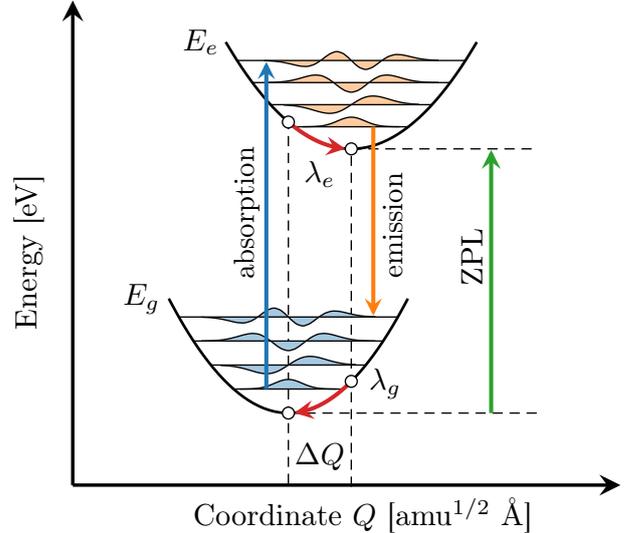$$\Delta Q = \sqrt{\sum_\alpha m_\alpha \Delta R_\alpha^2} \qquad (20)$$

116

Figure 8. Schematic CC diagram of a ground state to excited state transition. Absorption and emission to and from the excited state (blue, orange) as well as their respective reorganization energies $\lambda_{e,g}$ as a function of energy and configuration coordinate Q. The zero phonon line transition (ZPL) is visualized in green.

where sum runs over all the atoms in the supercell. The partial Huang-Rhys factors are then computed by

$$S_k = \frac{1}{2\hbar}\omega_k Q_k^2 \qquad (21)$$

where Q_k is the projection of the lattice displacement on the normal co-ordinates of the ground state described by phonon mode k. Then electron-phonon spectral function, which depends on the coupling between lattice displacement and vibrational degrees of freedom is then given by summation over all the modes k

$$S(\omega) = \sum_k S_k\delta(\omega - \omega_k) \qquad (22)$$

The integral over the electron-phonon spectral function gives the (total) HR factor of the transition. The above electron-phonon spectral function fed into a generating function [60] generates the PL line shape.

3. QPOD workflow and database

3.1. The QPOD workflow

The backbone of the QPOD database is represented by a high-throughput framework based on the Atomic Simulation Recipes (ASR) [27] in connection with the MyQueue [28] task and workflow scheduling system. Numerous recipes, designed particularly for the evaluation of defect properties, have been implemented in ASR and have been combined in a
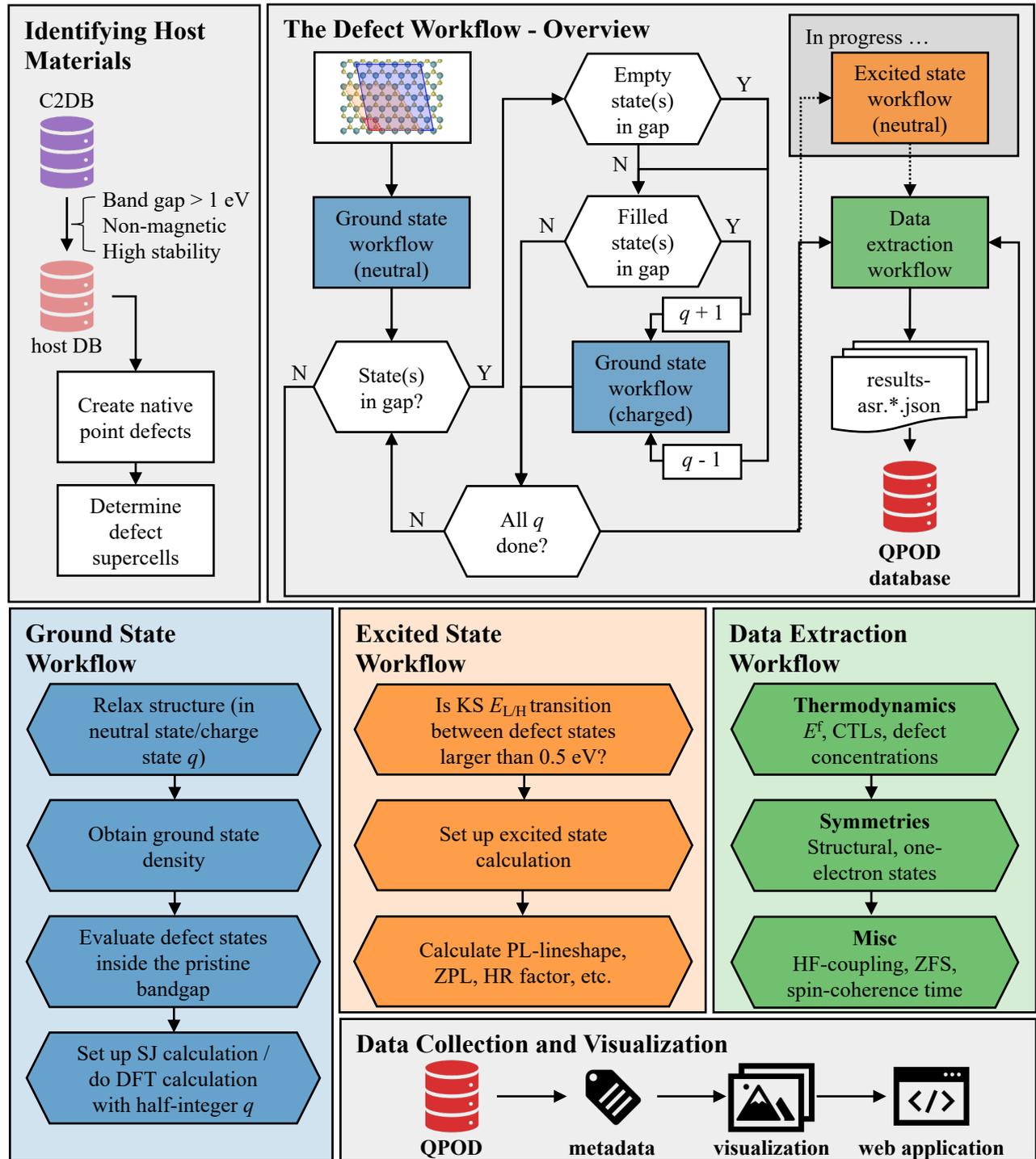
Figure 9. The workflow behind the defect database. First, starting from C2DB suitable host materials are identified. With the ASR recipe for defect generation the defect supercells are set up and enter the ground state workflow for neutral systems. Afterwards, depending on the nature of the defect states inside the gap, charged calculations are conducted within the charged ground state workflow. If all charged calculations for a specific system are done ($|q| > 3$), the excited state workflow is executed, before physical meaningful results are collected using the data extraction workflow and thereafter saved in the defect database. The database is equipped with all necessary metadata, and together with various visualization scripts the browseable web application is created.

central MyQueue workflow to generate all data for the QPOD database.

The underlying workflow is sketched in Fig. 9, and will be described briefly in the following. As a preliminary step the C2DB [29, 30] is screened to obtain the set of host materials. Only non-magnetic, thermodynamically and dynamically stable materials with a PBE band gap of $E_{\mathrm{gap}}^{\mathrm{PBE}} > 1$ eV are selected as host materials. These criteria result in 281 materials of which we select 82 from a criterion of $N_e \times V_{\mathrm{supercell}} < 0.9$ Å^3 combined with a few handpicked experimentally known and relevant 2D materials (MoS_2, hBN, WS_2, $MoSe_2$). It is important to mention, that some host materials exist in different phases (same chemical formula and stoichiometry, but different symmetry), *e.g.* 2H-MoS_2 and 1T-MoS_2. In these cases we only keep the most stable one, *i.e.* 2H-MoS_2 for the previous example.

For each host material, all inequivalent vacancies and antisite defects are created in a supercell as described in Section 2.2. Each defect enters the ground state workflow, which includes relaxation of the neutral defect structure, calculation of a well-converged ground state density, identification and extraction of defect states within the pristine band gap, and SJ calculations with half-integer charges q. If there are no states within the gap, the defect system directly undergoes the data extraction workflow and is stored in the QPOD database.

For systems with in-gap states above (below) E_{F}, an electron is added (removed) and the charged structures are relaxed, their ground state calculated and the states within the band gap are examined again up to a maximum charge of $+3/-3$. Once all charge states have been relaxed and their ground state density has been evaluated, the data extraction workflow is executed. Here, general defect information (defect name, defect charge, nearest defect-defect distance, *etc.*), charge transition levels and formation energies, the equilibrium self-consistent Fermi level, equilibrium defect concentration, symmetries of the defect states within the gap, hyperfine coupling, zero field splitting, transition dipole moments, *etc.* are calculated and the results are stored in the database. The data is publicly available and easy to browse in a web-application as will be described in chapter 3.3.

We note, that selected systems have been subject to excited state calculations in order to obtain ZPL energies, PL spectra, HR factors, *etc.* enabling the identification of promising defect candidates for optical applications as is discussed in chapter 4.7.

3.2. The QPOD database

The QPOD database uses the ASE DB format [61] which currently has five backends: `JSON`, `SQLite3`,

`PostgreSQL`, `MySQL`, `MariaDB`. An ASE DB enables simple querying of the data *via* the in-built `ase db` command line tool, a Python interface, or a webapp (see chapter 3.3). With those different possibilities to access and interact with the data, we aim to give users a large flexibility based on their respective technical background and preferences.

Each row of QPOD is uniquely defined by defect name, host name, and it's respective charge state. Furthermore, the fully relaxed structure as well as all of the data associated with the respective defect is attached in the form of key-value pairs or `JSON`-formatted raw data.

3.3. The QPOD webapp

A defining feature of the QPOD database is its easy accessibility through a web application (webapp). For each row of the database, one can browse a collection of web-panels designed to highlight the various computed properties of the specific defect. Specific elements of the web-panels feature clickable "?" icons with explanatory descriptions of the content to improve the accessibility of the data.

Directing between different entries of the database is either possible by using hyperlinks between related entries, or using the overview page of the database, where the user can search for materials, and order them based on different criteria. Furthermore, we ensure the direct connection to C2DB with hyperlinks between a defect material and its respective host material counterpart in C2DB in case users want to find more information about the defect-free systems.

3.4. Overview of host materials

In total, 82 host crystals have been selected to be the basis of our systematic study of intrinsic point defects. The set of host materials span across a large range of symmetries, stoichiometries, and band gaps (see Fig. 10). Among those, 9 have already been experimentally realized: As_2, BN, Bi_2I_6, C_2H_2, MoS_2, $MoSe_2$, Pd_2Se_4, WS_2, ZrS_2. We include both small band gap materials (*e.g.* Ni_2Se_4, ZrS_2) and wide band gap materials such as BN, $BaCl_2$ and $MgCl_2$ making our set of starting host crystals particularly heterogeneous.

4. Results

In this section we first present some general illustrations and analyses of the data in QPOD. We then leverage the data to address three specific scientific problems, namely the identification of: (i) Defect tolerant semiconductors with low concentrations of midgap states. (ii) Intrinsically p-type or n-type semicon-
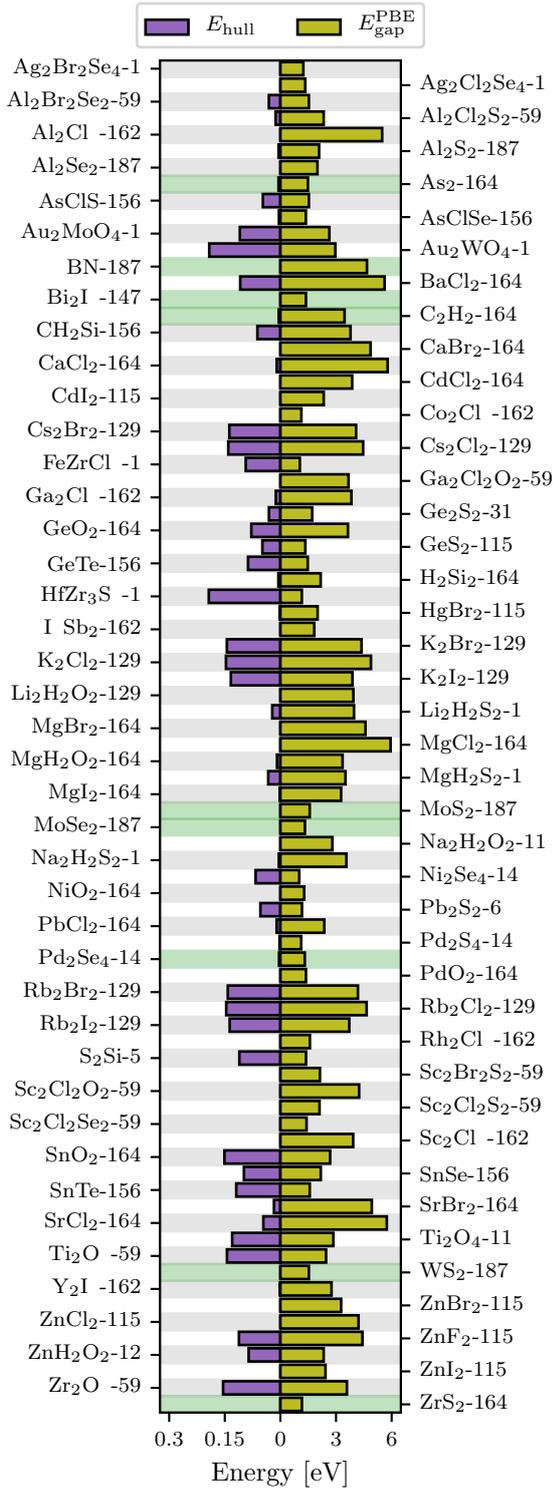
Figure 10. Overview of host crystals. Energy above convex hull in eV/atom E_{hull} (purple bars, left) and PBE-calculated band gap E_{gap}^{PBE} (yellow bars, right) of the 83 host crystals of QPOD. Monolayers which have been realized experimentally are highlighted in green.

ductors. (iii) Optically accessible high-spin defects for quantum technological applications.

4.1. Relaxation of defect structures

A major part of the computational efforts to create the QPOD went to the relaxation of the defect structures in a symmetry broken supercell. As discussed in Section 2.2 the strategy to actively break the symmetry of the host crystal by the choice of supercell was adopted to enable defects to relax into their lowest energy configuration.

Figure 11 shows the gain in total energy due to the relaxation for the over 1800 vacancy and antisite defects (different charge states included). Not unexpectedly, the relaxation has the largest influence on antisite defects while vacancy structures in general show very weak reorganization relative to the pristine structure as can be seen in the left panel of Fig. 11. The relaxations for charged defects have always been started from the neutral equilibrium configuration of the respective defect. As a result, the reorganization energies for charged defects is significantly lower (see right panel of Fig. 11).

4.2. Charge transition levels

In Section 2.4 we described how we obtain the CTLs by combining Slater-Janak transition state theory on a static lattice with geometry relaxations in the final state. For "negative" transitions, e.g. (0/-1), the effect of the relaxation is to lower the energy cost of adding the electron, i.e. the reorganisation energy lowers the CTL. In contrast, the lattice relaxations should produce an upward shift for "positive" transitions, e.g. (0/+1), because in this case the CTL denotes the negative of the energy cost of removing the electron.

Figure 12 shows the 0/+1 and 0/-1 CTLs for a small subset of defects. Since the energies are plotted relative to the vacuum level, the CTLs correspond to the (negative) ionisation potential (-IP) and electron affinity (EA), respectively. Results are shown both with and without the inclusion of relaxation effects. As expected, the relaxation always lowers the EA and the IP (the 0/+1 CTL is always raised). The reorganisation energies can vary from essentially zero to more than 2 eV, and are absolutely crucial for a correct prediction of CTLs and (charged) defect formation energies.

We notice that the CTLs always fall inside the band gap of the pristine host (marked by the grey bars) or very close to the band edges. This is clearly expected on physical grounds, as even for a defective system the CTLs cannot exceed the band edges (there are always electrons/holes available at the VBM/CBM sufficiently far away from the point defect). However,
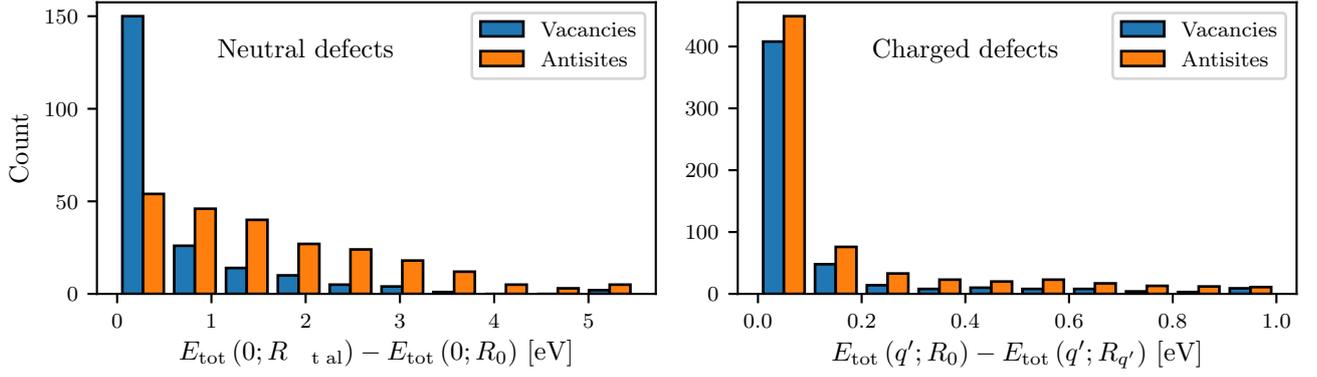
Figure 11. Relaxation effects for the formation of charged and neutral defects. Left: histogram of the reorganization energy from the initial defect substitution to the neutral equilibrium configuration. Low values on the x-axis correspond to small reorganization of defect structures upon addition of a defect to the pristine host crystal. Right: histogram of the reorganization energy between a neutral defect and its charged counterpart.
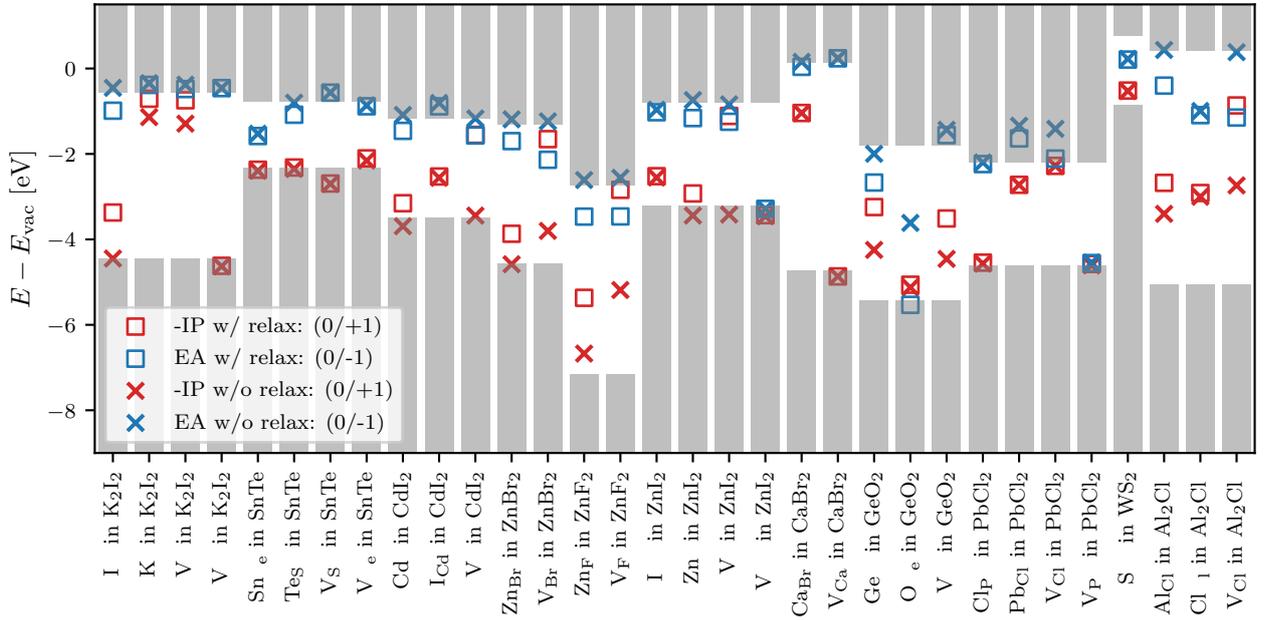

Figure 12. Relaxation effects for ionisation potentials and electron affinities. Energies of -IP (red symbols) and EA (blue symbols) with and without relaxation effects included (boxes and crosses, respectively). The energies are all referenced to the vacuum level of the pristine crystal and grey bars represent the valence/conduction band of the individual host crystals.

for small supercells such behavior is not guaranteed as the band gap of the defective crystal could deviate from that of the pristine host material. Thus the fact that the CTLs rarely appear outside the band gap is an indication that the employed supercells are large enough to represent an isolated defect.

When the Fermi level is moved from the VBM to the CBM one expects to fill available defects states with electrons in a stepwise manner, i.e. such that $CTL(q) < CTL(q-1)$. In particular, we expect $-IP < EA$. Interestingly, for a few systems, e.g. V_F in ZF_2, the ordering of IP and EA is inverted.

The physical interpretation of such an ordering is that the neutral charge state becomes thermodynamically unstable with respect to positive and negative charge states [20]. This results in a direct transition from positive to negative charge state in the formation energy diagram.

4.3. Intrinsic carrier concentrations

For many of the potential applications of 2D semiconductors, e.g. transistors[62], light emitting devices[63], or photo detectors[64], the question of

dopability of the semiconductor material is crucial. Modulation of the charge carrier concentration is a highly effective means of controlling the electrical and optical properties of a semiconductor. This holds in particular for 2D semiconductors whose carrier concentrations can be modulated in a variety of ways including electrostatic or ionic gating[65, 66], ion intercalation[67], or surface functionalisation[68]. In general, these methods are only effective if the material is not too heavily doped by its ubiquitous native defects, which may pin the Fermi level close to one of the band edges. For applications relying on high carrier conduction rather than carrier control, a high intrinsic carrier concentration may be advantageous - at least if the native defects do not degrade the carrier mobility too much, see Sec. 4.6.

Figure 13 shows the calculated position of the equilibrium Fermi level (at room temperature) for all the 2D materials considered as defect hosts in this work. It should be noted that the Fermi level position depends on the number of different defect types included in the analysis. Consequently, the results are sensitive to the existence of other types of intrinsic defects with formation energies lower than or comparable to the vacancy and antisite defects considered here. Fermi level regions close to the VBM or CBM (indicated by red/blue colors) correspond to $p-$type and $n-$type behavior, respectively, while Fermi levels in the central region of the band gap correspond to intrinsic behavior.

Clearly, most of the materials present in the QPOD database show either intrinsic or n-type behavior. An example of a well known material with intrinsic behavior is MoS_2, where native defects pin the Fermi level deep within the band gap resulting in very low electron and hole carrier concentrations in good agreement with previous observations[69, 70, 71, 45]. As an example of a natural n-type semiconductor we highlight the Janus monolayer AsClSe, which presents an impressive electron carrier concentration of 9.5×10^{11} cm^{-2} at 300 K, making it an interesting candidate for a high-conductivity 2D material. For the majority of the materials, the equilibrium carrier concentrations are in fact relatively low implying a high degree of dopability. We note that none of the materials exhibit intrinsic p-type behavior. This observation indicates that the challenge of finding naturally p-doped semiconductors/insulators, which is well known for bulk materials[72, 73, 74, 75], carries over to the class of 2D materials.

4.4. Defect formation energies: Trends and correlations

The formation energy is the most basic property of a point defect. Figure 14 the distribution of the
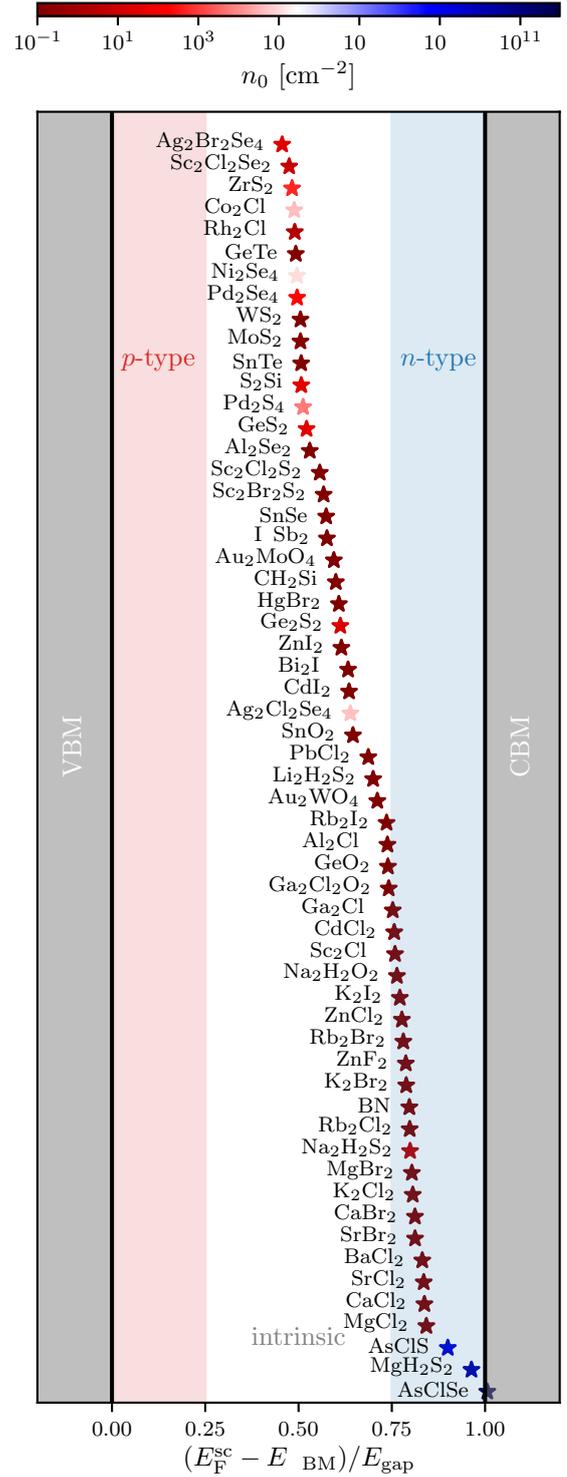

Figure 13. Self-consistent Fermi-level position for host materials within the pristine band gap. The energy scale is normalized with respect to each individual pristine band gap. The regions close to the CBM, VBM correspond to the p-, n-type dopable regimes, respectively. The colorcode of the markers represents the equilibrium electron carrier concentration n_0.
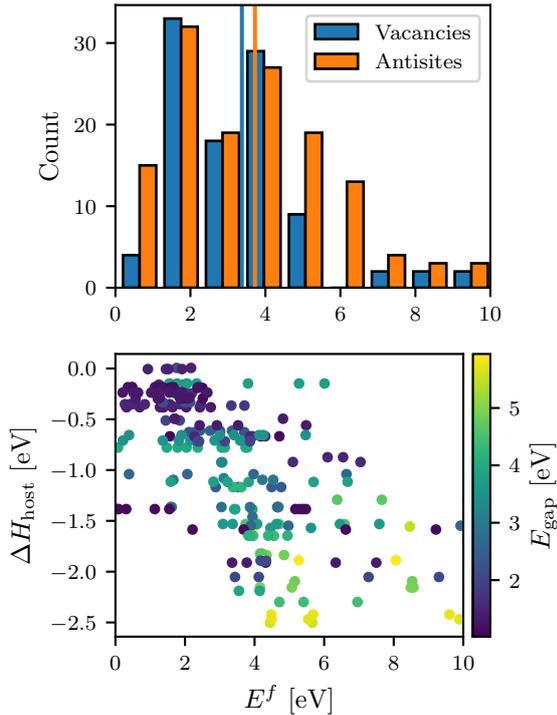
Figure 14. Distribution of defect formation energies in the neutral charge state. Top: histogram of neutral formation energies E^f wrt. standard states for vacancy and substitutional defects. The vertical lines represent the mean value. Bottom: heat of formation ΔH_{host} of the pristine monolayer as a function of the neutral formation energies with the pristine PBE-calculated band gap as a color code.

Figure 15. Distribution of point groups for relaxed defects. Point groups are ordered from the lowest symmetry group (i.e. C_1) to the highest occurring symmetry group (D_{3d}).

calculated formation energies of neutral defects for both vacancy (blue) and antisite (orange) defects. For the chemical potential appearing in Eq. (3) we used the standard state of each element. There is essentially no difference between the two distributions, and both means (vertical lines) are very close to 4 eV. Roughly half (44 %) of the neutral point defects have a formation energy below 3 eV and 28 % are below 2 eV. This implies that many of the defects would form readily during growth and underlines the importance of including intrinsic defects in the characterization of 2D materials. As a reference, the NV center in diamond shows formation energies on the order of 5 eV to 6 eV[76] (HSE value).

The lower panel of Figure 14 shows the defect formation energy relative to the heat of formation of the pristine host material, ΔH_{host}. There is a clear correlation between the two quantities, which may not come as a surprise since the ΔH_{host} measures the gain in energy upon forming the material from atoms in the standard states. More stable materials, i.e. materials with more negative heat of formation, are thus less prone to defect formation than less stable
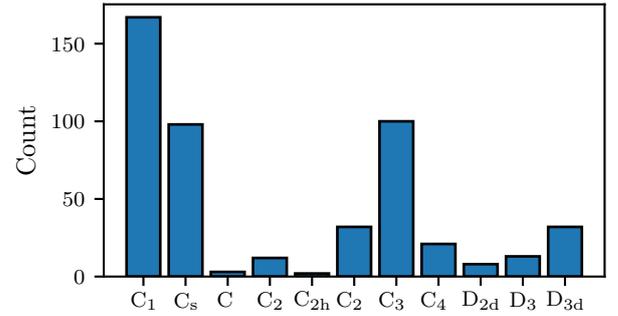
materials. It is also interesting to note the correlation with the band gap of the host material, indicated by the color of the symbols. A large band gap is seen to correlate with a large (negative) ΔH_{host} and large defect formation energies, and vice versa a small band gap is indicative of a smaller (negative) ΔH_{host} and low defect formation energies. These trends are somewhat problematic as low band gap materials with long carrier lifetimes, and thus low defect concentrations, are required for many applications in (opto)electronics, while large band gap host materials with high density of (specific types of) defects are required for many color center-based quantum technology applications.

4.5. Point defect symmetries

Another basic property of a point defects is its local symmetry, which determines the possible degeneracies of its in-gap electronic states and defines the selection rules for optical transitions between them. Figure 15 shows the distribution of point groups (in Schönflies notation) for all the investigated neutral defect systems. A large fraction of the defects break all the symmetries of the host crystal (34 % in C_1) or leave the system with only a mirror symmetry (20 % in C_s). A non-negligible number (31 %) of defects can be characterized by 2, 3, or 4-fold rotation axis with vertical mirror planes (C_{2v}, C_{3v} and C_{4v}). Naturally, C_{3v} defects (20 % overall) often stem from hexagonal host structures, some examples being V_S and W_S in WS_2, V_{Se} in SnSe, and Ca_{Br} in $CaBr_2$. Such defects share the symmetry group of the well known NV center in diamond[77] and might be particularly interesting candidates for quantum technology applications. Relatively few defect systems (3 %) incorporate perpendicular rotation axes, e.g. the D_3 symmetry of I_{Sb} in I_6Sb_2.

Adding or removing charge to a particular defect system can influence the structural symmetry as it was previously observed for the negatively charged sulfur

vacancy in MoS_2[78]. We find that 10 % of the defects in QPOD undergo a change in point group when adding (removing) an electron to (from) the neutral structure and relaxing it in its respective charge state.

4.6. Defect tolerant materials

Although defects can have useful functions and applications, they are often unwanted as they tend to deteriorate the ideal properties of the perfect crystal. Consequently, finding defect tolerant semiconductors[53, 79], i.e. semiconductors whose electronic and optical properties are only little influenced by the presence of their native defects, is of great interest.

When discussing defect tolerance of semiconductors one should distinguish between two different situations: (i) For transport applications where the system is close to equilibrium, defects act as scattering centres limiting the carrier mobility. In this case, charged defects represent the main problem due to their long range Coulomb potential, which leads to large scattering cross sections. (ii) For opto-electronic applications where relying on photo-excited electron-hole pairs, deep defect levels in the band gap represent the main issue as they facilitate carrier capture and promote non-radiative recombination. In the following we examine our set of host materials with respect to type (ii) defect tolerance.

Figure 16 shows the positions of charge transition levels of vacancy (blue) and antisite (orange) defects as a function of the Fermi level normalized to a host material's band gap. The neutral formation energy of the defects is shown in the middle panel, where we have also marked regions of shallow defect states that lie within 10% of the band edges (black dashed lines). A host material is said to be type (ii) defect tolerant if all of its intrinsic defects are shallow or all its deep defects have high formation energy. A number of defect tolerant host materials are revealed by this analysis, including the ionic halides K_2Cl_2, Rb_2I_2, and Rb_2Cl_2. With formation energies lying about 150 meV/atom above those of their cubic bulk structures, these materials may be challenging to realise in atomically thin form. Nevertheless, their defect tolerant nature fits well with the picture of deep defect states having larger tendency to form in covalently bonded insulators with bonding/anti-bonding band gap types compared to ionic insulators with charge-transfer type band gaps[53].

The data in Figure 16 suggests that vacancy and antisite defects have different tendencies to form shallow and deep defect states, respectively: While 55 % of the vacancy defects form shallow defect states, this only happens for 30 % of the antisites. The trend can be seen in the top and bottom histograms of Fig.
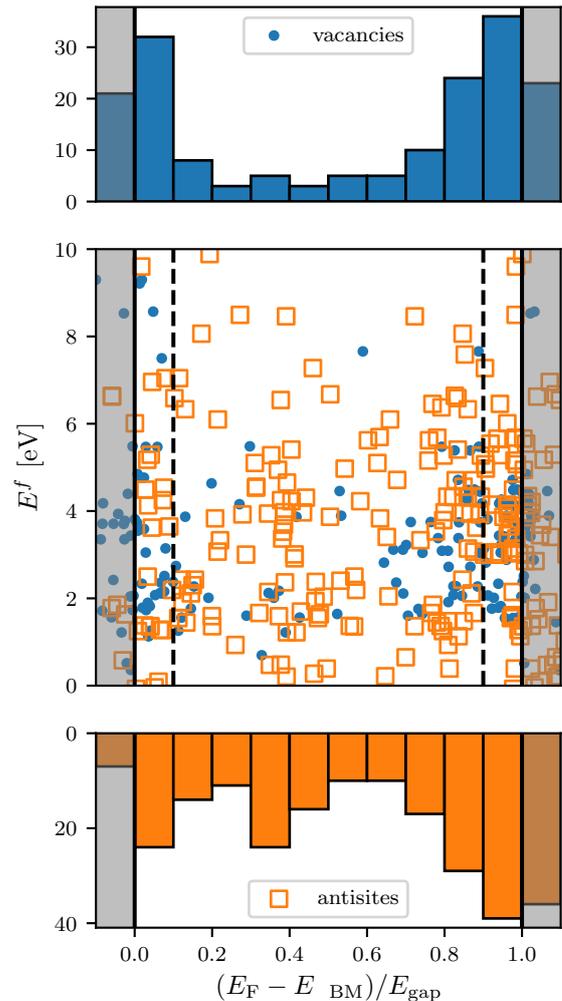
Figure 16. Defect tolerances for vacancy and substitutional defects. Position of defect charge transition levels as a function of Fermi energy (normalized with respect to the band gap) for antisite defects (orange squares) and vacancies (blue dots). Top (bottom): histogram of the occurance of a CTL for vacancies (antisites) within a certain energy range. If a defect's CTL lies above 0.9/below 0.1 (dashed black vertical lines in the middle plot) it corresponds to a CTL not having a detrimental effect on the host's properties. Materials with these CTL are considered *defect-tolerant* wrt. optical properties. The grey areas on the left and right hand side represent the valence band and conduction band, respectively.

16 where the vacancy distribution (top) shows fewer CTL around the center of the band gap compared to the antisite distribution (bottom). Based on this we conclude that vacancy defects are, on average, less detrimental to the optical properties of semiconductors than antisite defects.

4.7. Potential Defects for Quantum Technological Applications

The defect systems with high spin ground state are widely sought after, as such systems can be spin initialized, manipulated and can be read out. Such systems can be exploited in Optically detected Magnetic Resonance spectroscopy to act as qubits, nano-scale sensors, brain magnetometers and other quantum technological applications etc. An overview of the distribution of total magnetic moments for all the defect systems in show in Fig 17. In the present work, we have found a total of 80 high spin-systems, which are then screened for a (PBE) KS gap in the range 0.5-1.5 eV for the excited state calculations, with the motivation of finding system exhibiting sharp radiative transitions in the visible range of energies. In this section we present a few high spin systems i.e. with triplet ground state and exhibit sharp photo-luminescence i.e. small Huang-Rhys factors. Our results, not only can motivate the experimental research community to investigate these host systems for such defects, but can also help to identify these defects, by a comparision with experimental photo-luminescence spectra[80]. We present the calculated PL lineshape using generating function approach Pb_{Cl}^-, C_H^- and Si_H^- defect systems. The upper and bottom panels in the figure 18 show the calculated lineshape for HOMO to LUMO transition in the majority spin channel for Pb_{Cl}^- and C_H^- defects, respectively. The Huang-Rhys factor for these transition is extremely small. Infact the radiative transitions with such small systems have seldom been reported previously [6].The 2nd panel shows the PL spectra for HOMO to LUMO tranistion in the minority spin channel for Si_H^- defect in SiCH2 host material, while the third panel shows the PL spectra for the HOMO to LUMO transition within the singlet manifold of the same defect. The symmetry of this defect i.e. C_3v and KS level structure is somewhat similar to N_V^-1 centre in diamond. The fact that the calculated huang-Rhys factor is also close to the one reported for N_V^-1 centre in diamond [60], makes this defect system very interesting, as it could act as a 2D counter part of N_V^-1 in diamond for quantum technological applications.

5. Summary and Outlook

In summary, more than 1500 intrinsic point defects hosted by 82 different 2D semiconductors and insulators have been relaxed and characterised by means of DFT calculations. The 1500+ point defects comprise about 500 structurally distinct vacancy and antisite defects in different charge states. The thousands of DFT calculations were orchestrated by a computational workflow build using the Atomic
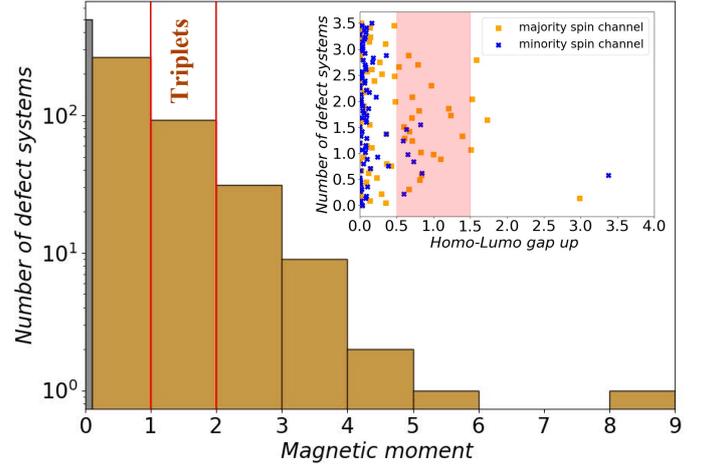

Figure 17. An overview of distribution of total magnetic moments for all defect systems. The bin corresponding to systems with triplet ground state are highlighted between two red lines, The inset shows the distribution of Kohn-Sham HOMO-LUMO gap, on the energy scale, for the triplets. The shaded area in the inset represents the energy range from where systems are picked for excited state calculations.

Simulation Recipes (ASR)[27] and executed with the MyQueue[28] task scheduler.

The ASR defect workflow includes the calculation of formation energies, charge transition levels, equilibrium defect and carrier concentrations, point group symmetry labels of in-gap states, transition dipole moments, hyperfine coupling, and zero field splittings. All DFT calculations were performed using the semilocal PBE xc-functional. While the PBE should yield accurate structural and thermodynamic properties, its tendency to underestimate band gaps can result in high-energy defect levels being missed, if they lie above the PBE band gap. We emphasise, however, that charge transition levels appearing within the PBE band gap should be well described by our PBE-based Slater-Janak transition state approach.

The analysis of defect formation energies showed that many of the investigated defects are likely to be present in the host material in appreciable concentrations thus confirming their relevance and importance for the general properties of the host materials. Based on the thermodynamic and electronic characterisation of all the intrinsic defects, we identified a number of defect tolerant ionic insulators among the host materials. While these specific materials may be challenging to synthesise due to competing bulk phases, the results can be used to guide future searches for defect tolerant semiconductors for high-performance opto-electronic 2D devices. Among the 82 host materials, we found several semiconductors with high intrinsic electron carrier concentrations, e.g.
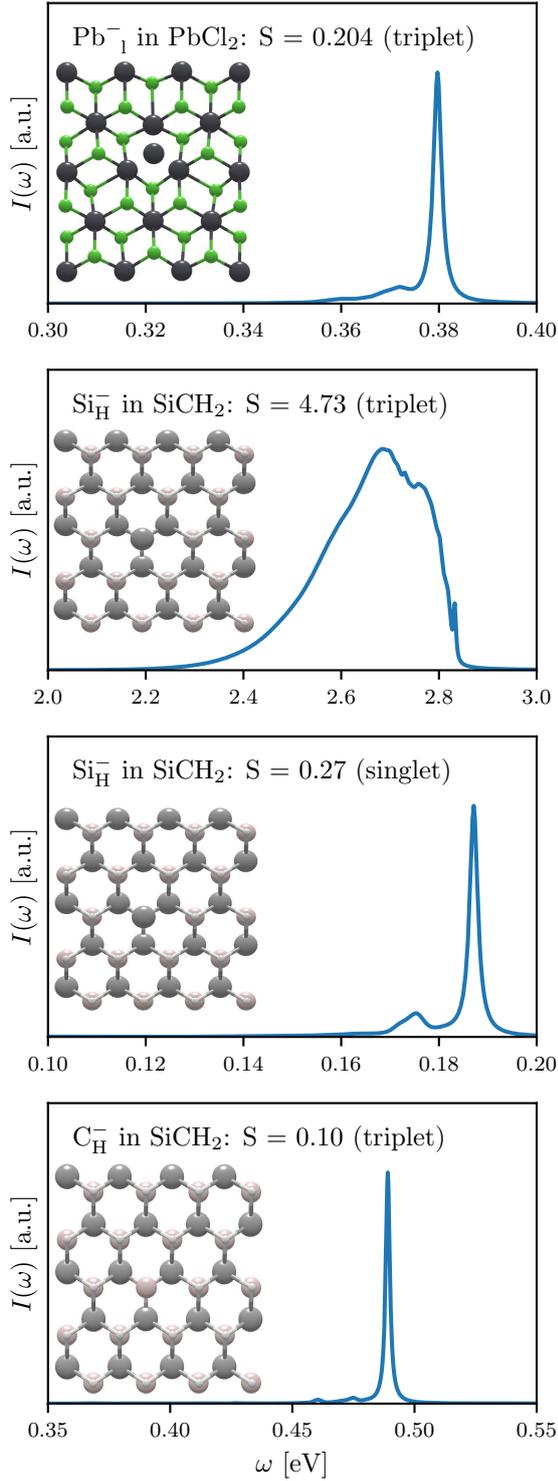
Figure 18. PL lineshapes for a few selected systems using generating function approach. A few selected systems with small Huang-Rhys factors are shown. The top panel show the PL line shape for HOMO to LUMO transition in the majority spin channel for Pb_{Cl}^- defect. The middle two panels shows the PL spectra for HOMO to LUMO transition in the minority spin channel (2nd panel) and HOMO to LUMO transition within the singlet manifold (3rd panel) for Si_H^- defect in $SiCH_2$ host material. The bottom panel shows the calculated spectra for C_H^- defect in $SiCH_2$ for HOMO to LUMO transition in the majority spin channel. The calculated radiative life times of these transitions (top to bottom) are 4.6×10^{-4} ns, 4.9×10^{-2} ns, 2.9×10^{-4} ns and 9.0×10^{-4} ns, repectively.

MgS_2H_2 and the Janus materials AgClS and AgClSe, whereas no the intrinsic *p*-type materials were found.

Out of the > 1500 defects, only around 100 adopt a high-spin ($S > 1/2$) ground state. This indicates that the simple defect types considered in this work are not likely to yield useful spin defects for quantum technology applications. On the other hand, a number of non-magnetic (singlet) defects showed low Huang-Rhys factors and correspondingly narrow photoluminescence (PL) spectra making them relevant as single-photon emitters. In this work, the excited state properties were calculated and analyzed manually. Incorporating this part into the automated workflow will be an important future extension of the current methodology that is critical to enable a systematic and rational design of defects with ideal excited state properties including transition energies, excited state lifetimes/dynamics, and emission line shapes.

The web-based presentation of the QPOD database and its seamless integration with the Computational 2D Materials Database (C2DB) makes a unique platform for exploring the physics of point defects in 2D materials, which should be useful both as a convenient lookup table and as a benchmark reference for computational studies. Moreover, the possibility to download the entire database makes it applicable for machine learning purposes, which has a large yet untapped potential for establishing structure-property relationships for point defects.

Looking ahead, there are many possible extensions and improvements of the current work. First of all, it would be interesting to move beyond the PBE to more advanced xc-functionals, such as screened hybrids. This would not only enhance the data quality/accuracy but also provide the basis for a systematic and statistically significant assessment of the performance of the PBE for defect calculations. It would also be relevant to expand the set of host materials beyond the 82 materials considered here. The C2DB currently contains about 500 monolayers exfoliated from experimentally known layered van der Waals crystals and a similar number of predicted highly stable monolayers providing ample opportunities for selecting ideal 2D host crystals. As mentioned, the simple defects considered in this work turned out to be mostly non-magnetic. Thus for applications relying on spinful defect centers, e.g. magnetic field sensors or qubits, it seems important to incorporate more complex defects such as divacancies and vacancy-substitutional defect pairs.

6. Acknowledgments

References

[1] Polman A, Knight M, Garnett E C, Ehrler B and Sinke W C 2016 *Science* **352**

[2] Awschalom D D, Bassett L C, Dzurak A S, Hu E L and Petta J R 2013 *Science* **339** 1174–1179

[3] Eckstein J N and Levy J 2013 *MRS Bulletin* **38** 783–789

[4] Lovchinsky I, Sushkov A, Urbach E, de Leon N P, Choi S, De Greve K, Evans R, Gertner R, Bersin E, Müller C *et al.* 2016 *Science* **351** 836–841

[5] Bradley C, Randall J, Abobeih M H, Berrevoets R, Degen M, Bakker M, Markham M, Twitchen D and Taminiau T H 2019 *Physical Review X* **9** 031045

[6] Sajid A, Ford M J and Reimers J R 2020 *Reports on Progress in Physics* **83** 044501

[7] Grosso G, Moon H, Lienhard B, Ali S, Efetov D K, Furchi M M, Jarillo-Herrero P, Ford M J, Aharonovich I and Englund D 2017 *Nature communications* **8** 1–8

[8] Ferrari A C, Bonaccorso F, Fal'Ko V, Novoselov K S, Roche S, Bøggild P, Borini S, Koppens F H, Palermo V, Pugno N *et al.* 2015 *Nanoscale* **7** 4598–4810

[9] Lin Z, Carvalho B R, Kahn E, Lv R, Rao R, Terrones H, Pimenta M A and Terrones M 2016 *2D Materials* **3** 022002

[10] Aharonovich I, Englund D and Toth M 2016 *Nature Photonics* **10** 631–641

[11] Tran T T, Bray K, Ford M J, Toth M and Aharonovich I 2016 *Nature Nanotechnology* **11** 37–41

[12] Fischer M, Caridad J, Sajid A, Ghaderzadeh S, Ghorbani-Asl M, Gammelgaard L, Bøggild P, Thygesen K S, Krasheninnikov A, Xiao S *et al.* 2021 *Science Advances* **7** eabe7138

[13] Bertolazzi S, Bonacchi S, Nan G, Pershin A, Beljonne D and Samorì P 2017 *Advanced materials* **29** 1606760

[14] Chakraborty C, Kinnischtzke L, Goodfellow K M, Beams R and Vamivakas A N 2015 *Nature nanotechnology* **10** 507–511

[15] Koperski M, Nogajewski K, Arora A, Cherkez V, Mallet P, Veuillen J Y, Marcus J, Kossacki P and Potemski M 2015 *Nature nanotechnology* **10** 503–506

[16] Zhukov I, Anishchik S and Molin Y N 2017 *Applied Magnetic Resonance* **48** 1461–1469

[17] Chejanovsky N, Mukherjee A, Geng J, Chen Y C, Kim Y, Denisenko A, Finkler A, Taniguchi T, Watanabe K, Dasari D B R *et al.* 2021 *Nature materials* 1–6

[18] Ye G, Gong Y, Lin J, Li B, He Y, Pantelides S T, Zhou W, Vajtai R and Ajayan P M 2016 *Nano Letters* **16** 1097–1103

[19] Xie J, Yang X and Xie Y 2020 *Nanoscale* **12** 4283–4294

[20] Freysoldt C, Grabowski B, Hickel T, Neugebauer J, Kresse G, Janotti A and Van de Walle C G 2014 *Reviews of modern physics* **86** 253

[21] Janotti A and Van de Walle C G 2007 *Physical Review B* **76** 165202

[22] Neugebauer J and Van de Walle C G 1996 *Applied Physics Letters* **69** 503–505

[23] Dreyer C E, Alkauskas A, Lyons J L, Janotti A and Van de Walle C G 2018 *Annual Review of Materials Research* **48** 1–26

[24] Gupta S, Yang J H and Yakobson B I 2018 *Nano letters* **19** 408–414

[25] Jain A, Ong S P, Chen W, Medasani B, Qu X, Kocher M, Brafman M, Petretto G, Rignanese G M, Hautier G *et al.* 2015 *Concurrency and Computation: Practice and Experience* **27** 5037–5059

[26] Pizzi G, Cepellotti A, Sabatini R, Marzari N and Kozinsky B 2016 *Computational Materials Science* **111** 218–230

[27] Gjerding M, Skovhus T, Rasmussen A, Bertoldo F, Larsen A H, Mortensen J J and Thygesen K S 2021 *arXiv preprint arXiv:2104.13431*

[28] Mortensen J J, Gjerding M and Thygesen K S 2020 *Journal of Open Source Software* **5** 1844

[29] Haastrup S, Strange M, Pandey M, Deilmann T, Schmidt P S, Hinsche N F, Gjerding M N, Torelli D, Larsen P M, Riis-Jensen A C *et al.* 2018 *2D Materials* **5** 042002

[30] Gjerding M, Taghizadeh A, Rasmussen A, Ali S, Bertoldo F, Deilmann T, Holguin U, Knøsgaard N, Kruse M, Manti S *et al.* 2021 *arXiv preprint arXiv:2102.03029*

[31] Broberg D, Medasani B, Zimmermann N E, Yu G, Canning A, Haranczyk M, Asta M and Hautier G 2018 *Computer Physics Communications* **226** 165–179

[32] Péan E, Vidal J, Jobic S and Latouche C 2017 *Chemical Physics Letters* **671** 124–130

[33] Goyal A, Gorai P, Peng H, Lany S and Stevanović V 2017 *Computational Materials Science* **130** 1–9

[34] Davidsson J, Ivády V, Armiento R and Abrikosov I A 2021 *Computer Physics Communications* 108091

[35] Perdew J P, Burke K and Ernzerhof M 1996 *Physical review letters* **77** 3865

[36] Janak J F 1978 *Physical Review B* **18** 7165

[37] Levi G, Ivanov A V and Jonsson H 2020 *Faraday Discussions*

[38] Mori-Sánchez P, Cohen A J and Yang W 2008 *Physical Review Letters* **100** 146401

[39] Lyons J L and Van de Walle C G 2017 *NPJ Computational Materials* **3** 1–10

[40] Enkovaara J, Rostgaard C, Mortensen J J, Chen J, Dułak M, Ferrighi L, Gavnholt J, Glinsvad C, Haikola V, Hansen H *et al.* 2010 *Journal of Physics: Condensed Matter* **22** 253202

[41] Pulay P 1980 *Chemical Physics Letters* **73** 393–398

[42] Van de Walle C G, Laks D, Neumark G and Pantelides S 1993 *Physical Review B* **47** 9425

[43] Zhang S and Northrup J E 1991 *Physical review letters* **67** 2339

[44] Van de Walle C G and Neugebauer J 2004 *Journal of applied physics* **95** 3851–3879

[45] Komsa H P and Krasheninnikov A V 2015 *Physical Review B* **91** 125304

[46] Komsa H P, Berseneva N, Krasheninnikov A V and Nieminen R M 2014 *Physical Review X* **4** 031044

[47] Liu X, Gao Z, Wang V, Luo Z, Lv B, Ding Z and Zhang Z 2020 *ACS Applied Materials & Interfaces* **12** 17055–17061

[48] Xia S, Wang D, Chen N K, Han D, Li X B and Sun H B 2020 *Annalen der Physik* **532** 1900318

[49] Xiao J, Yang K, Guo D, Shen T, Deng H X, Li S S, Luo J W and Wei S H 2020 *Physical Review B* **101** 165306

[50] Göransson C, Olovsson W and Abrikosov I A 2005 *Physical Review B* **72** 134203

[51] Sanna S, Frauenheim T and Gerstmann U 2008 *Physical Review B* **78** 085201

[52] Li Y, Sanna S and Schmidt W G 2014 *The Journal of Chemical Physics* **140** 234113

[53] Pandey M, Rasmussen F A, Kuhar K, Olsen T, Jacobsen K W and Thygesen K S 2016 *Nano letters* **16** 2234–2239

[54] Buckeridge J 2019 *Computer Physics Communications* **244** 329–342

[55] Kaappa S, Malola S and Häkkinen H 2018 *The Journal of Physical Chemistry A* **122** 8576–8584

[56] A T 2009 Spglib URL `https://atztogo.github.io/spglib/`

[57] Cornwell J F 1997 *Group Theory in Physics: An Introduction* (Academic press)

[58] Stoneham A M 2001 *Theory of defects in solids: electronic structure of defects in insulators and semiconductors* (Oxford University Press)

[59] Sajid A, Thygesen K S, Reimers J R and Ford M J 2020 *Communications Physics* **3** 153 ISSN 2399-3650 URL `https://doi.org/10.1038/s42005-020-00416-z`

[60] Alkauskas A, Buckley B B, Awschalom D D and Van de Walle C G 2014 *New Journal of Physics* **16** 073026

[61] Larsen A H, Mortensen J J, Blomqvist J, Castelli I E, Christensen R, Dułak M, Friis J, Groves M N, Hammer B, Hargus C *et al.* 2017 *Journal of Physics: Condensed Matter* **29** 273002

[62] Radisavljevic B, Radenovic A, Brivio J, Giacometti V and Kis A 2011 *Nature Nanotechnology* **6** 147–150

[63] Withers F, Del Pozo-Zamudio O, Mishchenko A, Rooney A, Gholinia A, Watanabe K, Taniguchi T, Haigh S J, Geim A, Tartakovskii A *et al.* 2015 *Nature Materials* **14** 301–306

[64] Koppens F, Mueller T, Avouris P, Ferrari A, Vitiello M and Polini M 2014 *Nature Nanotechnology* **9** 780–793

[65] Wang Q H, Kalantar-Zadeh K, Kis A, Coleman J N and Strano M S 2012 *Nature Nanotechnology* **7** 699–712

[66] Wang F, Stepanov P, Gray M, Lau C N, Itkis M E and Haddon R C 2015 *Nano Letters* **15** 5284–5288

[67] Yue Z 2021 *Nature Reviews Physics* **3** 306–306

[68] Xiang D, Han C, Wu J, Zhong S, Liu Y, Lin J, Zhang X A, Hu W P, Özyilmaz B, Neto A C *et al.* 2015 *Nature Communications* **6** 1–8

[69] Shang M H, Hou H, Zheng J, Yang Z, Zhang J, Wei S, Duan X and Yang W 2018 *The journal of physical chemistry letters* **9** 6032–6037

[70] Splendiani A, Sun L, Zhang Y, Li T, Kim J, Chim C Y, Galli G and Wang F 2010 *Nano letters* **10** 1271–1275

[71] Noh J Y, Kim H and Kim Y S 2014 *Physical Review B* **89** 205417

[72] Fortunato E, Barquinha P and Martins R 2012 *Advanced materials* **24** 2945–2986

[73] Kormath Madam Raghupathy R, Wiebeler H, Kühne T D, Felser C and Mirhosseini H 2018 *Chemistry of Materials* **30** 6794–6800

[74] Xu J, Liu J B, Wang J, Liu B X and Huang B 2018 *Advanced Functional Materials* **28** 1800332

[75] Bhatia A, Hautier G, Nilgianskul T, Miglio A, Sun J, Kim H J, Kim K H, Chen S, Rignanese G M, Gonze X *et al.* 2016 *Chemistry of Materials* **28** 30–34

[76] Deák P, Aradi B, Kaviani M, Frauenheim T and Gali A 2014 *Physical review B* **89** 075203

[77] Lenef A and Rand S 1996 *Physical Review B* **53** 13441

[78] Tan A M Z, Freysoldt C and Hennig R G 2020 *Physical Review Materials* **4** 064004

[79] Walsh A and Zunger A 2017 *Nature materials* **16** 964–967

[80] Sajid A and Thygesen K S 2020 *2D Materials* **7** 031007

References

[1] G. Galilei, *The Language of Nature: Reassessing the Mathematization of Natural Philosophy in the Seventeenth Century.* 2016.

[2] P. W. Anderson, "More is different," *Science*, vol. 177, no. 4047, pp. 393–396, 1972.

[3] K. S. Novoselov, A. K. Geim, S. V. Morozov, D. Jiang, Y. Zhang, S. V. Dubonos, I. V. Grigorieva, and A. A. Firsov, "Electric field effect in atomically thin carbon films," *Science*, vol. 306, no. 5696, pp. 666–669, 2004.

[4] S. Iijima, "Helical microtubules of graphitic carbon," *Nature*, vol. 354, pp. 56–58, Nov 1991.

[5] D. Bimberg, M. Grundmann, and N. N. Ledentsov, *Quantum dot heterostructures.* John Wiley & Sons, 1999.

[6] K. F. Mak, C. Lee, J. Hone, J. Shan, and T. F. Heinz, "Atomically thin mos_2: A new direct-gap semiconductor," *Phys. Rev. Lett.*, vol. 105, p. 136805, Sep 2010.

[7] L.-J. Lu, Ang-Yu et. al., "Janus monolayers of transition metal dichalcogenides," *Nature Nanotechnology*, vol. 12, pp. 744–749, Aug 2017.

[8] L. C. Bassett, A. Alkauskas, A. L. Exarhos, and K.-M. C. Fu, "Quantum defects by design," *Nanophotonics*, vol. 8, no. 11, pp. 1867–1888, 2019.

[9] G. Davies, M. F. Hamer, and W. C. Price, "Optical studies of the 1.945 ev vibronic band in diamond," *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, vol. 348, no. 1653, pp. 285–298, 1976.

[10] M. W. Doherty, N. B. Manson, P. Delaney, F. Jelezko, J. Wrachtrup, and L. C. Hollenberg, "The nitrogen-vacancy colour centre in diamond," *Physics Reports*, vol. 528, no. 1, pp. 1–45, 2013. The nitrogen-vacancy colour centre in diamond.

[11] L. Childress and R. Hanson, "Diamond nv centers for quantum computing and quantum networks," *MRS Bulletin*, vol. 38, no. 2, p. 134–138, 2013.

[12] R. Schirhagl, K. Chang, M. Loretz, and C. L. Degen, "Nitrogen-vacancy centers in diamond: Nanoscale sensors for physics and biology," *Annual Review of Physical Chemistry*, vol. 65, no. 1, pp. 83–105, 2014. PMID: 24274702.

[13] W. F. Koehl, B. B. Buckley, F. J. Heremans, G. Calusine, and D. D. Awschalom, "Room temperature coherent control of defect spin qubits in silicon carbide," *Nature*, vol. 479, pp. 84–87, Nov 2011.

[14] D. D. Falk, Abram L. et. al., "Polytype control of spin qubits in silicon carbide," *Nature Communications*, vol. 4, p. 1819, May 2013.

[15] T. T. Tran, K. Bray, M. J. Ford, M. Toth, and I. Aharonovich, "Quantum emission from hexagonal boron nitride monolayers," *Nature Nanotechnology*, vol. 11, pp. 37–41, Jan 2016.

[16] G. et. al. and Vladimir, "Initialization and read-out of intrinsic spin defects in a van der waals crystal at room temperature," *Nature Materials*, vol. 19, pp. 540–545, May 2020.

[17] P. M. Ci, Lijie et. al., "Atomic layers of hybridized boron nitride and graphene domains," *Nature Materials*, vol. 9, pp. 430–435, May 2010.

[18] J. Klein, L. Sigl, S. Gyger, K. Barthelmi, M. Florian, S. Rey, T. Taniguchi, K. Watanabe, F. Jahnke, C. Kastl, V. Zwiller, K. D. Jöns, K. Müller, U. Wurstbauer, J. J. Finley, and A. W. Holleitner, "Engineering the luminescence and generation of individual defect emitters in atomically thin mos2," *ACS Photonics*, vol. 8, no. 2, pp. 669–677, 2021.

[19] C. Mitterreiter et. al., Kastl, "The role of chalcogen vacancies for atomic defect emission in mos2," *Nature Communications*, vol. 12, p. 3822, Jun 2021.

[20] L. Yu, M. Deng, J. L. Zhang, S. Borghardt, B. Kardynal, J. Vučković, and T. F. Heinz, "Site-controlled quantum emitters in monolayer mose2," *Nano Letters*, vol. 21, no. 6, pp. 2376–2381, 2021. PMID: 33689386.

[21] B. Schuler, D. Y. Qiu, S. Refaely-Abramson, C. Kastl, C. T. Chen, S. Barja, R. J. Koch, D. F. Ogletree, S. Aloni, A. M. Schwartzberg, J. B. Neaton, S. G. Louie, and A. Weber-Bargioni, "Large spin-orbit splitting of deep in-gap defect states of engineered sulfur vacancies in monolayer ws_2," *Phys. Rev. Lett.*, vol. 123, p. 076801, Aug 2019.

[22] L. Linhart, M. Paur, V. Smejkal, J. Burgdörfer, T. Mueller, and F. Libisch, "Localized intervalley defect excitons as single-photon emitters in wse_2," *Phys. Rev. Lett.*, vol. 123, p. 146401, Sep 2019.

[23] M. Born and R. Oppenheimer, "Zur quantentheorie der molekeln," *Annalen der Physik*, vol. 389, no. 20, pp. 457–484, 1927.

[24] G. Grosso and G. Parravicini, *Solid State Physics.* Elsevier Science, 2000.

[25] G. m. H. Thiering and A. Gali, "Ab initio calculation of spin-orbit coupling for an nv center in diamond exhibiting dynamic jahn-teller effect," *Phys. Rev. B*, vol. 96, p. 081115, Aug 2017.

[26] C. J. Ciccarino, J. Flick, I. B. Harris, M. E. Trusheim, D. R. Englund, and P. Narang, "Strong spin–orbit quenching via the product jahn–teller effect in neutral group iv qubits in diamond," *npj Quantum Materials*, vol. 5, p. 75, Oct 2020.

[27] H. A. Jahn, E. Teller, and F. G. Donnan, "Stability of polyatomic molecules in degenerate electronic states - i—orbital degeneracy," *Proceedings of the Royal Society of London. Series A - Mathematical and Physical Sciences*, vol. 161, no. 905, pp. 220–235, 1937.

[28] P. Hohenberg and W. Kohn, "Inhomogeneous electron gas," *Phys. Rev.*, vol. 136, pp. B864–B871, Nov 1964.

[29] W. Kohn and L. J. Sham, "Self-consistent equations including exchange and correlation effects," *Phys. Rev.*, vol. 140, pp. A1133–A1138, Nov 1965.

[30] D. M. Ceperley and B. J. Alder, "Ground state of the electron gas by a stochastic method," *Phys. Rev. Lett.*, vol. 45, pp. 566–569, Aug 1980.

[31] J. P. Perdew, K. Burke, and M. Ernzerhof, "Generalized gradient approximation made simple," *Phys. Rev. Lett.*, vol. 77, pp. 3865–3868, Oct 1996.

[32] A. Alkauskas, P. Broqvist, and A. Pasquarello, "Defect levels through hybrid density functionals: Insights and applications," *physica status solidi (b)*, vol. 248, no. 4, pp. 775–789, 2011.

[33] J. Heyd, G. E. Scuseria, and M. Ernzerhof, "Hybrid functionals based on a screened coulomb potential," *The Journal of Chemical Physics*, vol. 118, no. 18, pp. 8207–8215, 2003.

[34] A. V. Krukau, O. A. Vydrov, A. F. Izmaylov, and G. E. Scuseria, "Influence of the exchange screening parameter on the performance of screened hybrid functionals," *The Journal of Chemical Physics*, vol. 125, no. 22, p. 224106, 2006.

[35] J. E. et. al., "Electronic structure calculations with GPAW: a real-space implementation of the projector augmented-wave method," *Journal of Physics: Condensed Matter*, vol. 22, p. 253202, jun 2010.

[36] P. E. Blöchl, "Projector augmented-wave method," *Phys. Rev. B*, vol. 50, pp. 17953–17979, Dec 1994.

[37] A. A. Maradudin, E. W. Montroll, G. H. Weiss, and I. Ipatova, *Theory of lattice dynamics in the harmonic approximation*, vol. 3. Academic press New York, 1963.

[38] S. Baroni, S. de Gironcoli, A. Dal Corso, and P. Giannozzi, "Phonons and related crystal properties from density-functional perturbation theory," *Rev. Mod. Phys.*, vol. 73, pp. 515–562, Jul 2001.

[39] D. Alfè, "Phon: A program to calculate phonons using the small displacement method," *Computer Physics Communications*, vol. 180, no. 12, pp. 2622–2633, 2009. 40 YEARS OF CPC: A celebratory issue focused on quality software for high performance, grid and novel computing architectures.

[40] Y. Wang, J. J. Wang, W. Y. Wang, Z. G. Mei, S. L. Shang, L. Q. Chen, and Z. K. Liu, "A mixed-space approach to first-principles calculations of phonon frequencies for polar materials," *Journal of Physics: Condensed Matter*, vol. 22, p. 202201, apr 2010.

[41] M. Gjerding, T. Skovhus, A. Rasmussen, F. Bertoldo, A. H. Larsen, J. J. Mortensen, and K. S. Thygesen, "Atomic simulation recipes – a python framework and library for automated workflows," *Computational Materials Science*, vol. 199, p. 110731, 2021.

[42] A. Togo and I. Tanaka, "First principles phonon calculations in materials science," *Scripta Materialia*, vol. 108, pp. 1 – 5, 2015.

[43] K. Kaasbjerg, K. S. Thygesen, and K. W. Jacobsen, "Phonon-limited mobility in n-type single-layer mos$_2$ from first principles," *Phys. Rev. B*, vol. 85, p. 115317, Mar 2012.

[44] A. Alkauskas, Q. Yan, and C. G. Van de Walle, "First-principles theory of nonradiative carrier capture via multiphonon emission," *Phys. Rev. B*, vol. 90, p. 075202, Aug 2014.

[45] C. Freysoldt, B. Grabowski, T. Hickel, J. Neugebauer, G. Kresse, A. Janotti, and C. G. Van de Walle, "First-principles calculations for point defects in solids," *Rev. Mod. Phys.*, vol. 86, pp. 253–305, Mar 2014.

[46] C. Freysoldt, J. Neugebauer, and C. G. Van de Walle, "Electrostatic interactions between charged defects in supercells," *physica status solidi (b)*, vol. 248, no. 5, pp. 1067–1076, 2011.

[47] H.-P. Komsa, N. Berseneva, A. V. Krasheninnikov, and R. M. Nieminen, "Charged point defects in the flatland: Accurate formation energy calculations in two-dimensional materials," *Phys. Rev. X*, vol. 4, p. 031044, Sep 2014.

[48] M. Pandey, F. A. Rasmussen, K. Kuhar, T. Olsen, K. W. Jacobsen, and K. S. Thygesen, "Defect-tolerant monolayer transition metal dichalcogenides," *Nano Letters*, vol. 16, no. 4, pp. 2234–2239, 2016. PMID: 27027786.

[49] J. F. Janak, "Proof that de/dn i= ε in density-functional theory," *Physical Review B*, vol. 18, no. 12, p. 7165, 1978.

[50] A. Alkauskas, M. D. McCluskey, and C. G. Van de Walle, "Tutorial: Defects in semiconductors—combining experiment and theory," *Journal of Applied Physics*, vol. 119, no. 18, p. 181101, 2016.

[51] J. Elliott and P. Dawber, *Symmetry in Physics*. No. v. 1 in Symmetry in Physics, Macmillan, 1979.

[52] M. Tinkham, *Group Theory and Quantum Mechanics*. Dover Books on Chemistry and Earth Sciences, Dover Publications, 2003.

[53] S. Altmann, "Point-group theory tables," 2011.

[54] W. Hergert and R. Geilhufe, "Group theory in solid state physics and photonics: Problem solving with mathematica," 2018.

[55] M. Cini, *Topics and methods in condensed matter theory*. Springer, 2007.

[56] S. Kaappa, S. Malola, and H. Häkkinen, "Point group symmetry analysis of the electronic structure of bare and protected metal nanocrystals," *The Journal of Physical Chemistry A*, vol. 122, no. 43, pp. 8576–8584, 2018.

[57] e. a. Hong, "Exploring atomic defects in molybdenum disulphide monolayers," *Nature Communications*, vol. 6, p. 6293, Feb 2015.

[58] H.-P. Komsa and A. V. Krasheninnikov, "Native defects in bulk and monolayer mos_2 from first principles," *Phys. Rev. B*, vol. 91, p. 125304, Mar 2015.

[59] B. Bransden and C. Joachain, *Physics of Atoms and Molecules.* Prentice Hall, 2003.

[60] J. R. Maze, A. Gali, E. Togan, Y. Chu, A. Trifonov, E. Kaxiras, and M. D. Lukin, "Properties of nitrogen-vacancy centers in diamond: the group theoretic approach," *New Journal of Physics*, vol. 13, p. 025025, feb 2011.

[61] A. Sajid, J. R. Reimers, and M. J. Ford, "Defect states in hexagonal boron nitride: Assignments of observed properties and prediction of properties relevant to quantum computation," *Phys. Rev. B*, vol. 97, p. 064101, Feb 2018.

[62] Ádám Gali, "Ab initio theory of the nitrogen-vacancy center in diamond:," *Nanophotonics*, vol. 8, no. 11, pp. 1907–1943, 2019.

[63] J. R. Weber, W. F. Koehl, J. B. Varley, A. Janotti, B. B. Buckley, C. G. Van de Walle, and D. D. Awschalom, "Quantum computing with defects," *Proceedings of the National Academy of Sciences*, vol. 107, no. 19, pp. 8513–8518, 2010.

[64] G. Zhang, Y. Cheng, J.-P. Chou, and A. Gali, "Material platforms for defect qubits and single-photon emitters," *Applied Physics Reviews*, vol. 7, no. 3, p. 031308, 2020.

[65] C. Chakraborty, C. J. Ciccarino, and P. Narang, "Dynamic modulation of phonon-assisted transitions in quantum defects in monolayer transition-metal dichalcogenide semiconductors," 2020.

[66] T. Olsen, "Designing in-plane heterostructures of quantum spin hall insulators from first principles: 1t'-mos_2 with adsorbates," *Phys. Rev. B*, vol. 94, p. 235106, Dec 2016.

[67] S. Altmann, *Rotations, Quaternions, and Double Groups.* Dover books on mathematics, Dover Publications, 2005.

[68] T. Bradley, C. Bradley, and A. Cracknell, *The Mathematical Theory of Symmetry in Solids: Representation Theory for Point Groups and Space Groups.* Clarendon Press, 1972.

[69] C. M. Gilardoni, T. Bosma, D. van Hien, F. Hendriks, B. Magnusson, A. Ellison, I. G. Ivanov, N. T. Son, and C. H. van der Wal, "Spin-relaxation times exceeding seconds for color centers with strong spin–orbit coupling in SiC," *New Journal of Physics*, vol. 22, p. 103051, oct 2020.

[70] G. Zhang, Y. Cheng, J.-P. Chou, and A. Gali, "Material platforms for defect qubits and single-photon emitters," *Applied Physics Reviews*, vol. 7, no. 3, p. 031308, 2020.

[71] E. Londero, G. m. H. Thiering, L. Razinkovas, A. Gali, and A. Alkauskas, "Vibrational modes of negatively charged silicon-vacancy centers in diamond from ab initio calculations," *Phys. Rev. B*, vol. 98, p. 035306, Jul 2018.

[72] V. Adamyan and V. Zavalniuk, "Phonons in graphene with point defects," *Journal of Physics: Condensed Matter*, vol. 23, p. 015402, dec 2010.

[73] J. Mahanty, A. A. Maradudin, and G. H. Weiss, "Vibrational Thermodynamic Properties of Lattices with Defects, I: The Linear Lattice*," *Progress of Theoretical Physics*, vol. 20, pp. 369–394, 09 1958.

[74] R. J. Bell, "The dynamics of disordered lattices," *Reports on Progress in Physics*, vol. 35, pp. 1315–1409, sep 1972.

[75] A. Kosevich, *The Crystal Lattice: Phonons, Solitons, Dislocations, Superlattices.* Wiley, 2006.

[76] F. Eriksson, E. Fransson, and P. Erhart, "The hiphive package for the extraction of high-order force constants by machine learning," *Advanced Theory and Simulations*, vol. 2, no. 5, p. 1800184, 2019.

[77] U. Mishra, P. Parikh, and Y.-F. Wu, "Algan/gan hemts-an overview of device operation and applications," *Proceedings of the IEEE*, vol. 60, no. 6, pp. 1022–1031, 2002.

[78] V. I. P. V. N. Abakumov and I. N. Yassievich, *Nonradiative Recombinations in Semiconductors*. North-Holland, 1991.

[79] C. H. Henry and D. V. Lang, "Nonradiative capture and recombination by multiphonon emission in gaas and gap," *Phys. Rev. B*, vol. 15, pp. 989–1016, Jan 1977.

[80] C. H. Henry, "Deep level spectroscopy, low temperature defect motion and nonradiative recombination in gaas and gap," *Journal of Electronic Materials*, vol. 4, pp. 1037–1052, Oct 1975.

[81] K. Huang, A. Rhys, and N. F. Mott, "Theory of light absorption and non-radiative transitions in <i>f</i>-centres," *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, vol. 204, no. 1078, pp. 406–423, 1950.

[82] A. M. Stoneham, *Theory of defects in solids: electronic structure of defects in insulators and semiconductors*. Oxford University Press, 2001.

[83] R. Pässler, "Calculation of nonradiative multiphonon capture coefficients and ionization rates for neutral centres according to the static coupling scheme: I. theory," *physica status solidi (b)*, vol. 68, no. 1, pp. 69–79, 1975.

[84] L. Shi and L.-W. Wang, "Ab initio calculations of deep-level carrier nonradiative recombination rates in bulk semiconductors," *Phys. Rev. Lett.*, vol. 109, p. 245501, Dec 2012.

[85] L. Shi, K. Xu, and L.-W. Wang, "Comparative study of ab initio nonradiative recombination rate calculations under different formalisms," *Phys. Rev. B*, vol. 91, p. 205315, May 2015.

[86] Y. Xiao, Z.-W. Wang, L. Shi, X. Jiang, S.-S. Li, and L.-W. Wang, "Anharmonic multi-phonon nonradiative transition: An ab initio calculation approach," *Science China. Physics, Mechanics and Astronomy*, vol. 63, 5 2020.

[87] R. Borrelli, A. Capobianco, and A. Peluso, "Generating function approach to the calculation of spectral band shapes of free-base chlorin including duschinsky and herzberg–teller effects," *The Journal of Physical Chemistry A*, vol. 116, no. 40, pp. 9934–9940, 2012. PMID: 22978703.

135

[88] "Anharmonic lattice relaxation during nonradiative carrier capture," *Physical Review B-Condensed Matter*, vol. 100, July 2019.

[89] D. Wickramaratne, J.-X. Shen, A. Alkauskas, and C. G. Van de Walle, "Comment on "comparative study of ab initio nonradiative recombination rate calculations under different formalisms"," *Phys. Rev. B*, vol. 97, p. 077301, Feb 2018.

[90] L. Shi, K. Xu, and L.-W. Wang, "Reply to "comment on 'comparative study of ab initio nonradiative recombination rate calculations under different formalisms' "," *Phys. Rev. B*, vol. 97, p. 077302, Feb 2018.

[91] M. E. Turiansky, A. Alkauskas, M. Engel, G. Kresse, D. Wickramaratne, J.-X. Shen, C. E. Dreyer, and C. G. Van de Walle, "Nonrad: Computing nonradiative capture coefficients from first principles," *Computer Physics Communications*, vol. 267, p. 108056, 2021.

[92] N. Nepal, J. Li, M. L. Nakarmi, J. Y. Lin, and H. X. Jiang, "Temperature and compositional dependence of the energy band gap of algan alloys," *Applied Physics Letters*, vol. 87, no. 24, p. 242104, 2005.

[93] A. Alkauskas, J. L. Lyons, D. Steiauf, and C. G. Van de Walle, "First-principles calculations of luminescence spectrum line shapes for defects in semiconductors: The example of gan and zno," *Phys. Rev. Lett.*, vol. 109, p. 267401, Dec 2012.

[94] A. Alkauskas, Q. Yan, and C. G. Van de Walle, "First-principles theory of nonradiative carrier capture via multiphonon emission," *Phys. Rev. B*, vol. 90, p. 075202, Aug 2014.

[95] S. Curtarolo, G. L. W. Hart, M. B. Nardelli, N. Mingo, S. Sanvito, and O. Levy, "The high-throughput highway to computational materials design," *Nature Materials*, vol. 12, pp. 191–201, Mar 2013.

[96] N. Marzari, A. Ferretti, and C. Wolverton, "Electronic-structure methods for materials design," *Nature Materials*, vol. 20, pp. 736–749, Jun 2021.

[97] S. Curtarolo, W. Setyawan, G. Hart, M. Jahnatek, R. Chepulskii, R. Taylor, S. Wang, J. Xue, K. Yang, O. Levy, M. Mehl, H. Stokes, D. Demchenko, and D. Morgan, "Aflow: An automatic framework for high-throughput materials discovery," *Computational Materials Science*, vol. 58, pp. 218–226, 08 2013.

[98] M. Hellenbrandt, "The inorganic crystal structure database (icsd)—present and future," *Crystallography Reviews*, vol. 10, no. 1, pp. 17–22, 2004.

[99] A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, and K. a. Persson, "The Materials Project: A materials genome approach to accelerating materials innovation," *APL Materials*, vol. 1, no. 1, p. 011002, 2013.

[100] J. E. Saal, S. Kirklin, M. Aykol, B. Meredig, and C. Wolverton, "Materials design and discovery with high-throughput density functional theory: The open quantum materials database (oqmd)," *JOM*, vol. 65, pp. 1501–1509, Nov 2013.

[101] C. Draxl and M. Scheffler, "The NOMAD laboratory: from data sharing to artificial intelligence," *Journal of Physics: Materials*, vol. 2, p. 036001, may 2019.

[102] S. Haastrup, M. Strange, M. Pandey, T. Deilmann, P. S. Schmidt, N. F. Hinsche, M. N. Gjerding, D. Torelli, P. M. Larsen, A. C. Riis-Jensen, J. Gath, K. W. Jacobsen, J. J. Mortensen, T. Olsen, and K. S. Thygesen, "The computational 2d materials database: high-throughput modeling and discovery of atomically thin crystals," *2D Materials*, vol. 5, p. 042002, sep 2018.

[103] N. Mounet, M. Gibertini, P. Schwaller, D. Campi, A. Merkys, A. Marrazzo, T. Sohier, I. E. Castelli, A. Cepellotti, G. Pizzi, and N. Marzari, "Two-dimensional materials from high-throughput computational exfoliation of experimentally known compounds," *Nature Nanotechnology*, vol. 13, pp. 246–252, Mar 2018.

[104] A. Sajid, K. S. Thygesen, J. R. Reimers, and M. J. Ford, "Edge effects on optically detected magnetic resonance of vacancy defects in hexagonal boron nitride," *Communications Physics*, vol. 3, p. 153, Aug 2020.