Semi-empirical and ab-initio simulation of quantum transport in nanoscale devices

Adel M'Foukh

université PARIS-SACLAY

Semi-empirical and *ab-initio* simulation of quantum transport in nanoscale devices

Simulation semi-empirique et ab-initio du transport quantique dans les composants nanométriques

Thèse de doctorat de l'université Paris-Saclay

École doctorale n°575 : electrical, optical, bio : physics and engineering (EOBE)

Spécialité de doctorat : Electronique, Photonique et Micro-Nanotechnologies

Graduate School : Sciences de l'Ingénierie et des Systèmes

Référent : Faculté des sciences d'Orsay

Thèse préparée dans l'unité de recherche **Centre de Nanosciences et de Nanotechnologies** (Université Paris-Saclay, CNRS), sous la direction de **Marco PALA**, directeur de recherche, et le co-encadrement de **Philippe DOLLFUS**, directeur de recherche

Thèse soutenue à Paris-Saclay, le 13 février 2023, par

Adel M'FOUKH

Composition du jury

Membres du jury avec voix délibérative

Frédéric ANIEL	Président
Professeur des universités, Université Paris Saclay	
Fabienne MICHELINI	Rapporteure & examinatrice
Professeure des universités, Aix-Marseille Université	
Mathieu LUISIER	Rapporteur & examinateur
Professeur, ETH Zurich	
Demetrio LOGOTETA	Examinateur
Chercheur, Université de Rome La Sapienza	

THESE DE DOCTORAT

NNT : 2023UPAST026

université PARIS-SACLAY
ÉCOLE DOCTORALE
Physique et ingénierie : Electrons, Photons, Sciences du vivant (EOBE)

Titre: Simulation semi-empirique et *ab-initio* du transport quantique dans les composants nanométriques

Mots clés: Interaction électron-phonon, Fonctions de Green hors équilibre, Ondes planes, Transport quantique, Théorie de la fonctionnelle de la densité, Pseudo-potentiels

Résumé: La diminution agressive de la taille des transistors menée par l'industrie microélectronique a atteint l'échelle nanométrique. Les performances des composants à cette échelle sont fortement impactées par des phénomènes quantiques (effet tunnel, confinement, interférence,..) mais le contrôle électrostatique devient également plus délicat. Pour gagner en performance, il devient indispensable de prendre ces phénomènes en compte et d'explorer des alternatives au MOSFET à base de silicium qui est la référence actuelle. En support de cette tâche, il est important de mettre en place une approche théorique rigoureuse pour étudier les propriétés électroniques des matériaux et simuler les performances des composants. Dans cette thèse, nous proposons de développer un modèle du transport électronique basé sur le formalisme des fonctions de Green hors équilibre. La particularité de notre approche est l'utilisation de Hamiltoniens de type pseudopotentiels développés sur des ondes planes. Ces Hamiltoniens peuvent être de type soit semi-empirique soit *ab-initio*. Ils permettent d'obtenir une description atomistique du matériau et ainsi d'avoir une description précise à l'échelle microscopique.

Dans un premier temps, nous développons le modèle des Hamiltoniens pseudopotentiels empiriques pour étudier des hétérostructures de matériaux III/V (InAs/GaSb) avec une interface non abrupte. Cette méthode a été appliquée à l'étude de composants basés sur l'effet tunnel, la diode Esaki et le tunnel FET. Les résultats montrent une dégradation du courant lorsque la région de transition est longue. Une comparaison avec une étude expérimentale a permis également de montrer la précision de nos simulations.

Pour la suite, la thèse se concentre sur les Hamiltoniens obtenus à partir de calculs *ab-initio*. Nous avons commencé avec le développement d'un modèle pour étudier les Hamiltoniens des hétérostructures en étudiant deux approches *ab-initio*, une basée sur le Hamiltonien d'un des matériaux isolé et la deuxième basée sur le Hamiltonien d'une super cellule de l'hétérostructure. Les modèles ont été appliquées à deux composants, un MOSFET composé de monocouche et bicouche de $PtSe_2$ et un tunnel FET composé de monocouche $PdSe_2$ et de monocouche SnS_2. Cela a permis de montrer que l'étude de l'interface entre deux matériaux distincts requiert l'utilisation d'approches sophistiquées pour décrire le couplage entre les matériaux.

Pour améliorer le modèle, nous avons ensuite inclus les interactions électron-phonon. Nous avons utilisé d'abord des potentiels de déformation, dérivés de calculs *ab-initio*, pour décrire les couplages avec les phonons sous l'approximation de Born auto-cohérente. Ce modèle a été appliqué à l'étude d'un tunnel FET fait avec l'hétérostructure d'une monocouche $HfSe_2$ et d'une monocouche SnS_2. Deux architectures différentes ont été étudiées, une avec un empilement vertical des matériaux 2D et l'autre avec une hétérostructure latérale. Les résultats montrent que les phonons jouent un rôle important sur les propriétés de transmission en générant une augmentation du courant via une transmission inter-vallée. Enfin, l'hétérostructure verticale ressort comme la candidate la plus prometteuse avec une pente sous le seuil de 40 mV/dec pour un courant à l'état ON de 580 mA/μm à $V_{DD} = 0.35$V.

Finalement, une description plus rigoureuse du couplage électron-phonon pour le transport est proposée. Nous avons développé un modèle utilisant le calcul *ab-initio* des éléments de matrice électron-phonon et des fréquences des phonons. Ce modèle a été comparé à la littérature et à la méthode des potentiels de déformation par l'étude du monocouche MoS_2 avec des calculs de mobilité. Enfin, l'étude de dispositifs CMOS a été réalisée à partir de deux MOSFETs de type n (monocouche MoS_2) et de type p (monocouche WSe_2). Malgré les phonons, les composants satisfont les prédictions pour les prochaines générations de transistor.

université PARIS-SACLAY

ÉCOLE DOCTORALE
Physique et ingénierie : Electrons, Photons, Sciences du vivant (EOBE)

Title: Semi-empirical and *ab-initio* simulation of quantum transport in nanoscale devices

Keywords: Electron-phonon interaction, Non-equilibrium Green's functions, Plane waves, Quantum transport, Density functional theory, Pseudopotentials

Abstract: The aggressive shrinking of transistor sizes by the microelectronics industry has reached the nanoscale. The performance of components at this scale is strongly impacted by quantum phenomena (tunneling, interference, etc.) but also the electrostatic control becomes more delicate. In order to improve the performance of components, it is essential to take these phenomena into account and to explore alternatives to the silicon-based MOSFET that is the current reference. This task can be supported by a rigorous theoretical approach for electronic material properties and thus simulate the performance of the devices. In this thesis, we propose to extend the state of the art in this field by developing an electronic transport model based on the non-equilibrium Green's function formalism. The particularity of our approach is the use of pseudopotential Hamiltonians based on plane waves obtained by either a semi-empirical or an *ab-initio* approach. This type of Hamiltonian allows us to obtain an atomistic description of the material and thus to have an accurate description at the microscopic scale.
In a first part, we developed the empirical pseudopotential Hamiltonian model to study III/V material heterostructures (InAs/GaSb) with a non-abrupt interface. This method has been applied to the study of tunneling based devices, Esaki diode and tunnel FET. The results show a degradation of the current when the transition region is long. A comparison with an experimental study also showed the accuracy of our simulations.
In the following parts of the thesis, the focus is on the Hamiltonians obtained from *ab-initio* calculations. We started with the development of a model to study the Hamiltonians of heterostructure by investigating two approaches, one based on the Hamiltonian of isolated material and the second based on the Hamiltonian of a supercell of the heterostructure. The models were applied to two components, a MOSFET composed of monolayer and bilayer $PtSe_2$ and a tunnel FET composed of monolayer $PdSe_2$ and monolayer SnS_2. This highlighted that the study of the interface between two distinct materials requires the use of sophisticated approaches to describe the coupling between the materials.
To improve the model, we then included electron-phonon interactions in our simulations. We first used deformation potentials, fully derived from *ab-initio* calculations, to describe the coupling with phonons under the self-consistent Born approximation. This model has been applied to the study of a tunnel FET made with the heterostructure of a monolayer $HfSe_2$ and a monolayer SnS_2. Two different architectures were studied, one with a vertical stacking of the 2D materials and a lateral heterostructure. The results show that phonons play an important role for the transmission properties by introducing an increase in current through inter-valley transmission. Finally, the vertical heterostructure emerges as the most promising candidate with a sub-threshold slope of 40 mV/dec for an ON state current of 580 mA/μm at $V_{DD} = 0.35$V.
In the last part of the thesis, a more rigorous description of the electron-phonon coupling for transport was proposed. We have developed a model using *ab-initio* calculations of the electron-phonon matrix elements and phonon frequencies. This model was compared to the literature and to the deformation potential method by studying the monolayer MoS_2 with mobility calculations. Finally, the study of a CMOS was carried out using an n-type (with monolayer MoS_2) and a p-type (with monolayer WSe_2) MOSFET. Despite the phonon scattering, the devices satisfy the predictions for the next generation of transistors.

À mon père

Acknowledgements

Je remercie chaleureusement mon directeur de thèse, Marco Pala, mon co-encadrant, Philippe Dollfus mais également David Esseni pour leur soutien. C'est grâce à eux que j'ai commencé mes travaux de recherche, en commençant par mon stage à Udine. Je n'oublierai pas ce qu'ils m'ont apportés, que ce soit scientifiquement ou humainement, pour me permettre d'atteindre mes objectifs.

Je remercie également tous les membres de l'équipe COMICS avec qui j'ai partagé beaucoup de bon moment. Pour les permanents, Jérôme Saint-Martin, Davide Romanin et Arnaud Bournel avec qui j'ai pu discuter de l'enseignement à l'université. Je remercie également tous les non-permanents qui sont passés ou sont encore dans l'équipe. Junbun Park, Alessandro Pilotto, Remi Helleboid, Mohammad Ghanem, Thibault Cazimajou, Zhizhong Zhang ainsi que tous les stagiaires qui ont partagés le bureau avec moi et m'ont accompagné lors de cette thèse. Merci à tous pour les moments partagés, que ce soit autour d'un café ou au bureau !

Finalement, je remercie mes proches. Je remercie ma mère pour tout ce qu'elle a fait pour moi. Je pense à mes frères et sœur, Emen, Dounia et Gabriel ainsi que mon beau-père Laurent. Je remercie enfin Margaux, ma compagne, avec qui j'ai partagé ces trois dernières années et qui a sue m'aider dans les moments difficiles.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

This chapter introduces the context and the key concepts at stake in the present work. It starts with a focus on the history and evolution of the transistor. We also define the figures of merit used to characterize the devices and necessary to introduce and discuss the limits and trends in this field. Then, we present the general perspectives for logic devices, in terms of material, and the different computational methods developed in order to model such devices. We will conclude by presenting the motivation and the main lines of the present theoretical PhD work.

1.1 Context

The year 2022 marks the 75th anniversary of the transistor. In 1947, John Bardeen, William Shockley and Walter Brattain at the Bell labs invented the first transistor that started the era of microelectronics, which later on was boosted by the invention of microchips and integrated circuits.

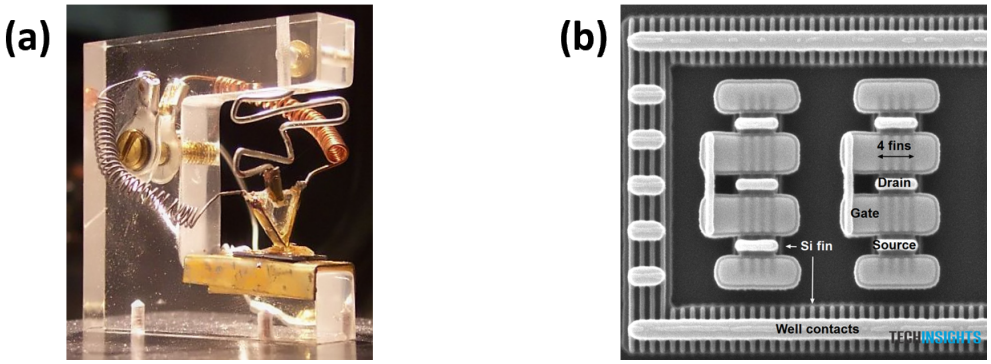

Figure 1.1: (a) Replica of the first transistor exposed at the Heinz Nixdorf MuseumsForum [4] (b) Picture of Samsung 14nm FinFET transistor (source: Samsung 14 nm Exynos 7 7420 Logic Detailed Structural Analysis, TechInsights [1].

This semiconductor device quickly became the most fabricated product in the world as it can be found in all

modern electronic systems where it serves as the main building block. Nowadays, Its major role is to act as a switch for digital logic but it can also be used as an electric signal amplifier. Its success comes from the potential of the technology to be produced at an industrial scale and from the ability to largely shrink its dimensions in order to increase the performance of the microchips.

During the 75 years since its invention, the dimension went from the centimeter large piece of germanium in Fig. 1.1(a) to the silicon based nanoscale device on Fig. 1.1(b). A lot of challenges have been overcome to go from the old device to the recent one. Semiconductor industry made it possible to respect the usually called Moore's law [160]. This law results from a proposition of Gordon Moore, in 1965, that the number of transistors in integrated circuits will double every two years on average. Up to now, this law has been verified by semiconductor industry as shown in Fig. 1.2. Periodically, a roadmap called international roadmap for devices and systems (IRDS) [3] is published by specialists of the field in order to identify the challenges to overcome and to target the next technological nodes to reach. However, this strategy is starting to falter as the technology is reaching the physical limits of the nanoscale transistor.

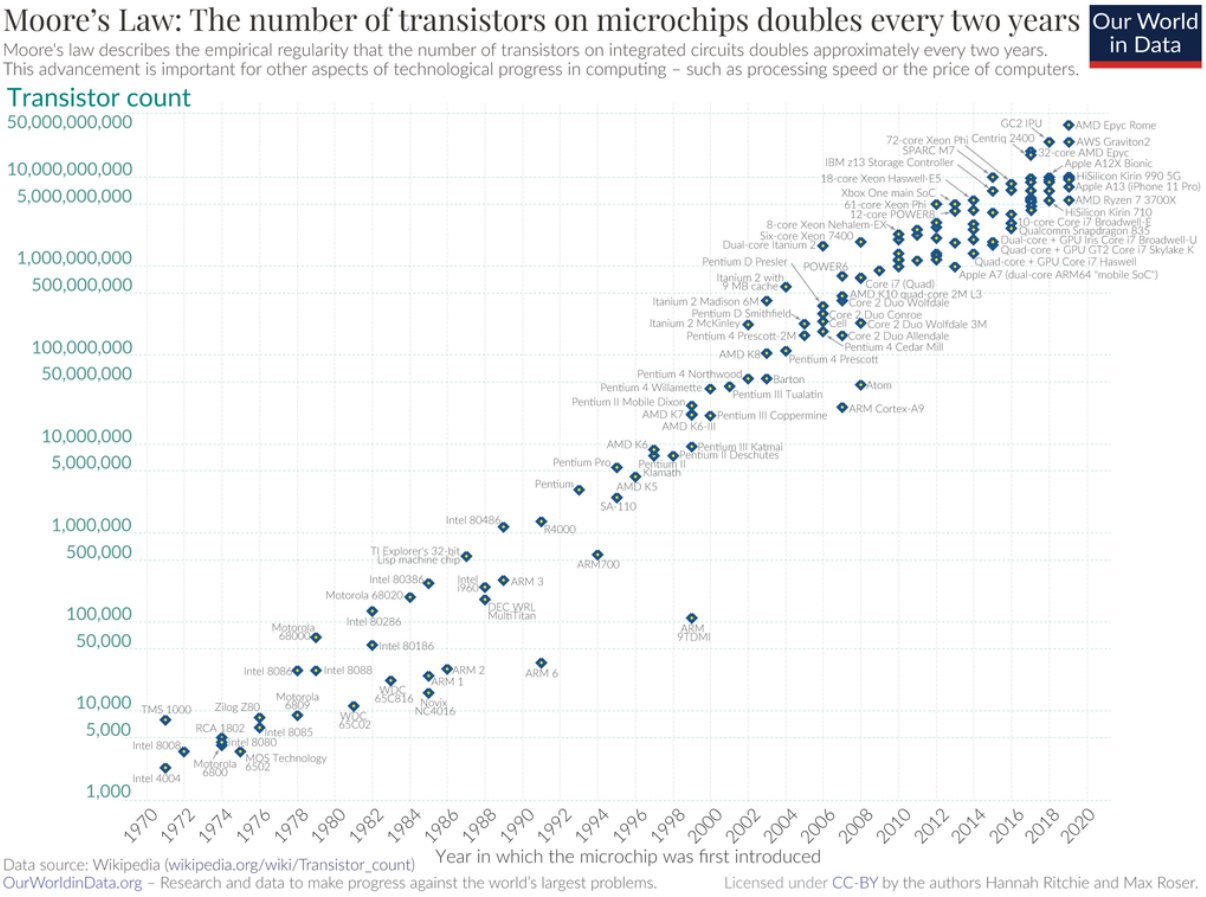

Figure 1.2: Graph presenting the evolution of the number of transistor in microchips over the years from [160].

1.2 MOSFET: Principle

Different kinds of transistor exist, each one with a more or less different working principle, but the most commonly found nowadays is the metal oxide semiconductor field effect transistor (MOSFET) invented in 1956 by Mohamed Atalla and Dawon Kahng [11]. The working principle of the MOSFET is based on the semiconductor nature of the material used to construct it. Semiconductors have the particularity to intrinsically conduct a small amount of current, but under specific circumstances it is possible to increase the conductivity of electrons or holes (electron vacancies in the valence band of the crystal). This is obtained by using an electric field, to add or remove free electrons from a specific region, or chemical doping, that is the implantation of impurities (atoms of a different species) that will add or remove free electrons. A semiconductor with an excess of electrons is qualified as n-doped and p-doped in the opposite situation when it lacks electrons (or has an excess of holes).

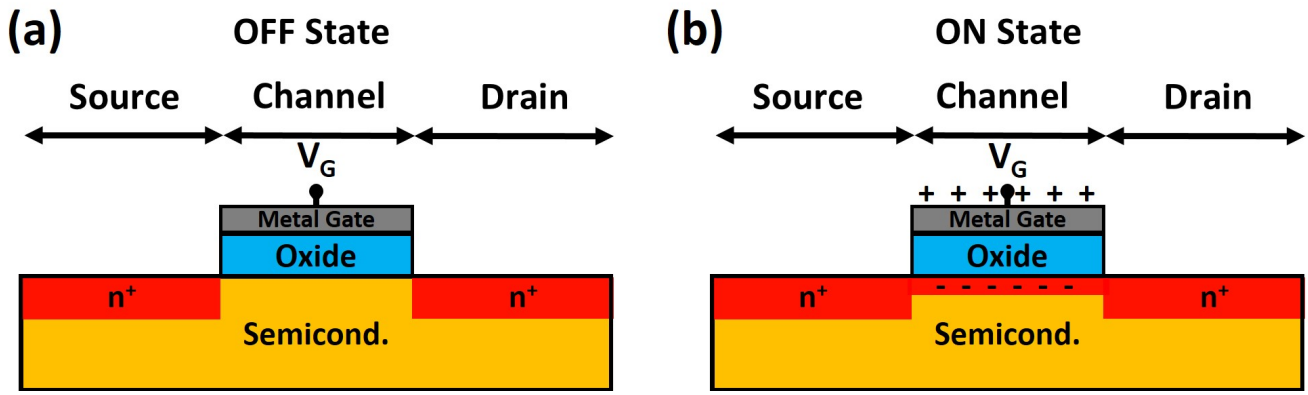

Figure 1.3: Sketch of a n-type MOSFET in OFF-state (a) and ON-state (b) where the gate generate an inversion in the channel to create a path for the current.

In the sketch of Fig. 1.3, the red region corresponds to the electron-doped semiconductor region. The MOSFET is divided into 3 regions, the source, the channel and the drain. An electric field is generated between the source and the drain to drive the charges from the source to the drain. These two regions are also chemically $n(p)$-doped to reduce their resistances and have a better carrier injection/detection from the source/drain contacts. The channel is the region used to switch from an OFF-state to an ON-state by applying a voltage to a metallic gate. An oxide is interposed between the gate and the semiconductor in order to suppress any leaking current. For an n-type MOSFET, an electron current is injected from the n-doped source and collected in the drain region. In the OFF-state, low to no current can flow in the device due to the lack of electron in the central region. This region is controlled by an electric field to populate it with free electrons, which creates an electron current that can now flow from the source to the drain.

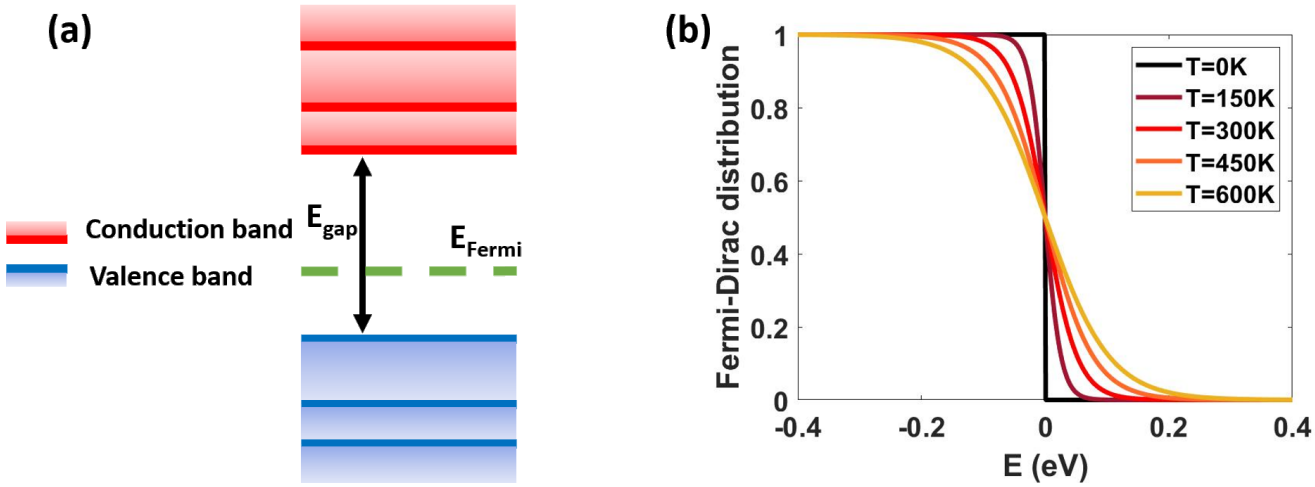

Figure 1.4: (a) Sketch of the bands in a semiconductor. (b) Fermi-Dirac distribution for different temperature.

To get an in depth understanding of the working principle of a MOSFET and semiconductor physics in general it is necessary to use the concept of bandstructure, describing the energy dispersion of electrons in crystals. From quantum physics theory, it is known that an electron linked to an isolated atom can only take some specific energy values. In a crystalline solid, due to the large number of atoms and electrons, the energy values that an electron can take the form of a series of interval of values that are called "bands" as sketched in Fig. 1.4 (a). Since electrons are fermions, these bands are then filled according to the Fermi-Dirac distribution defined by Eq. (1.1) that gives the occupation statistics for an energy E.

$$f(E) = \frac{1}{1 + e^{\frac{E-\mu}{k_B T}}} \tag{1.1}$$

Where μ is the chemical potential, k_B is the Boltzmann constant and T is the temperature. The Fermi-Dirac distribution is illustrated in Fig. 1.4 (b). It can be observed that the increase of temperature spreads the electron distribution. At 0 K, we define the Fermi level as the highest energy value occupied by electrons for metals. In intrinsic semiconductors, it is useful to mention that this Fermi level is located in the energy band gap where there is no electron state. The valence bands correspond to the band filled with electrons (all the bands below the Fermi level) and conduction bands are empty bands where electrons can move freely if we can put electrons on these bands.

In an n-type MOSFET, the current comes from free electrons meaning that they occupy states in the conduction bands. Thus it is interesting to study the conduction bands along the device to understand their transport properties. In Fig. 1.5(a), the profile of the bottom of conduction band along an n-type MOSFET are displayed for different values of V_{GS}. It can be understood that the OFF-state is controlled by the potential barrier present in the channel region resulting from the gate potential, the work function of the metal used and also the difference of doping between channel and source/drain regions. Tuning the gate voltage modulates the height of the barrier to suppress it making the current flow possible between source and drain.

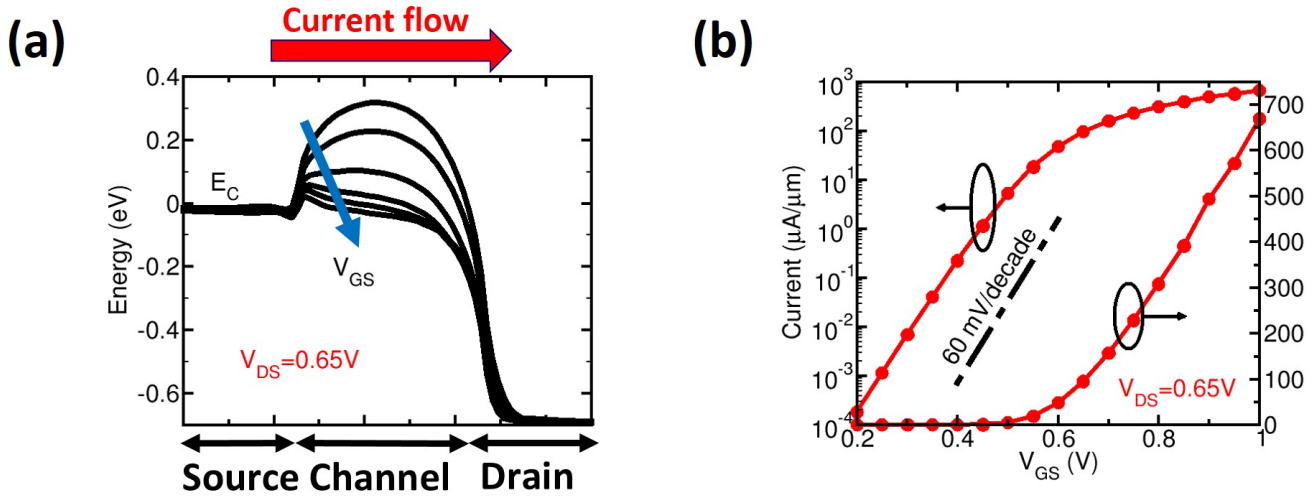

Figure 1.5: (a) Band profile along the transport direction for a n-type MOSFET for different V_{GS}. (b) Transfer characteristics of n-type MOSFET at $V_{DS} = 0.65$ V.

Most of the figures of merit of a transistor can be defined from Fig. 1.5. (b) showing the transfer characteristics of a typical n-type MOSFET. Since it aims at acting as a switch, we can cite the following important figures of merit to look at.

- The ratio between ON and OFF-current: $\frac{I_{ON}}{I_{OFF}}$
- The ON-current: I_{ON}
- the sub-threshold swing: SS

The subthreshold swing of the Current-Voltage (I-V) curve is required to be as steep as possible. SS is defined as follows[193]:

$$\mathrm{SS} = \left[\frac{\partial \log_{10}(I_{DS})}{\partial V_{GS}}\right]^{-1} \tag{1.2}$$

It is evaluated in mV/decade as it represents the swing of voltage that is needed to increase the current by one decade. In the MOSFET, this value is intrinsically limited by the device physics. Indeed, by neglecting quantum phenomena discussed in the next sub-section, since in the OFF-state only electrons belonging to the tail of the Fermi distribution of the source contribute to the thermionic current beyond the barrier, the I-V curve of the MOSFET presents an exponential dependence on the gate voltage. Hence, the lower limit value of the SS is:

$$\mathrm{SS}_{limit} = \frac{k_B T \times \ln 10}{e} = 60\mathrm{mV/dec} \text{ at } T = 300 \text{ K}. \tag{1.3}$$

Where e is the absolute value of elementary electron charge. In Fig. 1.5(b), the exponential growth of the current in sub-threshold regime can be appreciated. As expected the subthreshold swing remains higher than the theoretical limit.

1.3 MOSFET: Limits and challenges

The IRDS, mentioned in Sec. 1.1, observed that the dimensional scaling of silicon MOSFET technology will soon reach some physical limits. In respect of charge transport, two main obstacles arise: short-channel effects and power consumption.

1.3.A Short channel effect (SCE)

Directly linked to the shrinking of the channel length, several phenomena are likely to degrade the performance of the MOSFET [191, 192]. For instance, an important effect of the short channel is the drain-induced barrier lowering (DIBL). It appears when the channel size is comparable to the size of depletion region at the channel/drain junction. It results in a competition between the drain voltage and the gate voltage on the control of the channel barrier height. The higher the drain voltage is, the lower the barrier height will be pushing down. Hence, for small scale devices, the sub-threshold swing is increased as barrier length in the channel is reduced since the control of the gate on the channel is degraded. A solution to limit the effect of DIBL is to reduce to reduce the junction depth in the channel, which can be done by increasing the doping concentration. Another one is to improve the electrostatic control over the channel region. Generally, it is realised by decreasing the oxide thickness. However, one must be cautious with the gate leakage current. This solution has led the semiconductor industry to change the convenient oxide SiO_2 that possesses a dielectric constant of $3.9\epsilon_0$ with high-κ oxides such as HfO_2 having a dielectric constant of $25\epsilon_0$ [201, 159]. They produce a larger gate capacitance with low leakage currents. For example, a layer of 3.2nm of HfO_2 is equivalent to a layer of SiO_2 of 0.5nm. We define the equivalent oxide thickness (EOT) as:

$$\mathrm{EOT} = \frac{\epsilon_{SiO_2} t_{Ox}}{\epsilon_{Ox}} \tag{1.4}$$

A second short-channel effect makes it more and more difficult to achieve a low current in the OFF-state. The origin of this degradation is the source-to-drain tunnelling. The channel reaches a length small enough to allow the carrier to tunnel through, which generates a relatively weak current. The importance of the source to drain tunnelling can be linked to the effective mass of the material used, along the transport direction. For silicon MOSFETs, a channel length below 10 nm is predicted to be the lower limit [206]. One solution is to have a larger effective mass along the transport direction compared to the classical Si MOSFET. It can be done by applying strain to the material or to choose another crystal orientation [123]. Another solution is to use a different material with a larger effective mass than silicon.

1.3.B The power consumption scaling

As the device scales down, the power consumption is expected to follow a similar scaling by decreasing the power supply voltage V_{DD}. The power consumption of a MOSFET corresponds to the contribution from the internal resistance in OFF-state and the contribution of the power required by switching operation. It can be estimated with the following formula:

$$P = I_{OFF} V_{DD} + f_c C_L V_{DD}^2 \tag{1.5}$$

Where f_c is the clock frequency and C_L is the load capacitance, which includes the capacitance from the gate and the wire. The first term is related to the consumption of the transistor in OFF-state and the second one, during a switching operation.

In order to decrease the power consumption, from Eq. (1.5), several strategies are possible. Since reducing the clock frequency is not an option for logic devices, one could decrease the supply voltage. Consequently, in order to reduce the supply voltage, it is necessary that the new device outperform the previous node in terms of figures of merit (SS if the limit of 60 mV/dec is not reached and I_{ON}). For silicon MOSFETs, as the channel length decreases, it can be done by improving the electrostatic control of the channel.

For that two solutions exist. As for solving the DIBL problem, one could reduce the oxide thickness to provide a better electrostatic control of the channel. The second option focuses on gate design. To increase the gate control on the channel, double gate, FinFET (trigate) and gate-all-around (GAA) gate have been proposed[44, 22, 175]. The drawback of this approach is about the thermal management since the thermal dissipation is more difficult in confined architectures. Moreover, as previously explained, the MOSFET subthreshold swing cannot be smaller than 60 mV/dec which means that this scaling of the MOSFET is limited since we need to maintain both high on-current and low off-current.

1.4 Beyond CMOS devices

In the previous section, we mentioned the difficulties that the semiconductor industry has to overcome in order to follow the roadmap proposed by IRDS and presented in Tab.1.1. In this table, two categories of applications are presented, the high performance (HP) logic and the low power or also high density (HD) logic. As seen previously for this device, the voltage supply as well as the subthreshold swing need to be decreased for the next technological nodes while keeping an I_{OFF} at 10 nA/μm for HP and 100 pA/μm for HD and still achieving a ON-current high enough.

Table 1.1: Projected Electrical Specifications of Logic Core Device [3]

Year of production	2022	2025	2028	2031	2034	2037
Power supply voltage V_{DD}	0.7	0.65	0.65	0.6	0.6	0.6
Subthreshold slope(mV/dec) - HP	82	72	70	70	70	70
Subthreshold slope(mV/dec) - HD	75	67	67	65	65	65
I_{ON} (μA/μm) at I_{OFF}=10nA/μm - HP	874	787	851	753	737	753
I_{ON} (μA/μm) at I_{OFF}=100pA/μm - HD	644	602	656	551	532	547

Since the dimensional scaling of silicon MOSFET is no longer a viable solution as we reach its physical limits, next-generation electronics need to be redesigned. Two solutions can be explored in order to achieve this goal. The architecture of the device can be changed to obtain better electrostatic performance similarly to what has been done with the FinFET or by changing the paradigm of the device to overcome the intrinsic limitations of the MOSFET for power scaling. Silicon can also be changed with another material with better properties enabling the continuation of the scaling.

To increase the density of transistors in a chip, the architecture of the device is a key points to be optimised. This problem concerns the ways to overcome the limitation of scaling of the MOSFET. It includes improvements regarding the increased chip density and the enhanced channel control or the adoption of a new device working principle.

1.4.A Architecture evolution of the MOSFET

An idea to improve the MOSFET performance is to increase the gate surface surrounding the channel region. This has led to the adoption of double gate MOSFET[44] and later on the FinFET[22], with one gate that covers three sides of the channel, which was introduced in the market recently. The FinFET marked also the use of 3D architecture. The projected evolution of the MOSFET architecture is presented in Fig. 1.6.

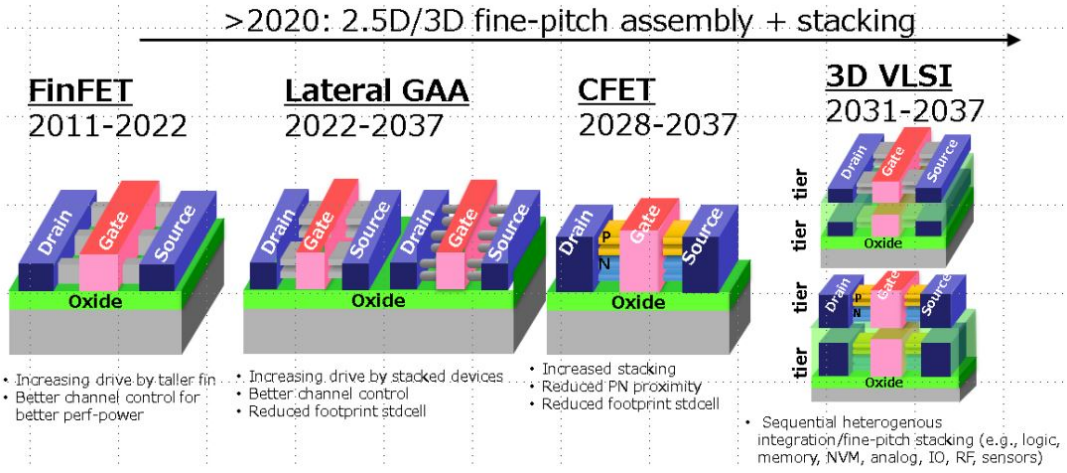

Figure 1.6: Evolution of device architectures in the IRDS More Moore roadmap [3].

The natural evolution after the FinFET is the GAAFET [175, 13] where the gate is now all around the channel region which can be obtained either with a nanowire or a nanosheet configuration. This architecture greatly improves the electrostatic control allowing the reduction of the supply voltage and thus the power consumption. Moreover, due to the confined architecture of these devices, where each of them is isolated from each other, GAA technology allows the devices to be vertically stacked on top of each other [18]. This approach permits the pursuit of the Moore's law as the density of devices per surface unit will continue to increase. Moreover, this open the way to the next step of complementary FET (CFET) [84] where the N and p-type MOSFETs are vertically stacked and then to the 3D very large scale integration (3DVLSI).

As promising these new designs are, there are many problems to solve before their industrial production. From the economical point of view, the process complexity of these devices will make them even more expensive to produce due to the increased number of masks needed for the 3D integration. From a technological point of view, thermal management could become a serious issue in these confined devices in high density. Last but not least, the power scaling is still limited by the working principle of the MOSFET with a subthreshold current that cannot increase faster than 60 mV/dec. The latter point is one of the reasons justifying the increasing research on a sub-thermionic paradigm that we will develop in the next sub-section.

1.4.B Tunnel FET

A special attention is given to the tunnel field effect transistor (TFET) since it is one of the most explored alternatives to replace the MOSFET. The first TFET was made in 1965 [82] but at that time, it was just the firstfruits of the technology. In 2004, a TFET with a sub-thermionic subthreshold slope of 40mV/dec was reported [9]. The strength of the TFET resides in the fact that its SS can theoretically be smaller than the thermionic limit of 60 mV/dec and thus outperform the MOSFET as a low power device. This is of great importance in a time where the power supply scaling is ending.

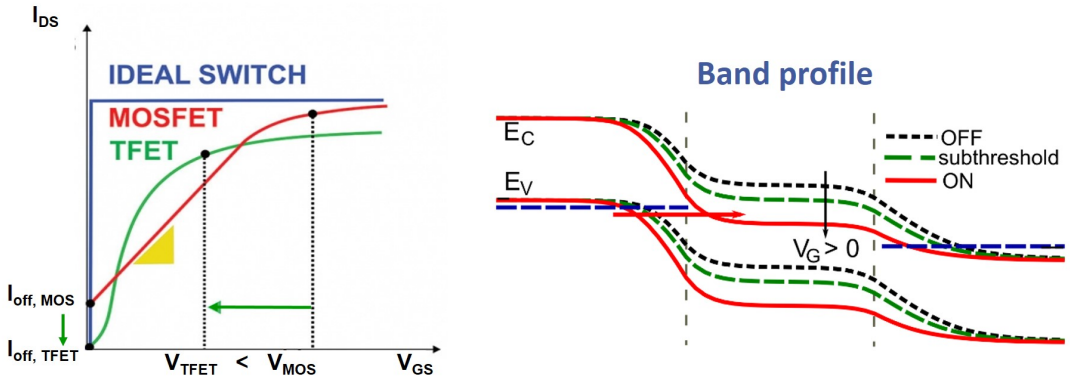

Figure 1.7: (left) I-V curves of the TFET and MOSFET compared to the ideal characteristic. (right) Typical band profile along a TFET device for different gate voltage.

Similarly to the MOSFET, the TFET is controlled by a gate, but the source and the drain have an opposed type of doping, p-i-n for a n-type transistor and n-i-p for the p-type counterpart. The current of the TFET comes from band-to-band tunneling between the valence band to the conduction band as illustrated in Fig. 1.7 (right). In the OFF-states, the bandgap of the channel material filters the electron flow from the valence band of the source to the conduction band of the drain. This mechanism can potentially produce a sub-thermionic subthreshold slope since the electrons belonging to the Fermi tail, which are responsible of the MOSFET SS limit, are filtered by the bandgap of the semiconductor in the source region.

In the left picture of Fig. 1.7, the typical I-V curves of a TFET, a MOSFET and an ideal switch are presented. As previously mentioned, the SS is steeper for the TFET, which allows it to work with a smaller supply voltage. However, the drawback of this device is that the ON-current is smaller than that of MOSFET due to the high resistance related to the tunneling barrier. In order to reduce this resistance and hence to boost the on-state current, it is necessary to consider small gap materials and to highly dope the source region to suppress the depletion within the source/channel junction. This is why the use of a lateral heterostructure may be preferred for this device. Using heterostructures is technically challenging as it requires a clean interface between the two materials due to the sensibility of TFET on interface traps [134] with trap assisted tunneling that can degrade the subthreshold swing.

To summarize, the TFET presents the advantage to potentially solve the power scaling limit of the MOSFET with steep subtheshold slope achievable. However, the low ON-current resulting from the tunnel barrier that the electrons must overcome and the use of a heterostructure with clean interface remains challenging. Further design, technological and material improvement are certainly needed to consider the TFET as a viable candidate in the near future.

1.5 Material aspects

Another approach for future nanoelectronics is to replace the silicon with another material or to introduce new materials in silicon technology. It is a difficult task as silicon is a well known material with many advantages for a channel material and has a strong industrial maturity. This means that in order to be considered for electronic application, the candidate must present much better properties at low cost in what concerns better drive current and electrostatic control. It should be a semiconductor with a bandgap large enough to minimize the OFF-current and it must be compatible with CMOS technology (possibility to grow on silicon with low defect density). The other options that we can exploit is in the use of heterostructures of materials. It may allow us to tune the electronic properties in a device by using different materials in a specific region. An important aspect of heterostructures is the band alignment between the two materials.

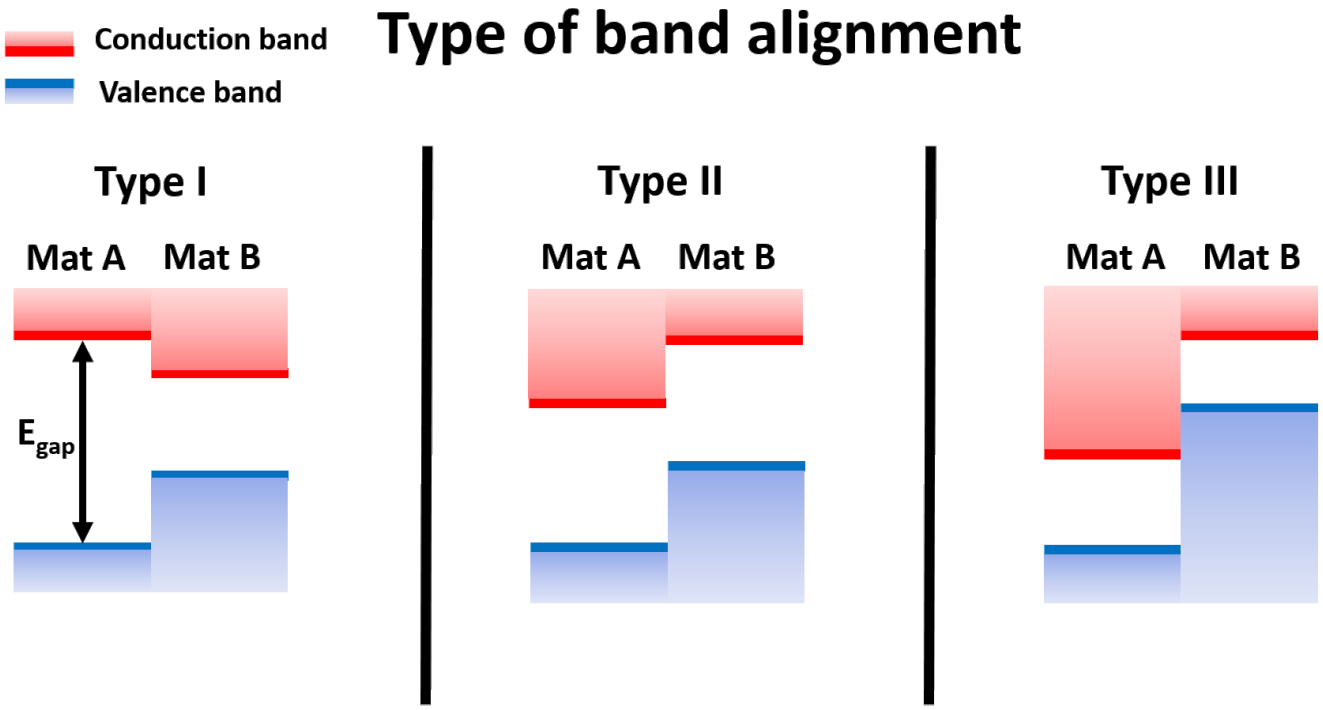

Figure 1.8: Sketch of the different type of band alignment possible in heterostructure.

In Fig. 1.8, we represent the three different kinds of possible band alignment. Type I corresponds to the case where one of the two materials has its conduction and valence band edges in the bandgap of the second material. For the type II, only one band edge (from the conduction or valence bands) of one material is located in the bandgap energy range of the second material. The two first type of band alignments are the most commonly found among materials. The third one, also called broken gap, corresponds to the situation where one material has its conduction band edge lower that the valence band edge of the second material, i.e. with no overlap between the bandgaps. The latter one is the most interesting for tunnel FET, as the tunnel barrier is suppressed in this broken gap configuration, resulting in a higher ON-current.

Two classes of materials are promising candidates to improve TFET performance, the III-V and the 2D materials, in particular because they both offer many possibilities of designing heterostuctures.

1.5.A III-V materials

III-V semiconductors are a well known class of compound materials made of atoms coming from the third and fifth columns of the periodic table of elements (like GaAs,InAs,GaSb,...). They are the second most used semiconductors in the industry after silicon. Indeed, most of them present a direct bandgap in opposition to the indirect bandgap of silicon, which makes them widely used in optoelectronics for laser, LEDs, solar cells and more.

Regarding electron transport properties, they usually present a high electron mobility and a low effective mass

which may lead to a high ON-current needed to increase the frequency of the MOSFET. However, a low effective mass makes the dimensional scaling challenging since the source to drain tunneling current will be difficult to minimize. This identifies the III-V materials as good candidates for high performance transistors only for applications that do not need to scale to the size of current MOSFETs.

However, in the case of tunnel FET, a low effective mass is desired for the charges to go through the tunneling path and to get a high ON-current. It is also important to note that designing heterostructures of III-V compounds with type III configuration is also possible since a couple of them (InAs and GaSb) offer this opportunity to reduce drastically the tunneling path and boost the ON-current. It has been experimentally confirmed that the subthermionic subtheshold swing is achievable with this approach [63, 124, 125]. However, progresses are still needed to have clean interface and materials free of traps.

1.5.B 2D materials

The experimental isolation of graphene, the first two dimensional material ever isolated, in 2004, by Geim and Novoselov [131], started the research field on 2D materials. This class of materials is composed of lamellar materials that are one to a few atoms thick. The research interest for them is growing year after year because of the unique advantages of these materials. Their electronic properties may also vary depending on the number of layer. For instance, we may see a change from direct to indirect bandgap for MoS_2 [53] or a transition from semiconductor to semi-metal for $PtSe_2$ [8].

1.5.B.1 Graphene

Graphene is composed of a one carbon atom thick layer with a honeycomb structure that is shown in Fig. 1.9. Graphene has been extensively studied as it exhibits exotic properties with the presence a bandstructure with Dirac cones, i.e. a linear dispersion, at K points as illustrated in Fig. 1.9. This leads to massless low energy carriers and this induces several interesting properties like a mobility around 100000 cm^2/Vs [122] and intriguing physical phenomena like Klein tunneling that allows a perfect transmission of electrons through a barrier. However, graphene is a semi-metal which disqualifies it as a potential channel material for transistors. Because of that, a lot of energy has been devoted to open a bandgap in graphene.

Among many other ones [85], the most promising approach to do that is to use quantum confinement effect with nanoribbon of graphene. It has been predicted from first-principle calculation that is is possible to open a gap up to 3 eV [204]. Experimental works have confirmed this prediction with measurement of nanoribbon of graphene bandgap of 2.3 eV for a armchair nanoribbon of 7 dimer lines in the width direction [161]. However, every attempt of device fabrication resulted in poor results. In an experimental work, an I_{ON}/I_{OFF} ratio of 3.6×10^3 has been reached with a supply voltage of 1 V [20]. It would be related to the edge quality control leading to a lot of edge-defects. It is also important to mention that this edge disorder results into a strong degradation of mobility[21].

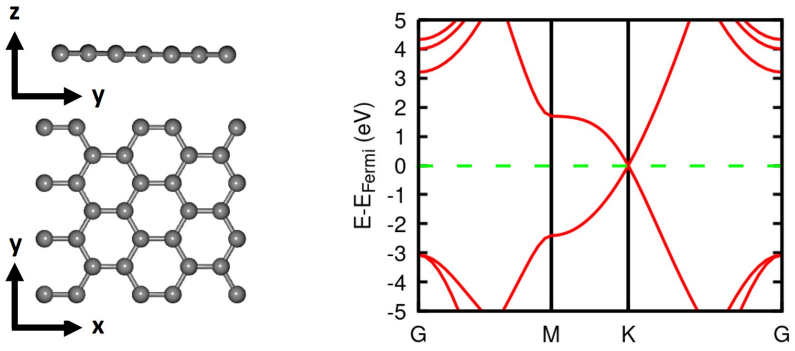

Figure 1.9: (left) Atomic structure of graphene. (right) Electronic band structure computed from the primitive cell of graphene.

1.5.B.2 Transition metal dichalcogenides

Transition metal dichalchogenides (TMDs) is a family of 2D materials that can be exfoliated from their bulk form, where the layers are lightly bonded by Van der Walls interactions. The first isolation of a 2D TMDs has been done in 2005[132]. The unit cell of a monolayer TMD consists in three layers of atoms, a metallic atom layer sandwiched between two chalcogene atom-layers. The most stable phases found in TMDs are the 2H phase where the metallic atom is bonded to three chalcogene atoms (monolayer MoS_2,WS_2, etc..) and the 1T phase where the metallic atom is bonded to six chalcogene atoms (monolayer $HfSe_2$, $PtSe_2$,...). However, it is possible to have several stable phases for the same TMD. For example, monolayer MoS_2 is stable for the 2H phase as well as the 1T [186]. Similarly, monolayer $MoTe_2$ [199] is stable for the phases 1T' and 2H. The electrical properties are impacted by the phase of the material. In monolayer MoS_2, the 2H phase is semiconducting while the 1T phase is metallic. A representation of the atomic structure of the 2H and 1T phases is shown in Fig. 1.10.

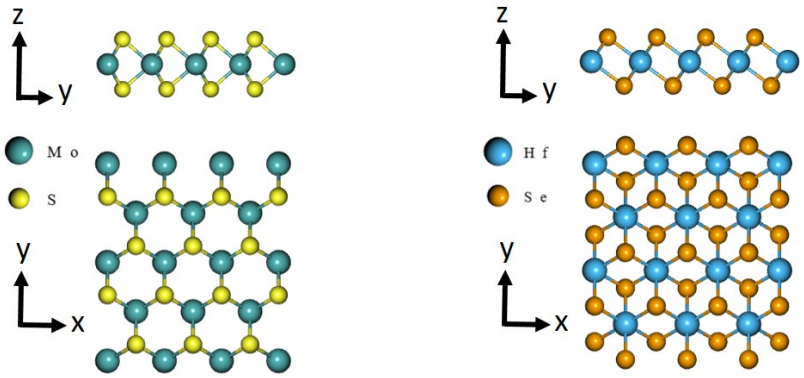

Figure 1.10: Atomic structure of monolayer 2H phase MoS_2 (left) and 1T phase $HfSe_2$ (right).

About their electronics properties, metals, semi-metals and semiconductors can be found in TMDs. It is then natural that the electronics research community focused on the semiconductor TMDs. Surely, the most famous

and studied among them is monolayer MoS_2. Monolayer MoS_2 is a direct band-gap semiconductor with a bandgap estimated to be 1.7 eV from first principle calculation (DFT) and 1.9 eV from experimental measurement of the optical bandgap [120]. Another difference compared to graphene is the finite effective mass found in the semiconductor TMDs which is in general much larger than that of silicon. This could be helpful regarding the channel scaling of MOSFET-like transistors since it would reduce the leakage current for the source to drain tunneling. This properties combined with the excellent electrostatic control of the 2D materials classify them as potential candidates to provide efficient new functionalities if not to replace silicon. Several devices have been fabricated and have shown that it is possible to achieve a I_{ON}/I_{OFF} ratio over 10^8 and a SS of 74 mV/dec [152]. Scaling has also been investigated and monolayer MoS_2 has proven to be able to go further than silicon with a 1nm gate length [49]. The challenges regarding TMDs MOSFET concern the contact region where large contact resistances are usually measured due to metal-induced gap states producing a Schottky barrier [181]. However, recent progresses have been made for n-type, with contact resistance of only 123 $\Omega \cdot \mu$m between monolayer MoS_2 and semi-metal bismuth[171], and p-type with a contact resistance of 3.3 k$\Omega \cdot \mu$m[202].

On the material synthesis aspect, two strategies are possible, the first one is exfoliation. It is from mechanical exfoliation from bulk crystals that the first 2D material has been isolated. However, even if it results in high quality material, the yield is low and the size of the flake limited to some micrometers. The other strategy to produce 2D materials is epitaxial approach with chemical vapor deposition (CVD)[170] or molecular beam epitaxy (MBE)[130, 23]. It is promising as large area sample are produced. Recently, Imec reported a 300mm wafer of WS_2 [10] via a MOCVD process.

1.5.B.3 Lateral and vdW heterostructures

The other major asset of 2D materials is the possibility to make heterostructures of them. A particularity of layered materials is the possibility to conceive vertical heterostructures that present almost defect free interface without dangling bonds. However, it can be subject to momentum mismatch [27] since the layer are bonded by Van der Waals interaction.

Despite the large bandgap and effective mass found in TMDs making them not directly suitable for TFET, as it results in a low ON-current [188], heterostructures of TMDs, on the other hand, present an interest for TFET. Indeed, they form a large family of materials and different types of band alignment are possible. In particular, the broken gap configuration, or a close one, can be achieved, as it can be seen in Fig. 1.11. Theoretical works suggest that using van der Waals (vdW) heterostructures, that present a band inversion between the lowest conduction band and highest valence band under the action of a vertical electric field, could provide a solution to improve the ON-current value in TFET. The band inversion results in the suppression of the bandgap that can allow a high ON-current since there is no tunnelling barrier anymore[188, 27, 6]. This band engineering has been observed recently with the quenching of the bandgap of the bilayer WSe_2 with an electric field[50].

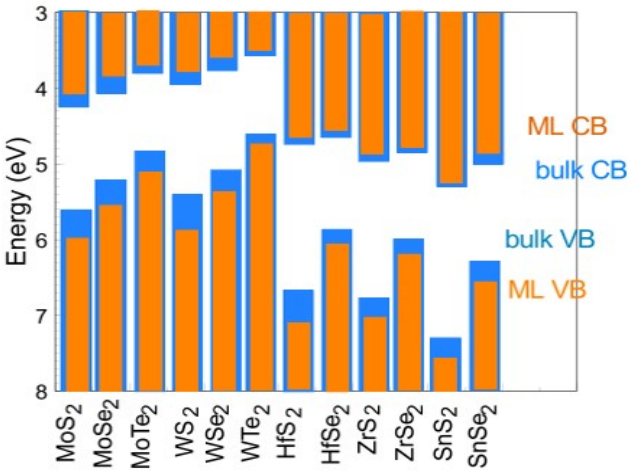

Figure 1.11: Band alignment of TMDs for bulk and monolayer. DFT calculations from ref [73]

About synthesis process, the field is still under development. However, some research groups have already reported results of lateral [106, 189, 105] and Van der Waals heterostructures [109, 60, 205]. For example, Li et al [106] have demonstrated the synthesis of a lateral heterostructure of WSe_2/MoS_2 via epitaxial growth that allows them to achieve a sharp interfaces between the two materials (see Fig. 1.12. (a)). As shown in Fig. 1.12. (b), the group of Lin et al [109] have reported the fabrication via metal-organic chemical vapor deposition (MOCVD) of a Van der Waals heterostructure made of $MoSe_2/WSe_2$.

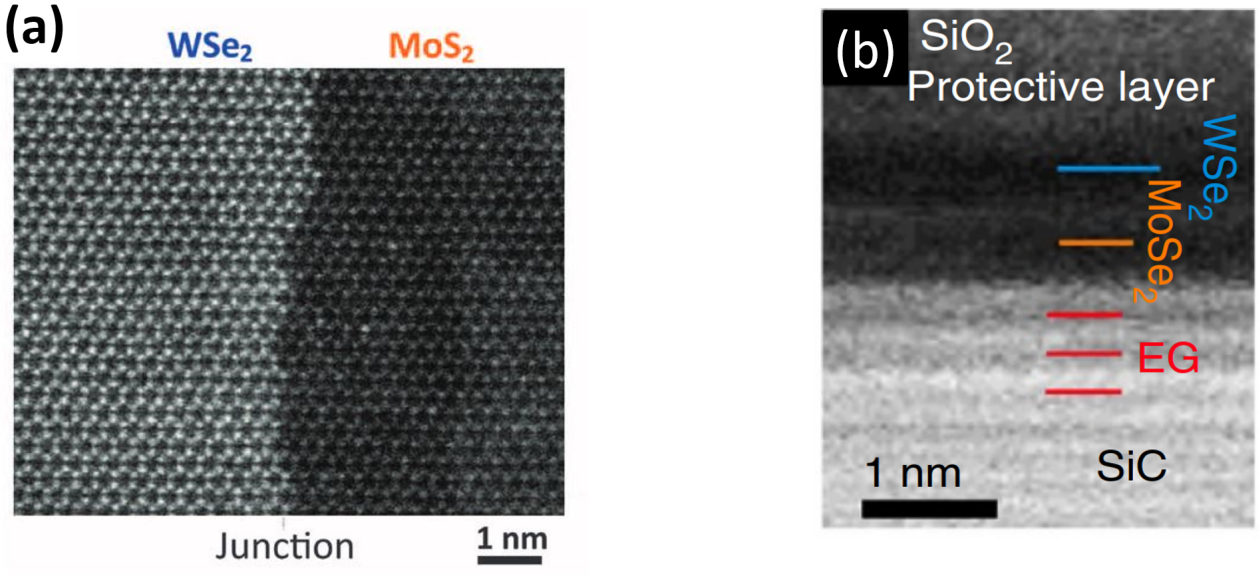

Figure 1.12: (a) High-resolution STEM images taken from the WSe2-MoS2 in-plane heterostructure from [106]. (b) STEM of MoSe2–WSe2–Epitaxial graphene vertical heterostructures from [109]

1.5.B.4 Other 2D materials

Other remarkable 2D materials are of interest for electronics. h-BN, for example, has the same atomic structure as graphene but presents insulator properties with low dielectric constant [101], which makes it a good insulator layer for 2D FET additionally it presents no dangling bonds at the interface and no charge impurity [102]. Another largely investigated material is black Phosphorus, a semiconductor, exhibiting excellent electrical properties with a strongly anisotropic valley in the conduction band, with a low effective mass along the longitudinal transport direction ($0.17m_0$) and a large one along the transverse directions ($1.12m_0$) [197], which produces remarkable MOSFETs with high ON-current [99, 112, 96]. However, it suffers from instability in the ambient air [87]. It is worth to mention that other 2D materials have also been theoretically studied and some have been predicted to present better performance than TMDs for TFET. The study by Szabo et al. mentions SnTe, As, TiNBr, and Bi [188] as they provide ON-current higher than 100 μA/μm with a supply voltage of 0.5 V which has not been achieved with TMDs.

This part concludes the perspectives about electron-based devices. It shows that many ways need to be explored to find the best candidate to replace the current technological node. However, experimental works and manufacturing developments of these proposals will take a lot of time and money. This effort should be reserved for the best candidates among all the propositions. To identify these candidates, modelling and simulations are essential. For device simulations, the problem can be divided in two goals. The first one is the description of the material to properly include all the relevant electronic properties like the electron energy states. The second one concerns the description of electron transport in the device in order to derive its performance. These two goals can be achieved with different degrees of accuracy depending on the system studied and the computational resources available. The different possible methods are presented in the next section.

1.6 Device modeling

The necessity of a theoretical tool to understand the physics of the devices and to perform predictive calculations is more than ever needed for the next generation of electron devices. The calculations should be able to provide accurate information on the electronic transport properties and performance of the device.

1.6.A Electron transport in nano device

For transport calculations in out-of-equilibrium systems, two approaches have been largely used in the study of electron devices. One corresponds to a semi-classical approach and the other to a full quantum description of the electron transport. The choice between both is done by considering the relevant physics of the studied system. For example, to study large systems (>30 nm), the semi-classical methods would be preferred when quantum transport effects are expected to be not significant.

1.6.A.1 Semi-classical methods

The most frequently used method to describe transport in semiconductor devices with a semi-classical approach is based on the Boltzmann transport equation (BTE) formalism. In this equation, carriers are considered as localised point-like particles. The BTE is derived from the local conservation of the number of particles and takes the form of an integro-differential equation on the probability distribution function. This equation also takes into account the interactions experienced by particles by using the Fermi golden rule to derive the corresponding scattering rate. To solve this equation, a traditional particle Monte-Carlo has been frequently and successfuly used and has proven its efficiency to simulate electron devices [88, 90, 55]. This method has been also extended to the quantum transport through the Wigner formalism from which the semi-classical Boltzmann formalism can be derived [151].

1.6.A.2 Full quantum approach

The limit of the previous semi-classical approach is the absence of description of the quantum nature of particles which neglects phenomena like tunnelling or interference. The non-equilibrium Green's function formalism (NEGF) is a powerful formalism derived from many-body physics. It is a full quantum approach that provides a coherent description of quantum transport by solving a set of equations giving access to the Green's functions that describe the dynamical and statistical evolution of the system [5]. The strength of this formalism comes from its generality. It can be used for all kinds of Hamiltonian and it can describe different regimes of transport from the ballistic to the diffusive cases. The simplicity to include external interactions in the system through the concept of self-energy is also one of its advantages. The drawback of such accurate formalism is the heavy computational burden, especially when interactions (e.g. with phonons) are included. However, efficient methods have been developed in order to be able to solve the set of equations of the method applied to devices [183, 117, 137]. In this manuscript, this method is used in every chapter. A special focus on the derivation and application of the method will be done in Chapter 3.

1.6.B Material description

The material description to be used in the transport formalism is provided by the Hamiltonian or the band structure of the material. To access these quantities, one needs to start with the Schrödinger equation associated to the studied system. The total Hamiltonian describing this system includes the kinetic energy of the electrons and nuclei plus the interaction between: electron-electron, electron-nucleus and nucleus-nucleus. To simplify this cumbersome many-body problem, it is assumed that the nuclei are heavier than electrons, which makes it possible to separate the wave function of the ions from the one of the electrons. This approximation is called the Born-Oppenheimer approximation. Knowing that the main goal is the calculation of the electron energy states, the contribution from the nucleus vibrations is neglected in the first instance and the nucleus positions are assumed to be fixed. These vibrations can be re-introduced later as a perturbation to describe the electron-phonon interaction in the transport

formalism. In this case for a system of N electrons in a solid, the equation to solve now reads:

$$[\underbrace{-\frac{\hbar^2}{2m}\sum_i \nabla_i^2}_{\hat{T}} + \underbrace{\sum_{i<j}^{N} V_C(|r_i - r_j|)}_{\hat{V}_{ee}} + \underbrace{\sum_{i,n} U_n(r_i - R_n)}_{\hat{V}_{en}}]\Psi = E\Psi \tag{1.6}$$

Where Ψ is the wave function of the electronic system, R is the position of the nucleus, r is the electron position and m is its mass. Indices i and j refer to electrons and n to nuclei, respectively. In the left part of the equation, the first term $\hat{T}$ corresponds to the kinetic energy of the electrons. In the electron-electron interaction $\hat{V}_{ee}$, V_C is the Coulomb repulsive energy between two electrons and in the electron-nuclei interaction $\hat{V}_{en}$, U_n is the potential energy from the electron-nucleus attraction with the n–th nuclei. Solving this equation is still out of reach since a typical system contains $N \approx 10^{23}$ electrons. Fortunately, the problem can be reduced to a one-electron problem. To achieve this, the Hartree approximation is assumed. The mutual Coulomb interaction between the electrons is simplified by considering a non-interacting electron that moves in a mean electric field arising from all the other electrons. The Schrödinger equation has now the following form:

$$\hat{H}\Psi = E\Psi \equiv [\hat{T} + \hat{V}]\Psi = E\Psi \tag{1.7}$$

Where the potential $\hat{V}$ is the operator that includes the average value of the interaction with all electrons and nuclei, and as such, it possesses the same periodicity as the crystal lattice. $\hat{T}$ is the kinetic energy operator for electrons.

Then the next step consists in estimating this Hamiltonian. Different approaches have been developed to compute it. The most general one is based on first-principles from which a full Hamiltonian is derived. It is free of empirical parameters and can be generalized to most materials. This fundamental approach is called *ab-initio*. It is usually very accurate but also computationally very demanding. The other possibility consists in empirical approaches where ad hoc functions and parameters are used to approximate the Hamiltonian and adjusted to reproduce experimental data. In this case, the accuracy of the model generally depends on the number of empirical parameters used and on the availability of relevant experimental data. The choice among these two kinds of method relies on the accuracy needed, the available computational resources and experimental data.

1.6.B.1 Density functional theory

The Density functional theory (DFT) is the most popular *ab-initio* method used to compute the electronic band structure of materials due to its efficiency and good accuracy. The basis used to describe the DFT Hamiltonian can be either a set of plane wave functions, which is the most natural option for periodic systems, or or a set of atomic orbitals. The latter is better for a chemical bonding description and is computationally more efficient because of to the large number of plane waves needed for the former basis. However, the plane-wave basis provides a better atomic

relaxation as it does not suffer for the Pulay force [149, 162]. Pulay force is a fictitious force that arises when we use the Hellmann-Feynman theorem [78, 57], which provides an expression for the ionic forces at the ground state of the system, to optimise the geometry parameters of our system with a basis set that depends on the position of atomic nuclei. It results from the incompleteness of such basis.

The general idea of DFT is to start with the one-electron Hamiltonian of Eq. (1.7) and to define the potential term as functional of the charge density which is computationally more efficient to solve in comparison with the initial many-body problem. From this expression of the Hamiltonian as a sum of charge density functional, it is possible to implement a self-consistent calculation resulting, at convergence, in the ground-state charge density of the system. Theoretically, the ground-state wave function obtained is exact and the correct band structure of the Hamiltonian should be obtained. However, the exchange-correlation functional, which contains all the many-body interactions, with non-local interaction, is unknown. Many approximations for this functional have been proposed to improve the bandgap estimation. They usually result in an increase of the computational burden since more calculations may be required to describe more accurately the exchange-correlation potential dependency with the charge density. This aspect will be more detailed in Chapter 2.

The DFT is a powerful tool to compute the full band description of the energy dispersion and to have access to an atomistic ground-state wave function of the system which provides an accurate description of the confined states. As an *ab-initio* method, all materials are theoretically accessible and no experimental data are needed. Moreover, it is naturally suited for incorporating more interactions derived from *ab-initio* methods (e.g. electron-phonon interaction). These results come at the cost of a large amount of plane waves needed to converge which produces large Hamiltonian matrices.

On the transport calculation aspect, DFT Hamiltonian has been considered unreachable for a long time due to the large plane-wave basis needed for the convergence. First attempts have used DFT with linear combination of atomic orbitals (LCAO) basis [174] and more recently Bloch function basis [137].

To make calculations easier and achievable in a reasonable time with reasonable resources, simplified empirical approaches have been developed. All of them are based on a specific simplification of the Hamiltonian including some empirical parameters adjusted to fit either *ab-initio* or experimental data. We now describe the most common of these approaches in the following sub-sections.

1.6.B.2 Tight-binding Hamiltonian

A more computationally efficient alternative of the DFT can be found in the tight-binding method, also called LCAO method, which also provides an atomistic approach to study the band structure of materials. To evaluate the Hamiltonian, electrons of the system are first considered tightly bounded to their atom. To represent this, an atomic-like orbitals basis is used to compute the wave function as a linear combination of them. The orbitals chosen are Löwdin orbitals [116], which are constructed to obtain a wave function orthogonal to the other wave functions

centred at different atomic sites in order to simplify the calculation. The coupling between the ions describing the ability of an electron to move between two neighbouring atoms is represented by considering the overlap between the orbitals. The parameters needed to construct this model are the hopping integrals and onsite energy of the isolated ions. These parameters are used to fit and reproduce experimental data. The accuracy of the method is directly linked to its computational efficiency as it depends on the number of orbitals used. The reference for this method is the 10 orbitals model called $sp^3d^5s^*$[89] which provides an accurate description of most diamond-like and zinc-blende-like semiconductors. In the case of a zinc-blende semiconductor with two atoms in its unit cell, it results in a 20×20 matrix if the spin-orbit is not included.

Similarly, to DFT, this method provides a full band description of the energy dispersion. It is attractive because it is computationally efficient as only a few orbitals per atom are usually sufficient to describe the band structure. Moreover, it gives access to an atomistic description of the potential which makes it suitable to describe confined states in nanostructures. Last but not least, An alternative approach to deriving a tight-binding Hamiltonian from *ab-initio* methods, which provides access to all materials, has also been developed. More precisely, DFT calculation with the Hamiltonian projected on a maximally localized Wannier functions (MLWF) basis [182] is used. The tight-binding Hamiltonian parameters are then extracted by post-processing. However, using an *ab-initio* Hamiltonian with a maximally localized Wannier functions basis is computationally more demanding and may reveal delicate to fit.

About the transport calculation aspect, ultra-scaled device simulations have been carried out successfully for silicon [118, 94] with the tight-binding method. It has been extended to many materials by the development of the MLWF tight-binding Hamiltonian. It represents the current state of the art for Hamiltonians used in quantum transport calculation [187, 188].

1.6.B.3 Empirical pseudopotential method

The Empirical pseudopotential method (EPM) is a quite successful method and it is also an alternative to DFT as it is computationally more efficient. It aims to provide an estimation of the atomistic crystal potential by fitting parameters. The calculations are realised on a plane waves basis. It allows to automatically satisfy the periodic boundary conditions at the border of the unit cell which insures that a correct crystal wave function is obtained if a large number of plane waves is used. To solve the problem of the number of plane waves, the notion of pseudopotential is introduced [141]. The idea is to cancel out the core electrons of the system, as they contribute weakly to the electric properties. It leads to a smooth varying wave function in the core region. This makes the convergence on plane wave basis possible with much less plane wave elements than the calculation of the Hamiltonian with the full potential, making it computationally more efficient.

In this method, the potential in the one-electron Hamiltonian of Eq. (1.7) is approximated as the sum of the atomic potential located at the corresponding atom position. To evaluate the pseudopotential, the idea of the

method is to expand the atomic potential in Fourier series. Then the form factors, which are the Fourier coefficients of the atomic potential, are adjusted to fit the simulation results against experimental data from photo-emission and reflectivity measurements [43]. The main advantages of the EPM is its surprising accuracy for a low number of vectors in reciprocal space used to fit experimental data. For diamond-like and zinc-blende-like semiconductors, only three and six parameters are needed [42], respectively, for the local EPM. The use of a non-local EPM only adds five more parameters for a diamond-like semiconductor [138, 34, 80]. The Hamiltonian matrix is large compared to other methods due to the number of plane waves in the basis required which is around one hundred plane waves for a zinc-blende semiconductor. It is also important to note that for some materials like Germanium, the valence bands and some gaps at different symmetry points are difficult to accurately determine, due to the angular momentum dependence of the form factors. However, it has been corrected with the non-local EPM [138, 34, 80] which adds only a few more parameters to obtain better density of states and bandgap description.

Moreover, it gives access to a full band description of the energy dispersion and enables a possible access to an atomistic wave function which provides a more accurate description for ultra-scaled devices. The drawback is that the matrix size is large, due to the plane wave basis, compared to the other empirical methods. It also suffers from the same problem as the other empirical methods which are dependent on the availability of experimental measurements. Empirical parameters are only available for some conventional semiconductors and metals, which limits the predictive character of this method for other materials.

For quantum transport calculation, it is still computationally demanding compared to the other empirical methods since a hundred of plane waves are generally required to achieve a correct band structure. However, recently, some attempts tried to compute quantum transport from this approach by using either an efficient calculation framework or/and the mode-space approximation have allowed to reduce the final basis to a few modes making it computationally much more efficient[135, 136]. The method has been used and extended for the study of heterostructure-based devices in the frame of this PhD thesis. This method is presented in details in chapter 2.

1.6.B.4 $\mathbf{k} \cdot \mathbf{p}$ method

An alternative to the atomistic methods described before is the $\mathbf{k} \cdot \mathbf{p}$ method, that aims to provide a simplified formulation of the energy dispersion. It is based on the perturbation theory to propose an extrapolation of the eigenvectors and eigenvalues in the neighbourhood of a high symmetry point of the Brillouin zone. In order to do that, the wave function of the system is taken in the form of Bloch waves defined by Bloch's theorem. In the case of a crystal, consisting in a periodic repetition in all space directions of a unit cell made of a few atoms, the potential is periodic with the same periodicity (see Eq. (1.7)). In this situation, the Bloch's theorem states that it is possible to define a basis for our system made of wave functions called Bloch states that are solution to the Schrödinger equation. These wave functions follow the periodicity of the crystal. Then by considering that the Bloch states of the high symmetry point make a complete set of function, one can express the Bloch states at different wave vectors

from it. It is possible to provide a simple expression for the energy as:

$$E_n(\mathbf{k}) = E_n(\mathbf{k}_0) + \frac{\hbar^2}{2m} \sum_{i,j} \frac{\mathbf{k}_i \mathbf{k}_j}{m_{ij}} \tag{1.8}$$

where n correspond to the band index and m_{ij} is the effective mass tensor expressed as:

$$\frac{1}{m_{ij}} = \delta_{ij} + \frac{2}{m} \sum_{n \neq n'} \frac{\langle u_n(\mathbf{k}_0) | \, p_i \, | u_{n'}(\mathbf{k}_0) \rangle \langle u'_n(\mathbf{k}_0) | \, p_j \, | u_n(\mathbf{k}_0) \rangle}{E_n(\mathbf{k}_0) - E_{n'}(\mathbf{k}_0)} \tag{1.9}$$

where u_k is the periodic term of the Bloch states that possess the same periodicity of the crystal and p_i is the momentum associated to the wave vector $\mathbf{k}_i$. Therefore, the energy at any $\mathbf{k}$ can be evaluated from unknown parameters which are the energy value at the $\mathbf{k}_0$ point and the terms $\langle u_n(\mathbf{k}_0) | \, p_i \, | u_{n'}(\mathbf{k}_0) \rangle$ which are deduced from experimental data.

In general, only a few bands are used, which keeps the number of parameters to adjust to reproduce experimental data manageable. The description provided is generally accurate only around one high symmetry points in the Brillouin zone. It could even be pushed to a full-band description of the band structure. It has been done by different groups [28, 126, 155, 156] for silicon and germanium and GaAs. However, it requires a non-negligible work to find the set of parameters that is valid for all the k points.

The advantages of this method is related to its versatility. Depending on the problem, a full band or a simple description of some bands around a high symmetry point are possible with this method. It is computationally efficient as the matrix size is linked to the number of bands. However, if we need more bands and to describe a larger volume of the Brillouin zone, it becomes difficult to determine the parameters for the $\mathbf{k} \cdot \mathbf{p}$ Hamiltonian. For instance, the full band description represents a difficult task to find the satisfactory set of parameters to describe accurately the bands. It is also important to mention that the atomistic nature and complex form of the wave function that we can find in the previous method is neglected which implies a more complex treatment of the Hamiltonian at ultra-confined scale.

Some quantum transport calculations have been reported by using the $\mathbf{k} \cdot \mathbf{p}$ method which is a good alternative to atomistic methods due to its computational efficiency when several bands are relevant for the transport [172, 47, 14].

1.7 Motivation and outline of the Thesis

In this context, predicting which architectures and which materials should be used for the next-generation electronics is far from simple. Exploring all possible options experimentally would be overpriced. One of the key to this problem is to use modelling and simulation. It provides an understanding over the semiconductor device physics which allows to create some predictive tools capable to estimate the potentially best candidates on which experimentalists and

industrialists should focus. Moreover, with the reach of the nanoscale and the rise of quantum effect in devices, a full quantum approach is crucial to investigate the quantum transport phenomena. Therefore, the development of such models to study the transport in devices is needed. From a material description aspect, a full band description of the energy dispersion is mandatory to describe accurately the transport properties. It allows to take into account the band structure of the new materials that can have a complex band structure with several relevant valleys and coupled bands. Moreover, mechanisms like intervalley transfer can play a key role to evaluate the current and need a full band description of the material to be rigorously accounted for. The $\mathbf{k} \cdot \mathbf{p}$ method can provide the full band description. However, the task required to find the set of parameters to obtain a satisfying band structure is difficult which made it not popular on this aspect. It is also important to mention that quantum phenomena like confinement need to be taken into account for ultra thin devices and the best accuracy is provided by an atomistic description of the wave function. $\mathbf{k}\cdot\mathbf{p}$ method can describe confinement through the enveloppe function approximation [17, 139] that includes a slowly varying confinement potential. However, this last condition is rarely met for ultra-thin devices[157]. All other methods (EPM, tight-binding and DFT) can consider these aspects. The empirical nature of EPM and tight-binding approaches can represent a difficulty for predictive calculation as they have been mainly developed for conventional semiconductors. Moreover, DFT is the most accurate method among them. The resulting ground-state wave function allows a rigorous description of the charge density in the material. Even if a tight-binding Hamiltonian can be derived from DFT calculation with the maximally localized Wannier functions, using the DFT Hamiltonian would represent an interesting alternative as it would be free of parametrization process and a wave function basis should be more accurate as it is not restricted to localised orbitals. For the transport calculation, since quantum phenomena emerge in ultra-scaled devices, the semi classical approach are no longer valid. For this reason, NEGF stands as the state of the art method for transport calculation with a full quantum approach. The main drawback is the computational burden. Thus a rigorous method to simulate electron devices would be to use the NEGF formalism coupled with DFT Hamiltonian. The main challenge is then to deal with the computational efficiency.

The purpose of this thesis is to continue to develop a full quantum transport model for carriers in devices. The model starts with the material which needs to be accurately described at an atomistic level. With this purpose, we focus on the use of atomistic Hamiltonians based on pseudopotential from the local empirical pseudopotential method and the density functional theory. These two approaches allow an atomistic description of the potential inside the material, which results in a full band description of the materials. This level of accuracy on the material description is needed for the computation of transport properties. The calculation of these properties are based on the state of the art formalism, the Non-equilibrium Green's function. It is a full quantum approach that provides a coherent description of the out-of-equilibrium quantum transport with the possibility to incorporate external perturbations to the system. Until some years, these two methods were not used together for computational reasons. However, the use of a reduced basis developed recently made it possible to do so. The combination of an atomistic Hamiltonian with the NEGF formalism would represent the state of the art of device simulation and is likely to provide an accurate

predictive tool for next-generation electronics.

After this introduction on the current state of electron devices and simulation methods, the two next chapters are focused on the theoretical tools used during this PhD work. We focus on the derivation of the atomic Hamiltonian in the chapter 2. Empirical pseudopotential method and the density functional theory are presented. Then, quantum transport theory is introduced, in chapter 3, with the description of the NEGF formalism. The kinetic equations are derived and a particular attention is given to the definition of the reduced basis in order to use the pseudopotential Hamiltonian.

After the theoretical chapters, the results obtained during this PhD are presented and analysed. The chapter 4 focuses on the development of the quantum transport model with empirical pseudopotential method. A model to study a specific III-V heterostructure, made of GaSb and InAs, which offers a broken gap configuration. The possibility to consider a non-abrupt interface has been introduced in the model. It is applied to two tunnelling devices, the Esaki diode and both n- and p-type ultra-thin body, heterojunction Tunnel FETs. For the latter, a method to include a local quantum confinement operator has also been used. The devices are then studied for different lengths of the transition region between the two materials of the heterostructure.

Chapter 5 is centred around the same thematic of heterostructure but focuses now on DFT Hamiltonian. Indeed, EPM parameters are not yet available for most 2D materials, that are of interest for electronics. It is then needed to move into an *ab-initio* method which is DFT in our case. This study presents the development of the description of lateral heterostructure of 2D materials. The developed model is explained and then compared to the state of the art methodology. For the comparison, two different heterostructures are used. The first is composed of the same materials but with a different number of layers in the different regions to study a MOSFET. The second is realised with two completely different but lattice-matched 2D materials (monolayer $PdSe_2$ and monolayer SnS_2) in order to study a TFET.

After the description of the heterostructures, the accuracy of the model is improved by including the electron-phonon interactions in our equations. The interaction is added by first using the deformation potential approximation which is fully derived from *ab-initio* calculation. This is developed in the chapter 6, where a special focus is given to the el-ph interaction impact on TFET performance. The robustness of two TFET architectures of 2D materials, the lateral and vertical heterostructures, are investigated.

Finally, the chapter 7 presents the development we did to improve the description of the electron-phonon interaction within the NEGF formalism. A new model is developed to incorporate electron-phonon matrix elements computed directly from density function perturbation theory. It provides an accurate description of the phonon dispersion as well as the phonon momentum on the electron-phonon coupling which is often neglected in the deformation potential approximation within NEGF. Transport calculations including this formulation of the interaction are then applied to compute the mobility in monolayer MoS_2 and to transistor simulation with an n-type MoS_2 FET and a p-type WSe_2 FET.

Chapter 2

Methods for material description

In this chapter, we present the theoretical tools used to model the materials under investigation. We continue the first developments given in Sec.1.6.B to derive the one-electron Hamiltonian. The generic form of the Hamiltonian operator solves the following Schödinger equation:

$$\hat{H}\Psi = \left[\hat{T} + \hat{V}\right]\Psi = E\Psi \tag{2.1}$$

where the potential $\hat{V}$ is the operator that includes the average value of the Coulomb interaction from all electrons and nuclei and $\hat{T}$ is the kinetic energy operator for electrons.

In this PhD thesis, we have first focused on pseudopotential Hamiltonians as they provide an efficient and accurate way to access an atomistic and full-band description of materials. To derive the Hamiltonian with the pseudopotential method, a plane wave basis is adopted. This choice is justified as the pseudopotential method is a nearly free electron representation. This representation of electrons is opposed to the tight-binding approach were the electrons are bounded to the atoms. The plane wave basis is the natural choice for this kind of representation since they are the eigenvalues of the free electron system. In the nearly free electron approach, the electrons, which are completely delocalized, are perturbed and localized by a periodic potential. Hence, the basis is well suited to describe the delocalized nature of electrons. The periodic potential, $\hat{V}$ in Eq. (2.1), originates from the periodic crystal structure which represents the Bravais lattice (Infinite periodic array of atoms generated by translation operations). The plane wave basis set is defined by:

$$\langle \mathbf{r}|\mathbf{k}+\mathbf{G}\rangle = \frac{1}{\sqrt{\Omega}} e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}} \tag{2.2}$$

where Ω is the volume of the crystal, $\mathbf{G}$ is a reciprocal lattice vector, $\mathbf{k}$ is the wave vector in the first Brillouin zone, which corresponds to the primitive unit cell of the reciprocal lattice, and $\mathbf{r}$ is the real space position. It should be emphasized that these plane waves have the periodicity of the Bravais lattice.

One major advantage of the plane wave basis is that it provides a simple parameters to improve the accuracy of the results with a cut-off energy on the kinetic energy which determines the number of plane waves representing one-particle Kohn-Sham wave functions. The **G** vectors are selected according to the condition:

$$|\mathbf{G}|^2 \leq \frac{2m_0 E_{\text{cut}}}{\hbar^2} \tag{2.3}$$

The choice of the **G** with the cut-off energy represent a truncation of the plane wave basis. Increasing the cut-off energy make it possible to increase the completeness of the basis. Moreover, it is an orthonormal basis by construction and the Hamiltonian elements have a simple form in this basis which makes the calculation easier. However, describing systems with localized function requires a large number of plane waves as it is the case with the charge density in the core region of atoms.

2.1 Pseudopotential

The notion of pseudopotential is the first step to derive the EPM Hamiltonian. The model used for the atoms in a solid within the pseudopotential method is represented in Fig. 2.1. In this model, the atoms are separated between core and valence electrons. In the core, we find the nucleus plus the core electrons which occupy the lowest energy levels of the atoms and are tightly bounds to their atom. The valence electrons, on the other hand, are responsible of the binding interaction between atoms.

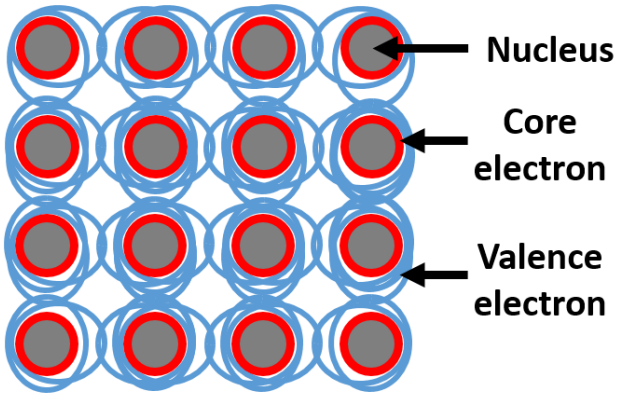

Figure 2.1: Sketch of the atomic model of a solid where the core and valence electrons are distinguished.

The idea of the pseudopotential model is to cancel out the core electrons of our system, as they contribute weakly to the electric properties. Cancelling them leads to a smoothly varying wave function in the core region. This makes the number of plane waves required to describe accurately the electronic properties much fewer than in the case with the core potential due to the important variation of the wave function in the core of the atom. Hence, it improves the computational efficiency.

One ways to do that was proposed by Phillips and Kleinmann in Ref. [141] where a wave function is constructed to be orthogonal to the cores states. The true wave function Ψ is expressed in terms of a smooth wave function ϕ at the core region by the following equation in order to respect the $\langle \phi_i | \Psi \rangle = 0$ condition:

$$\Psi = \phi - \sum_i \langle \phi_i | \phi \rangle \, \phi_i \tag{2.4}$$

where ϕ_i corresponds to the wave function of a core state i. Injecting Eq.(2.4) into (2.1) leads to:

$$H\phi - \sum_i E_i \phi_i \, \langle \phi_i | \phi \rangle = E\phi - \sum_i E \phi_i \, \langle \phi_i | \phi \rangle \tag{2.5}$$

which can be rewritten as:

$$\left[H + \sum_i (E - E_i) \, |\phi_i\rangle \, \langle \phi_i | \right] \phi = E\phi \tag{2.6}$$

The new Hamiltonian corresponds to a system with a new potential called pseudopotential, different from the true potential due to the second term in the left side of Eq.(2.6). Figure 2.2 represents the kind of eigenvectors obtained with pseudopotentials. The eigenvectors are now constructed to be a smooth wave function in the core region compared to the wave functions with all the electrons taken into account. Outside of the core region, the two wave functions must fit each other. However, the energy eigenvalues remain the same as the initial system ones.

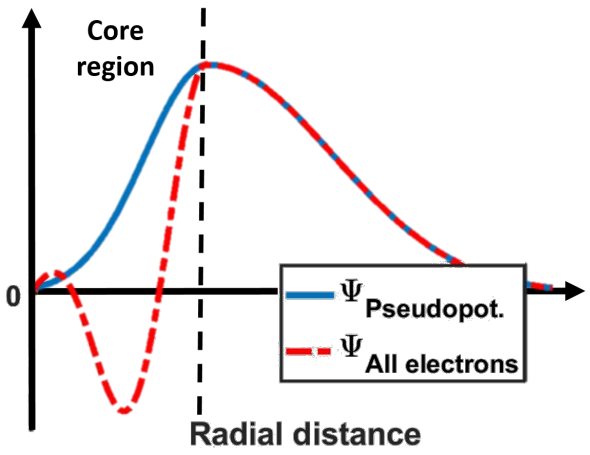

Figure 2.2: Pseudopotential wave function and and all electron wave function obtained by taking the core plus the valence electrons into account.

2.2 Empirical pseudo-potential method

Empirical pseudo-potential method is an approach to estimate the Hamiltonian in a material with a full band description and a few numbers of empirical parameters. It was first developed for metals but it works well also for

semiconductors[33].

2.2.A Local empirical pseudopotential

To evaluate the pseudopotential, we assume that the total potential is equal to the sum of the atomic potential, the pseudopotential is derived as:

$$V(\mathbf{r}) = \sum_{\mathbf{R},\tau} V_a(\mathbf{r} - \mathbf{R} - \tau) \tag{2.7}$$

where $\mathbf{R}$ is the lattice vector and τ is the basis vector. In reciprocal space, it reads:

$$V(\mathbf{r}) = \sum_{G} V_a(\mathbf{G}) S(\mathbf{G}) e^{i\mathbf{G}\cdot\mathbf{r}} \tag{2.8}$$

where $\mathbf{G}$ is a reciprocal lattice vector and S is the structure factor that is expressed as:

$$S(\mathbf{G}) = \frac{1}{N_a} \sum_{\tau} e^{-i\mathbf{G}\cdot\tau} \tag{2.9}$$

with N_a the number of atom in the unit cell. This term describes the crystal structure that depends only on the atom positions. Finally, the input values consist in the atomic form factors $V_a(\mathbf{G})$, which are the Fourier coefficients of the atomic potential that are adjustable parameters used to describe the atomic potential V_a. The results are confronted to experimental data from photo-emission and reflectivity measurement. The second input is the structure factor $S(\mathbf{G})$ used to describe the material lattice.

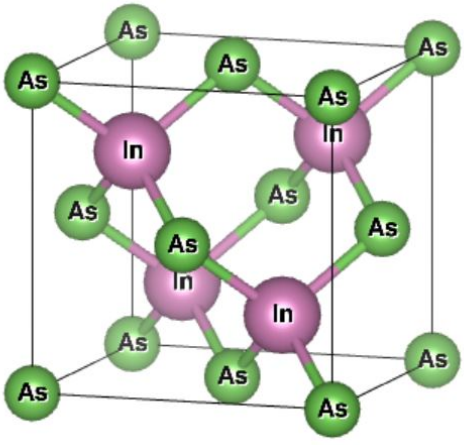

Figure 2.3: Zinc blende crystal structure of InAs.

For a diamond semiconductor (e.g. Si or Ge) or a zinc-blende compound (e.g. GaAs, InAs), the structure, sketched in Fig.2.3, consists in two interpenetrating fcc translated by $(1/4, 1/4, 1/4)\, a_0$, with a_0 the lattice constant.

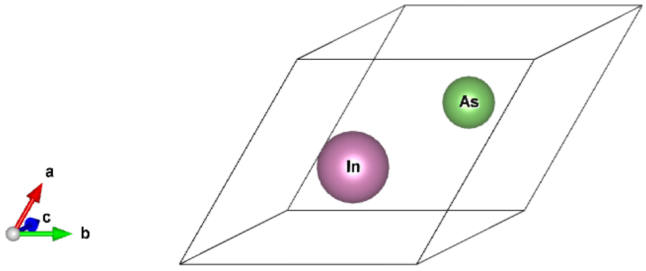

Figure 2.4: Zinc blende unit-cell of InAs.

The unit cell of this system, represented in Fig.2.4, is composed of two atoms and by choosing the origin halfway between the two fcc, the two atomic positions are given by $\boldsymbol{\tau}=\pm(1/8,1/8,1/8)\,a_0$. In this case, the spectral components of the pseudopotential can be written as [33, 42]

$$\begin{aligned} V(\mathbf{G}) &= V_1(\mathbf{G})e^{i\mathbf{G}\cdot\boldsymbol{\tau}} + V_2(\mathbf{G})e^{-i\mathbf{G}\cdot\boldsymbol{\tau}} \\ &= U_S(|\mathbf{G}|)\cos[\mathbf{G}\cdot\boldsymbol{\tau}] + iU_A(|\mathbf{G}|)\sin[\mathbf{G}\cdot\boldsymbol{\tau}] \end{aligned} \tag{2.10}$$

The material dependent $U_S(|\mathbf{G}|)$, $U_A(|\mathbf{G}|)$ denote respectively the symmetric and anti symmetric form factors and V_1 and V_2 are the atomic form factor of atom 1 and 2, respectively. The relation for $U_S(|\mathbf{G}|)$ and $U_A(|\mathbf{G}|)$ are given by:

$$\begin{aligned} U_S(\mathbf{G}) &= \frac{V_1(\mathbf{G}) + V_2(\mathbf{G})}{2} \\ U_A(\mathbf{G}) &= \frac{V_1(\mathbf{G}) - V_2(\mathbf{G})}{2} \end{aligned} \tag{2.11}$$

In the case of diamond materials, $\mathrm{U}_A(|G|)$ is null since the two atoms are identical. The pseudopotential calculated from Eq.(2.10) is presented in Fig.2.5. It is computed by using only three non-null $U_S(|\mathbf{G}|)$ and $U_A(|\mathbf{G}|)$ form factors, corresponding to $|\mathbf{G}|^2 = 3, 8, 11\,[2\pi/a_0]^2$ for $U_S(|\mathbf{G}|)$ and to $|\mathbf{G}|^2$=$3, 4, 11\,[2\pi/a_0]^2$ for $U_A(|\mathbf{G}|)$ (with a_0 being the lattice constant) [33]. The term at $|\mathbf{G}| = 0$ corresponds to a rigid shift in energy of the bands. Generally, it is taken such as the top of valence band corresponds to 0 eV.

With this potential, the full band dispersion of most technologically relevant semiconductors can be well described with EPM [33] as illustrated by Fig. 2.6 for Si and InAs. These two band structures are obtained with only 800 plane waves thanks to the smooth wave functions given by the pseudopotential Hamiltonian. For comparison, an *ab-initio* calculations would require more than ten of thousands plane waves to reach the convergence of the total energy of the system.

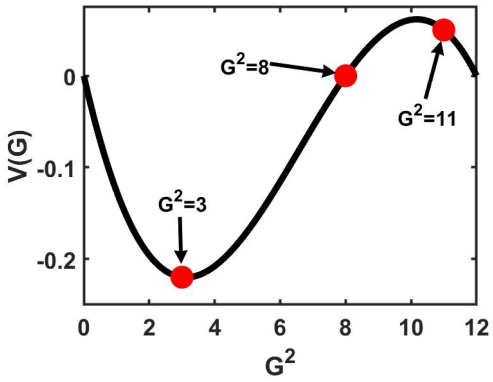

Figure 2.5: Pseudopotential in reciprocal space for InAs.

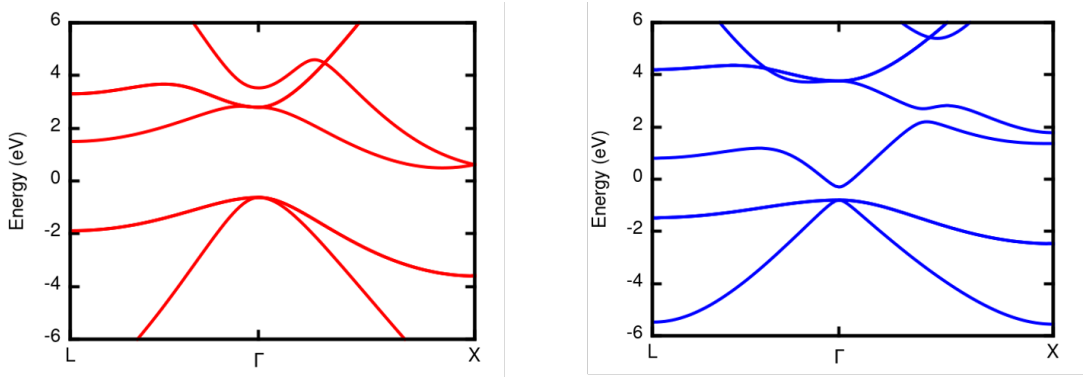

Figure 2.6: Band structure of (left) Si and (right) InAs computed from local empirical pseudopotential with 800 plane waves.

2.2.B Nonlocal pseudopotential correction

Local EPM provide excellent results in regards of electric and optical gaps, however comparison with experimental data highlighted an error in the valence band calculations. Indeed, it was observed that the valence band width is often underestimated. Moreover, for Ge and GaAs, experimental data showed that the band topology and some band gaps at high symmetry points were incorrect. A nonlocal correction is thus proposed to improve the results[79, 138, 34]. It allows us to take into account the variation of the pseudopotential for electrons with different angular momentum l. This is done by expanding the atomic form factor for each angular momentum component of the wave function which gives the new following equation for the potential:

$$V(\mathbf{r}) = \sum_l P_l^\dagger \, V_l(\mathbf{r}) P_l \tag{2.12}$$

where P_l is the projector on the corresponding angular momentum component. This new l-dependent component cannot just be computed from form factors due to its non-local nature. A form of the potential and the radial

symmetry must be assumed, which can be done with a Gaussian or a square well. This adds two new parameters, the well depth and the spatial range. For example, good results [80] are obtained by choosing

$$V_l(r) = \begin{cases} Are^{-\alpha \mathbf{r}} & \text{if } \mathbf{r} \leq \mathbf{R}_s \\ 0 & \text{if } \mathbf{r} > \mathbf{R}_s \end{cases} \tag{2.13}$$

where $\mathbf{R}_s$, is the estimated core radius which is in general around 0.2 Å, α and A are the two new fitting parameters.

On the plane wave basis, the l-th component of the potential depends also on the wave vector k and not just on the reciprocal lattice vector $\mathbf{G}$ due to the non-locality. Because of the radial symmetry, one uses the expansion in spherical harmonics for the plane wave components $|\mathbf{k}+\mathbf{G}\rangle$ and $\left|\mathbf{k}+\mathbf{G'}\right\rangle$. The form factor may be expressed as:

$$\langle \mathbf{k}+\mathbf{G}|\ V_l\left|\mathbf{k}+\mathbf{G'}\right\rangle = \frac{4\pi}{\Omega}(2l+1)P_l \cos\theta_{\mathbf{k}+\mathbf{G},\mathbf{k}+\mathbf{G'}} \int_0^{+\infty} \mathbf{r}^2\, V_l(\mathbf{r}) j_l(|\mathbf{k}+\mathbf{G}|\mathbf{r}) j_l(\left|\mathbf{k}+\mathbf{G'}\right|\mathbf{r})\, d\mathbf{r} \tag{2.14}$$

where Ω is the volume of the unit cell, θ is an angle between the two vectors, P_l is a Legendre polynomial and j_l is a spherical Bessel function.

2.3 Density functional theory

The previous method mentioned is dependent on empirical parameters, to be calibrated to recover experimental data. The possibility to obtain the electronic properties of a material without the use of empirical parameters is attractive in order to make predictive works and support experimentalists. It is the main goal behind the first-principles approach developed during the previous century[83, 97]. Density functional theory (DFT) succeeded in this task and the author Walter Kohn was even awarded a Nobel prize in 1998. The charge density is the core of this theory as its name suggests. It is a physical quantity from which all the ground state physical quantities of the system are derived. From this, DFT approximates a many-body problem of N electrons to the determination of the charge density and possibly its derivatives, depending on the functional used to approximate the exchange-correlation potential, which are physical quantities that depend only on the local position making the method computationally efficient.

2.3.A The Hohenberg-Kohn Theorem

To show that the determination of the ground state charge density is possible and permits the determination of all the electronic properties, two theorems must be exploited. The first one states:

- An external potential applied to a system is a unique functional of the electron density.

This means that if two systems have the same ground state charge density, they have also the same external potential. This can be demonstrated with proof by contradiction by assuming two systems with the same ground state charge

density but different potentials. If we know the corresponding functional, it is possible to determine any potential simply with the charge density. This leads to the same conclusion about the energy of the system which is also a unique functional of the electron density. The second theorem is related to this energy functional.

- The ground-state energy functional of the system minimizes the total energy if and only if the charge density is the ground-state density.

This theorem results directly from the first theorem. Indeed, from the variational theorem, we know that the ground state wave function minimizes the total energy of the system. The first Hohenberg-Kohn theorem tells us that there is a unique ground state electron density corresponding to this wave function, it means that this electron density corresponds to the lowest total energy of the system. This means that a variational principle can be used to determine the ground-state charge density. Indeed, if we consider a Hamiltonian with T, the kinetic energy, U the electron-electron interaction and V is the external potential, the functional of the total energy can be written as:

$$E[n(\mathbf{r})] = \langle\Psi|\ \hat{T} + \hat{U} + \hat{V}\ |\Psi\rangle = \langle\Psi|\ \hat{T} + \hat{U}\ |\Psi\rangle + \langle\Psi|\ \hat{V}\ |\Psi\rangle = F[n(\mathbf{r})] + \int d\mathbf{r}\, V(\mathbf{r}) n(\mathbf{r}) \tag{2.15}$$

In this case, F is the unknown universal functional and E is minimized only if n is the ground state charge density. A variational principle is well defined to determine the ground state charge density.

2.3.B Kohn-Sham equation

Kohn and Sham [83] used these theorems to determine the potential, and thus the Hamiltonian, for the one-electron Hamiltonian described in Sec.1.6.B. The Hamiltonian takes the following form:

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V_{KS}\right]\Psi_i = E_i\Psi_i \tag{2.16}$$

where i runs over the orbitals of the system,Ψ_i is the wave function of the i-th orbital and V_{KS} is the unknown potential called Kohn-Sham potential. For this system, the charge density is computed from the wave functions as:

$$n(\mathbf{r}) = \sum_i |\Psi_i(\mathbf{r})|^2 \tag{2.17}$$

From the first Hohenberg-Kohn theorem, we know that there is a unique functional of V_{KS} that leads to $n(\mathbf{r})$ as its ground state charge density. However, this functional has to be determined and for that the variational principle of the second theorem is used in order to minimize the total energy. The KS potential can be written as:

$$V_{KS} = V_H[n] + V_{xc}[n] + V_{PP} \tag{2.18}$$

where V_{PP} includes the *ab initio* pseudopotential representing the ion-electron interaction and is expressed like Eq.(2.7). V_H is the Hartree potential representing the electrostatic interaction of the electron with the mean field resulting from the repulsive Coulomb interaction from the other electrons. It is expressed as:

$$V_H(\mathbf{r}) = \frac{e^2}{2\pi\epsilon_0} \int d\mathbf{r}' \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \tag{2.19}$$

The second term of Eq.(2.19), V_{xc} corresponds to the exchange-correlation term. It contains all the many-body interaction effects, with non-local interactions. Its functional is unknown but some approximations have proven to be efficient. Kohn and Sham have proposed the local density approximation (LDA) where the functional is approximated by a function of the charge density. This assumes that the density is varying smoothly locally with respect to the real space. The functional in this case is:

$$E_{xc}^{LDA}[n(\mathbf{r})] = \int d\mathbf{r} \epsilon_{xc}^{LDA}[n(\mathbf{r})] n(\mathbf{r}) \tag{2.20}$$

where ϵ_{xc}^{LDA} is the exchange-correlation energy per particle and it is a functional of the charge density. This quantity has been parameterized to interpolate a precise quantum Monte Carlo calculation on a homogeneous electron gas [30]. A logical evolution of the functional is the generalized gradient approximation (GGA) where the gradient of the charge density is added to approximate the functional[140].

With the expression of the functional and from the two theorems of Hohenberg-Kohn, it is possible to define a self-consistent loop in order to minimize the total energy value to find the ground state charge density which will gives us the Kohn-Sham potential.

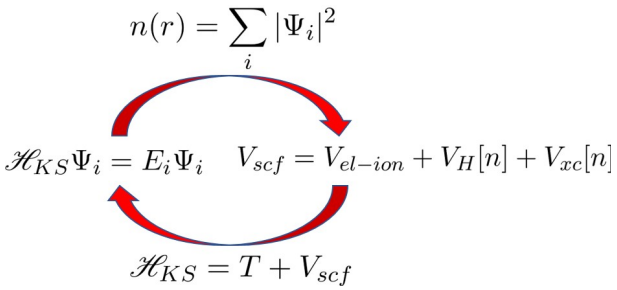

Figure 2.7: Scheme of the DFT self consistant loop.

This loop is sketched in Fig. 2.7. It starts with an initial guess of the charge density. From this guess, the Kohn-Sham potential is evaluated and used to compute the Kohn-Sham Hamiltonian. Solving the Schrödinger equation provides the wave functions of the system from which we compute a new charge density. The new charge density is then compared to the previous one. The loop continues until the difference between the total energy of the two last calculations is smaller than a chosen threshold.

2.3.C Density functional perturbation theory

Density functional perturbation theory (DFPT) [15, 16] is an extension of DFT to linear response calculation from a perturbation. It was developed to study the phonons, the lattice vibration, and thermal properties of materials. This perturbation takes the form of the displacement from equilibrium of ions in the system. For a perturbation of momentum $\mathbf{q}$, the displacement can be written as:

$$\mathbf{U}^{\nu}_{\mathbf{q}}(\mathbf{R}) = \mathbf{U}^{\nu}_{\mathbf{q}}(\mathbf{u}^{\nu}_1, ..., \mathbf{u}^{\nu}_{N_A}) e^{i\mathbf{q}\cdot\mathbf{R}} \tag{2.21}$$

where $\mathbf{R}$ is the lattice vector containing the position of each atom, N_A is the number of atoms, $\mathbf{u}^{\nu}_{\alpha}$ is the displacement pattern for the atom α and $\mathbf{U}^{\nu}_{\mathbf{q}}$ is the amplitude. These atomic displacements can be regrouped in 3 N_A phonon modes and ν is the label of these modes. If no symmetry is found, one can take the displacement along the Cartesian direction for each atom.

The displacement produces a variation of the total energy. For this system, the total energy corresponds to the full system where the ion-ion interaction is taken into account. The corresponding Hamiltonian is:

$$H_R(\mathbf{r}) = -\frac{\hbar^2}{2m}\sum_i \nabla_i^2 + V_H[n(\mathbf{r})] + V_{xc}[n(\mathbf{r})] + V(\mathbf{r}-\mathbf{R}) + V_{ion-ion}(\mathbf{R}) \tag{2.22}$$

where $\mathbf{R}$ is the position of the ions, $\mathbf{r}$ the position of the electrons, V is the ion-electron interaction potential and $V_{ion-ion}$ corresponds to the repulsive Coulomb interaction between the ions. An expansion of the total energy of this system on ion positions writes as:

$$E_{tot} = E_0 + \sum_{\alpha,m} \frac{\partial E}{\partial \mathbf{r}_{\alpha,m}} \cdot \mathbf{u}_{\alpha,m} + \frac{1}{2}\sum_{\alpha,\beta,m,n} \frac{\partial^2 E}{\partial \mathbf{r}_{\alpha,m}\partial \mathbf{r}_{\beta,n}} \cdot \mathbf{u}_{\alpha,m}\mathbf{u}_{\beta,n} ... \tag{2.23}$$

where $\mathbf{r}_{\alpha,m} = \mathbf{R}_m + \mathbf{r}_\alpha$ is the position of the atom α with $\mathbf{R}_m$ the lattice vector for the m-th unit cell and $\mathbf{r}_\alpha$ is the position of the ion in this unit cell. E_0 is the energy calculated with ions in their equilibrium position and $\mathbf{u}$ is the displacement. One can recognize the expression of the forces for the 1st order derivative term since:

$$F = -\frac{\partial E}{\partial \mathbf{r}} \tag{2.24}$$

The derivatives are evaluated at equilibrium, where the sum of all the forces is null. Under the Harmonic approximation that we consider here, the derivatives of order 3 and more are assumed negligible. The last remaining term, the second order derivative of total energy with respect to position, is called the interatomic force constant. Its Fourier

transform is defined by the dynamical matrix whose elements read:

$$D_{\alpha,\beta}(q) = \frac{1}{\sqrt{M_\alpha M_\beta}} \sum_n \frac{\partial^2 E}{\partial \mathbf{r}_{\alpha,0} \partial \mathbf{r}_{\beta,n}} e^{i\mathbf{q}\cdot\mathbf{R}_n} \tag{2.25}$$

with M_α the atomic mass of atom α. The dynamical matrix being hermitian, it can be diagonalised and provides real eigenvalues noted $\omega^2_{q,\nu}$.

$$\sum_\beta D_{\alpha,\beta}(\mathbf{q}) \mathbf{e}^\nu_\beta = \omega^2_{\mathbf{q},\nu} \mathbf{e}^\nu_\alpha \tag{2.26}$$

where ν is the label for the phonon mode and $\mathbf{e}^\nu_\alpha$, the eigenvector of the corresponding phonon mode. One can remark that the form of the total energy corresponds to that of a harmonic oscillator. By analogy with the Newton's law applied to harmonic oscillator in classical mechanics, $\omega_{\mathbf{q},\nu}$ corresponds to the frequency of the phonon. This means that in order to study the phonons, one needs to compute the dynamical matrix. In this order, it is convenient to use the Hellmann-Feynmann theorem [78, 57] which states in this case:

$$\begin{aligned} F = -\frac{\partial E}{\partial \mathbf{R}} &= -\langle \Psi_\mathbf{R}(\mathbf{r}) | \frac{\partial H_\mathbf{R}(\mathbf{r})}{\partial \mathbf{R}} | \Psi_\mathbf{R}(\mathbf{r}) \rangle \\ &= -\int n(\mathbf{r}) \frac{\partial V(\mathbf{r}-\mathbf{R})}{\partial \mathbf{R}} d\mathbf{r} - \frac{\partial E_{ion-ion}(\mathbf{R})}{\partial \mathbf{R}} \end{aligned} \tag{2.27}$$

where $H_R(r)$ is the Hamiltonian from Eq.(2.22), $\Psi_R(r)$ is the corresponding ground state wavefunction, V is the ion-electron interaction potential and $E_{ion-ion}$ corresponds to the energy of the repulsive Coulomb interaction between the ions. Since the interatomic force constant is equal to the derivative of the force with respect to the ion position, one can express it by differentiating Eq.(2.27) to obtain:

$$\frac{\partial^2 E}{\partial \mathbf{R}_\alpha \partial \mathbf{R}_\beta} = \int \frac{\partial n(\mathbf{r})}{\partial \mathbf{R}_\beta} \frac{\partial V(\mathbf{r}-\mathbf{R})}{\partial \mathbf{R}_\alpha} d\mathbf{r} + \delta_{\alpha,\beta} \int n(\mathbf{r}) \frac{\partial V(\mathbf{r}-\mathbf{R})}{\partial \mathbf{R}_\alpha \partial \mathbf{R}_\beta} d\mathbf{r} + \frac{\partial^2 E_{ion-ion}(\mathbf{R})}{\partial \mathbf{R}_\alpha \partial \mathbf{R}_\beta} \tag{2.28}$$

It shows that one can compute the interatomic force constant from charge density and its first derivative. DFPT intervenes in the calculation of the the first derivative of the charge density which is computed from the linearization of Eq.(2.17) and expressed as:

$$\frac{\partial n(\mathbf{r})}{\partial \mathbf{R}_\alpha} = 2\mathrm{Re}\left[\sum_i \Psi^*_i(\mathbf{r}) \frac{\partial \Psi_i(\mathbf{r})}{\partial \mathbf{R}_\alpha} \right] \tag{2.29}$$

To evaluate the 1st derivative of the wave function, we also linearize Eq.(2.16) which takes the form of a Stermheimer equation:

$$[H_{KS} - E_i] \frac{\partial \Psi_i(\mathbf{r})}{\partial \mathbf{R}_\alpha} = -\left[\frac{\partial V_{KS}(\mathbf{r})}{\partial \mathbf{R}_\alpha} - \frac{\partial E_i}{\partial \mathbf{R}_\alpha} \right] \Psi_i(\mathbf{r}) \tag{2.30}$$

where

$$\frac{\partial V_{KS}(\mathbf{r})}{\partial \mathbf{R}_\alpha} = e^2 \int d\mathbf{r}' \frac{\partial n(\mathbf{r}')}{\partial \mathbf{R}_\alpha} \frac{1}{|\mathbf{r}-\mathbf{r}'|} + \int d\mathbf{r}' \frac{\partial V_{xc}}{\partial n(\mathbf{r}')} \frac{\partial n(\mathbf{r}')}{\partial \mathbf{R}_\alpha} + \frac{\partial V(\mathbf{r}-\mathbf{R})}{\partial \mathbf{R}_\alpha} \tag{2.31}$$

and the 1st energy derivative is evaluated from Eq.(2.31) as:

$$\frac{\partial E_i}{\partial \mathbf{R}_\alpha} = \langle \Psi_i(\mathbf{r}) | \frac{\partial V_{KS}(\mathbf{r})}{\partial \mathbf{R}_\alpha} | \Psi_i(\mathbf{r}) \rangle \tag{2.32}$$

Similarly to DFT, Eq. (2.29) and Eq. (2.32) can be used to form a self consistent loop in order to compute the 1st derivative charge density from Eq.(2.29) which is updated from Eq.(2.30) at each loop until convergence. From this, one can compute the interatomic force constant which gives the dynamical matrix with the Fourier transform that describes the phonon properties. Another important quantity that can be derived from these equations is the electron-phonon matrix element. It is defined by using the first derivative over displacement of the Kohn-Sham potential. In the Bloch function basis, it reads

$$g_{m,n,\nu}(\mathbf{k},\mathbf{q}) = \left\langle u_{m,\mathbf{k}+\mathbf{q}} \right| \Delta_{\mathbf{q},\nu} V^{KS} \left| u_{n,\mathbf{k}} \right\rangle_{\text{uc}} \tag{2.33}$$

where $u_{m,\mathbf{k}}$ is the Bloch-periodic components of the Kohn-Sham electron wave functions for the band m and wave vector $\mathbf{k}$, $\Delta_{\mathbf{q},\nu} V^{KS}$ is the variation of the self-consistent potential experienced by the electrons from a phonon of wave vector $\mathbf{q}$ and corresponding to the mode ν and the subscript "uc" means that the integral is computed over one unit cell. Here, the phonon-induced variation of the potential is evaluated from DFPT with Eq.(2.31).

2.3.D Limit and perspectives

DFT and DFPT provide a methodology to compute electronic and thermal properties of materials from first-principles. In solid materials, these results are close to the experimental values. However, the bandgap and other optical properties computed are often incorrectly evaluated. This comes from the approximate functional as DFT provides the exact ground state properties of the system [83]. Hence, it should be able to estimate the correct bandgap energy. To solve this problem, many functionals have been introduced. More advanced functionals named meta-GGA [190, 185] have been developed to include the second derivative of the density in the exchange-correlation functional. We can also cite the hybrid functionals like HSE06 [81], that consist in a mix of the exact exchange energy from Hartree-Fock method with the exchange-correlation energy from GGA. In general, these methods result in a more accurate bandgap at the cost of a heavy computational burden.

Chapter 3

Quantum transport theory

The access to transport properties of materials and devices plays a key role in defining the best options for the next-generation electronics. Due to the low dimensional sizes of novel devices, classical methods like drift diffusion or even semi-classical method like Boltzmann transport equation reach their limit. Indeed, a full quantum approach is needed to accurately model electron transport in these systems due to the importance of quantum phenomena like tunnelling, electron-phase coherence and ballistic transport.

3.1 Non equilibrium Green's functions

The non-equilibrium Green's functions (NEGF) method is the state-of-the-art approach to study quantum transport in physical systems. It was developed in the 60's by Keldysh [92], Kadanoff and Baym [133], Martin and Schwinger [167] and Fujita [61] in order to derive of the out of equilibrium properties of the studied systesm. In the frame of the thesis, it is used to study quantum transport for nanoscale electron devices. In principle, equations of motion derived with NEGF can be solved for any Hamiltonian. The main strength of the method comes from the possibility to add a generic external perturbation through the use of self-energies that we will define later in this chapter. The main challenge comes from the computational resources needed to solve the associated equations limiting the size of the system to a few hundreds or thousands of atoms. However, ingenious strategies were developed to provide an efficient computational scheme to use NEGF with complex Hamiltonians such as the ones derived from DFT.

3.1.A Theoretical definition

In order to define the Green's functions, some concepts and definitions from quantum physics must be recalled first. Since the NEGF method was derived using the many-body perturbation theory, we first introduce the second quantization formalism. For the theoretical development in this section, we consider $\hbar = 1$.

3.1.A.1 Second quantization

The second quantization is a formalism used to study many-body systems composed of a large quantity of identical particles. Since quantum particles are indistinguishable, the wave function of the system must be symmetric, in the case of bosons, or antisymmetric for fermions. This implies that representing the many-body wave function with single state particles, as it is done in the classical representation, would require to symmetrize or antisymmetrize the wave function by considering all permutations which, in the case of many particles is a complex task.

In the second quantization, a Fock space is constructed by summing Hilbert spaces of all particle numbers. States from different Hilbert spaces are assumed to be orthogonal between them. We are then able to define operators that change the number of particles in our system. If we consider a system with $n-1$ particles, the creation operator $\hat{c}_k^\dagger$ will create a particle in state k and it reads:

$$\hat{c}_k^\dagger \left|\phi_1, .., \phi_{k-1}, \phi_{k+1}, .., \phi_n\right\rangle = \left|\phi_1, .., \phi_{k-1}, \phi_k, \phi_{k+1}, .., \phi_n\right\rangle \tag{3.1}$$

where $|\phi_k\rangle$ represent the system with one particle ϕ in the state k. Similarly, the annihilation operation $\hat{c}_k$ which destroys a particle in state k reads:

$$\hat{c}_k \left|\phi_1, .., \phi_{k-1}, \phi_k, \phi_{k+1}, .., \phi_n\right\rangle = \left|\phi_1, .., \phi_{k-1}, \phi_{k+1}, .., \phi_n\right\rangle \tag{3.2}$$

Since we study electrons which are fermions, the wave function must be anti-symmetric. This implies anti-commutation relations between the two operators:

$$\begin{aligned} \{\hat{c}_k \hat{c}_l^\dagger\} =& \hat{c}_k \hat{c}_l^\dagger + \hat{c}_l^\dagger \hat{c}_k = \delta_{k,l} \\ \{\hat{c}_k^\dagger \hat{c}_l^\dagger\} =& 0 \\ \{\hat{c}_k \hat{c}_l\} =& 0 \end{aligned} \tag{3.3}$$

where the braquets $\{\}$ represent the anti-commutator operators. For the Green's function definition, it is more convenient to work on the real space basis. In the second quantization, there is no restriction on the basis needed to define the creation and annihilation operator. Thus, the next step consists in defining the field operators, which are the creation $\psi^\dagger(\mathbf{r})$ and annihilation $\psi(\mathbf{r})$ operator transformed in the real space basis. It creates or annihilates a particle at the position $\mathbf{r}$. It is done with the following definition:

$$\begin{aligned} \hat{\psi}^\dagger(\mathbf{r}) &= \sum_k \phi_k^\dagger(\mathbf{r}) \hat{c}_k^\dagger \\ \hat{\psi}(\mathbf{r}) &= \sum_k \phi_k(\mathbf{r}) \hat{c}_k \end{aligned} \tag{3.4}$$

where $\phi_k(\mathbf{r})$ is the wave function of a fermion in state k in real space and it corresponds to the creation or the annihilation of a particle at the position $\mathbf{r}$.

3.1.A.2 Schrödinger, Heisenberg and interaction representations

In quantum physics, the dynamical evolution of a system can be formulated with three different pictures. The most common is the Schrödinger representation which is based on the time-dependent Schrödinger equation:

$$\frac{\partial \Psi(t)}{\partial t} = -i\hat{H}\Psi(t) \tag{3.5}$$

where Ψ is the wave function of the system. In this representation, operators such as the Hamiltonian $\hat{H}$ are considered to be time independent. It means that the solution to Eq. (3.5) is:

$$\Psi(t) = \Psi(0)e^{-i\hat{H}t} \tag{3.6}$$

The time evolution is then only described by the operator $e^{-i\hat{H}t}$.

In the Heisenberg representation, the situation is opposite. The problem is solved by considering a wave function independent of time. Operators are time-dependent and the time evolution of the operator is given by:

$$\hat{\mathcal{O}}_H(t) = e^{i\hat{H}t}\hat{\mathcal{O}}_H(0)e^{-i\hat{H}t} \tag{3.7}$$

where $\hat{H}$ is the Hamiltonian operator of the system and $\hat{\mathcal{O}}_H$ the operator in the Heisenberg representation. In this representation, wave function and operator are represented with a subscript H. The operator then obeys an equation of motion obtained by differentiating Eq. (3.7) over time:

$$\frac{\partial \hat{\mathcal{O}}_H(t)}{\partial t} = -i[\hat{\mathcal{O}}_H(t), \hat{H}] \tag{3.8}$$

The last representation is the interaction picture which is a mix of both previous representations, where the wave function and operator are both time-dependent. The Hamiltonian H is separated in two parts:

$$\hat{H} = \hat{H}_0 + \hat{V} \tag{3.9}$$

An unperturbed part denoted $\hat{H}_0$ and a perturbation term $\hat{V}$ containing the information about the interactions. In the interaction picture, operator and wave function are noted with a subscript I. In this representation, the time evolution of an operator depends on $\hat{H}_0$ which is time independent, while the wave function depends on the

interaction part $\hat{V}$. It is written as:

$$\begin{aligned}\hat{\mathcal{O}}_I(t) &= e^{i\hat{H}_0 t}\hat{\mathcal{O}}(0)e^{-i\hat{H}_0 t} \\ \Psi_I(t) &= e^{i\hat{H}_0 t}e^{-i\hat{H}t}\Psi_I(0)\end{aligned} \tag{3.10}$$

3.1.A.3 Time evolution operator and contour ordering

The next step is to define a time evolution operator $\hat{U}$, which describes the time evolution of the wave function:

$$\Psi(t) = \hat{U}(t)\Psi(0) \tag{3.11}$$

In the interaction representation, with Eq. (3.10), it is immediate to define $\hat{U}$ as:

$$\hat{U}(t) = e^{i\hat{H}_0 t}e^{-i\hat{H}t} \tag{3.12}$$

By differentiating Eq. (3.12), we obtain the equation of motion of $\hat{U}$:

$$\begin{aligned}\frac{\partial \hat{U}}{\partial t} &= ie^{i\hat{H}_0 t}(\hat{H}_0 - \hat{H})ie^{-i\hat{H}t} \\ &= -ie^{i\hat{H}_0 t}\,\hat{V}\,e^{-i\hat{H}_0 t}e^{i\hat{H}_0 t}e^{-i\hat{H}t} \\ &= -i\,\hat{V}_I\,\hat{U}(t)\end{aligned} \tag{3.13}$$

Integrating this equation leads to an expression of $\hat{U}$ as:

$$\hat{U}(t) = \mathcal{T}e^{-i\int_0^t dt'\,\hat{V}_I(t')} \tag{3.14}$$

where $\mathcal{T}$ is a time-ordering operator that rearranges the operators with decreasing time argument from left to the right. With $\hat{U}$ defined, we can introduce the operator $\hat{S}$:

$$\begin{aligned}\hat{S}(t_1,t_2) &= \hat{U}(t_1)\hat{U}^\dagger(t_2) \\ &= \mathcal{T}e^{-i\int_{t_1}^{t_2} dt'\,\hat{V}_I(t')}\end{aligned} \tag{3.15}$$

It can be shown that $\hat{S}$ obeys the following properties :

$$\begin{aligned}\hat{S}(t_1,t_3) &= \hat{S}(t_1,t_2)\hat{S}(t_2,t_3) \\ \Psi_I(t) &= \hat{S}(t,t')\Psi_I(t')\end{aligned} \tag{3.16}$$

These properties are useful for the Green's function formalism since we face the following problem. The formalism is based on the wave function of the fully interacting system which is not yet known. In the interaction representation, $\hat{H}_0$ is chosen to be simple enough to solve the Schrödinger equation to access Φ_0, the non-interacting ground state wave function. Then the link between the non-interacting wave function to the wave function Ψ, that we need, is given by the Gell-Mann and Low theorem [64] which states:

$$|\Psi_0\rangle = \hat{S}(0,-\infty)\,|\Phi_0\rangle \tag{3.17}$$

This means that the states at the time $t = -\infty$ is the ground state of the non-interacting system which is brought at a time t by adiabatically turning on the interaction. Then we assume that at a large time at $t = +\infty$ the interaction is adiabatically turned off and the system returns to the non-interacting the ground state:

$$|\Phi_0\rangle = \hat{S}(\infty,0)\,|\Psi_0\rangle \tag{3.18}$$

An alternative proposed by Schwinger (1961) is to consider that the time integral is divided into two parts, the first going from $t_0 = -\infty$ to t where the interacting ground state is evaluated, the second part going backward from t to $t_0 = -\infty$ when the system returns to the non-interacting the ground state. This makes a time loop called contour, illustrated in Fig. 3.1 and is composed of two branches that we will name $\mathcal{C}_1$ and $\mathcal{C}_2$. $\mathcal{C}_1$ is going from $t_0 = -\infty$ up to t while $\mathcal{C}_2$ from t to $t_0 = -\infty$.

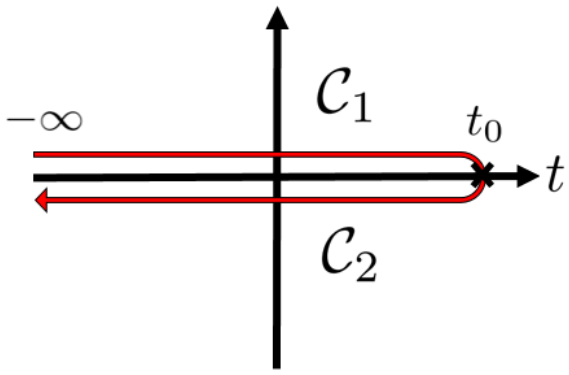

Figure 3.1: Time contour, proposed by Schwinger, represented by a red arrow and where the two branches $\mathcal{C}_1$ and $\mathcal{C}_2$ are identified

3.1.B Green's functions

NEGFs deal with a mathematical object called single-particle Green's function. In this part, the Green's function will be derived for the many-body Hamiltonian (for more details, see [119]) expressed with field operators defined in

Eq. (3.4):

$$H = \int dr\hat{\psi}^\dagger(\mathbf{r})H_0(\mathbf{r})\hat{\psi}(\mathbf{r}) + \frac{1}{2}\int d\mathbf{r}_1 \int d\mathbf{r}_2 \hat{\psi}^\dagger(\mathbf{r}_1,t)\hat{\psi}^\dagger(\mathbf{r}_2,t)v(\mathbf{r}_1-\mathbf{r}_2)\hat{\psi}(\mathbf{r}_2,t)\hat{\psi}(\mathbf{r}_1,t) \tag{3.19}$$

where H_0 is the one-electron Hamiltonian defined in the previous chapter and $v(\mathbf{r}_1-\mathbf{r}_2) = 1/|\mathbf{r}_1-\mathbf{r}_2|$ is the electron-electron interaction. The ground state of this system is noted Ψ_H and $\hat{\psi}$ and $\hat{\psi}^\dagger$ are the creation and annihilation field operators respectively.

3.1.B.1 Green's functions definition

In quantum mechanics, the contour-ordered Green's function is defined in the Heisenberg representation as:

$$\mathcal{G}(\mathbf{r},t,\mathbf{r}',t') = -i\left\langle\Psi_0\right|\mathcal{T}_\mathcal{C}\hat{\psi}_H(\mathbf{r},t)\hat{\psi}_H^\dagger(\mathbf{r}',t')\left|\Psi_0\right\rangle \tag{3.20}$$

where $\mathcal{T}$ is the time-ordering operation and Ψ_0 is the ground state wave function. The time evolution is evaluated on the contour defined previously where t and t' can belong to both $\mathcal{C}_1$ and $\mathcal{C}_2$. However, in order to keep track of the branch where the time is located, we define four different Green's functions.

$$\mathcal{G}(\mathbf{r},t,\mathbf{r}',t') = \begin{cases} \mathcal{G}_C(\mathbf{r},t,\mathbf{r}',t') = -i\left\langle\Psi_0\right|\mathcal{T}\hat{\psi}_H(\mathbf{r},t)\hat{\psi}_H^\dagger(\mathbf{r}',t')\left|\Psi_0\right\rangle \text{ if } t,t'\in\mathcal{C}_1 \\ \mathcal{G}_{\tilde{C}}(\mathbf{r},t,\mathbf{r}',t') = -i\left\langle\Psi_0\right|\mathcal{T}\hat{\psi}_H(\mathbf{r},t)\hat{\psi}_H^\dagger(\mathbf{r}',t')\left|\Psi_0\right\rangle \text{ if } t,t'\in\mathcal{C}_2 \\ \mathcal{G}^<(\mathbf{r},t,\mathbf{r}',t') = i\left\langle\Psi_0\right|\hat{\psi}_H^\dagger(\mathbf{r}',t')\hat{\psi}_H(\mathbf{r},t)\left|\Psi_0\right\rangle \text{ if } t\in\mathcal{C}_1,\ t'\in\mathcal{C}_2 \\ \mathcal{G}^>(\mathbf{r},t,\mathbf{r}',t') = i\left\langle\Psi_0\right|\hat{\psi}_H(\mathbf{r}',t')\hat{\psi}_H^\dagger(\mathbf{r},t)\left|\Psi_0\right\rangle \text{ if } t\in\mathcal{C}_2,\ t'\in\mathcal{C}_1 \end{cases} \tag{3.21}$$

where $\mathcal{G}_C$ and $\mathcal{G}_{\tilde{C}}$ are the causal and anti-causal Green's function respectively and $\mathcal{G}^<$ and $\mathcal{G}^>$ are the lesser than and greater than Green's functions, respectively. The last two Green's functions are useful as they contain the information about the statistical properties of the system like carrier density or current. The first two Green's functions are commonly replaced by two others that are the retarded and advanced Green's functions that provide information about the dynamical properties of the system like transmission, scattering rate and more. These Green's functions are defined as:

$$\begin{aligned} \mathcal{G}^r(\mathbf{r},t,\mathbf{r}',t') &= \mathcal{G}_C(\mathbf{r},t,\mathbf{r}',t') - \mathcal{G}^<(\mathbf{r},t,\mathbf{r}',t') = -i\Theta(t-t')\left\langle\Psi_0\right|\hat{\psi}_H(\mathbf{r},t)\hat{\psi}_H^\dagger(\mathbf{r}',t')\left|\Psi_0\right\rangle \text{ if } t,t'\in\mathcal{C}_1 \\ \mathcal{G}^a(\mathbf{r},t,\mathbf{r}',t') &= \mathcal{G}_C(\mathbf{r},t,\mathbf{r}',t') - \mathcal{G}^>(\mathbf{r},t,\mathbf{r}',t') = -i\Theta(t'-t)\left\langle\Psi_0\right|\hat{\psi}_H(\mathbf{r},t)\hat{\psi}_H^\dagger(\mathbf{r}',t')\left|\Psi_0\right\rangle \text{ if } t,t'\in\mathcal{C}_2 \end{aligned} \tag{3.22}$$

where we define a step function Θ as:

$$\Theta(t) = \begin{cases} 1 \text{ If } t>0 \\ 0 \text{ otherwise} \end{cases} \tag{3.23}$$

with Eq. (3.22), one can show that the four Green's functions obey the following relationship [76]:

$$\mathcal{G}^r - \mathcal{G}^a = \mathcal{G}^> - \mathcal{G}^< \tag{3.24}$$

3.1.B.2 Equation of motion

The time-evolution of the contour-ordered Green's function defined in Eq. (3.20) is evaluated by means of its equation of motion:

$$\begin{aligned}\frac{\partial \mathcal{G}(\mathbf{r},t,\mathbf{r}',t')}{\partial t} &= -i(\langle\Psi_0|\,\mathcal{T}_\mathcal{C}\frac{\hat{\psi}_H(\mathbf{r},t)}{\partial t}\hat{\psi}_H^\dagger(\mathbf{r}',t')\,|\Psi_0\rangle + \delta(t-t')\,\langle\Psi_0|\,\{\hat{\psi}_H(\mathbf{r},t)\hat{\psi}_H^\dagger(\mathbf{r}',t')\}\,|\Psi_0\rangle) \\ &= -i(\langle\Psi_0|\,\mathcal{T}_\mathcal{C}\frac{\partial\hat{\psi}_H(\mathbf{r},t)}{\partial t}\hat{\psi}_H^\dagger(\mathbf{r}',t')\,|\Psi_0\rangle + \delta(t-t')\delta(\mathbf{r}-\mathbf{r}')\end{aligned} \tag{3.25}$$

In order to evaluate the first derivative of the field operator, we use the Hamiltonian from Eq. (3.19). Then, we use Eq. (3.8) and the anticommutation relations of the field operators to obtain:

$$\begin{aligned}\frac{\partial\hat{\psi}_H(\mathbf{r},t)}{\partial t} &= -i[\hat{\psi}_H(\mathbf{r},t),\hat{H}] \\ &= -iH_0(\mathbf{r})\hat{\psi}_H(\mathbf{r},t) - i\int d\mathbf{r}'\hat{\psi}_H^\dagger(\mathbf{r}',t)v(\mathbf{r}-\mathbf{r}')\hat{\psi}_H(\mathbf{r}',t)\hat{\psi}_H(\mathbf{r},t)\end{aligned} \tag{3.26}$$

Injecting Eq. (3.26) into Eq. (3.25), we obtain:

$$\begin{aligned}i\hbar\frac{\partial\mathcal{G}(\mathbf{r},t,\mathbf{r}',t')}{\partial t} &= \delta(t-t')\delta(\mathbf{r}-\mathbf{r}') + H_0(\mathbf{r})\hat{\psi}_H(\mathbf{r},t)\hat{\psi}_H^\dagger(\mathbf{r}',t') - i\int d\mathbf{r}''v(\mathbf{r}-\mathbf{r}'')\hat{\psi}^\dagger(\mathbf{r}'',t'')\hat{\psi}^\dagger(\mathbf{r},t)\psi(\mathbf{r}'',t'')\hat{\psi}(\mathbf{r}',t') \\ &= \delta(t-t')\delta(\mathbf{r}-\mathbf{r}') + H_0(\mathbf{r})\mathcal{G}(\mathbf{r},t,\mathbf{r}',t') - i\int d\mathbf{r}''v(\mathbf{r}_1-\mathbf{r}_2)\mathcal{G}_2(\mathbf{r}'',t'',\mathbf{r},t,\mathbf{r}'',t'',\mathbf{r}',t')\end{aligned} \tag{3.27}$$

where $\mathcal{G}_2$ is the two-particles Green's function. By rearranging Eq. (3.27), we obtain:

$$\left[i\frac{\partial}{\partial t} - H_0(\mathbf{r})\right]\mathcal{G}(\mathbf{r},t,\mathbf{r}',t') + i\int d\mathbf{r}''\int dt''v(\mathbf{r}_1-\mathbf{r}_2)\mathcal{G}_2(\mathbf{r}'',t'',\mathbf{r},t,\mathbf{r}'',t'',\mathbf{r}',t') = \delta(t-t')\delta(\mathbf{r}-\mathbf{r}') \tag{3.28}$$

If we try to get rid of the two-particles Green's function by differentiating it in order to find a simpler formulation, it will only introduce a 3-particles Green's function and similarly with the 3-particles Green's function and so on.

3.1.B.3 Perturbation expansion of the Green's function

To solve that problem, the next step is to develop a perturbation expansion of the Green's function. We start with the Gell-Mann and Low theorem [64] to link the interacting state to the non-interacting one wave functions. Since this has to be done under the interaction representation, we separate the Hamiltonian to take the form of Eq. (3.9), where we move the time-dependent part of the Hamiltonian into $\hat{V}$. By applying that, the contour-ordered Green's

function expression takes the following form:

$$\mathcal{G}(\mathbf{r},t,\mathbf{r}',t') = -i\left\langle \Psi_{H_0}\right| \mathcal{T}_{\mathcal{C}}\{S_{\mathcal{C}}\hat{\psi}_{H_0}(\mathbf{r},t)\hat{\psi}^{\dagger}_{H_0}(\mathbf{r}',t')\}\left|\Psi_{H_0}\right\rangle \tag{3.29}$$

where the time-ordering operator acts to order t on the contour defined and Ψ_{H_0} is the ground state of the non interacting system and where we have:

$$\hat{S}_{\mathcal{C}} = e^{-i\int_{\mathcal{C}} dt'\,\hat{V}_I(t')} \tag{3.30}$$

Then $\hat{S}_{\mathcal{C}}$ is expanded in Taylor series to express the Green's function in term of power of $\hat{V}_I$, since the term $\hat{V}_I$ is expressed in general with field operators, like for the Coulomb interaction. The Green's function is then decomposed in products of quadratic averages by using the Wick theorem [200]. This theorem states that the time ordered product of field operators can be written as a sum over normal ordered products of the field operators where distinct pairs of field operators are contracted in all possible combinations. To apply the theorem, one requires only to have an unperturbed Hamiltonian that is quadratic in term of fermion operators. This makes it possible to express the perturbed Green's function in terms of unperturbed single particle Green's function that are much easier to calculate.

3.1.B.4 Self-energy and Dyson equation

Previously, we saw that with the Wick's theorem, it is possible to express the interacting Green's function as a sum of products of single-particle Green's functions. The method enables the introduction of the self-energy, symbolized by Σ, which is the functional of the single-particle Green functions and Eq. (3.28) takes now the following form:

$$\left[i\frac{\partial}{\partial t_1} - H_0(\mathbf{r})\right]\mathcal{G}(\mathbf{r}_1,t_1,\mathbf{r}_2,t_2) = \delta(t_1-t_2)\delta(\mathbf{r}_1-\mathbf{r}_2) + \int d\mathbf{r}_3\int dt_3\Sigma(\mathbf{r}_1,t_1,\mathbf{r}_3,t_3)\mathcal{G}(\mathbf{r}_3,t_3,\mathbf{r}_2,t_2) \tag{3.31}$$

where the self-energy is defined as:

$$\int d\mathbf{r}_3\int dt_3\ \Sigma(\mathbf{r}_1,t_1,\mathbf{r}_3,t_3)\mathcal{G}(\mathbf{r}_3,t_3,\mathbf{r}_1,t_1) = -i\int d\mathbf{r}_3\int dt_3\ v(\mathbf{r}_1-\mathbf{r}_3)\mathcal{G}_2(\mathbf{r}_1,t_1,\mathbf{r}_3,t_3,\mathbf{r}_2,t_2,\mathbf{r}_3,t_3) \tag{3.32}$$

For an unperturbed system, the Green's function follows:

$$\left[i\frac{\partial}{\partial t} - H_0(\mathbf{r})\right]\mathcal{G}_0(\mathbf{r},t,\mathbf{r}',t') = \delta(t-t')\delta(\mathbf{r}-\mathbf{r}') \tag{3.33}$$

Using this definition of the Green's function, we have $\int d\mathbf{r}_3\int dt_3\ \mathcal{G}_0(\mathbf{r}_1,t_1,\mathbf{r}_3,t_3)\mathcal{G}_0^{-1}(\mathbf{r}_3,t_3,\mathbf{r}_2,t_2) = \delta(\mathbf{r}_1-\mathbf{r}_2)\delta(t_1-t_2)$. From it, it is possible to show that:

$$\mathcal{G}(\mathbf{r}_1,t_1,\mathbf{r}_2,t_2) = \int d\mathbf{r}_3\int dt_3\int d\mathbf{r}_4\int dt_4\ \mathcal{G}_0(\mathbf{r}_1,t_1,\mathbf{r}_3,t_3)\mathcal{G}_0^{-1}(\mathbf{r}_3,t_3,\mathbf{r}_4,t_4)\mathcal{G}(\mathbf{r}_4,t_4,\mathbf{r}_2,t_2) \tag{3.34}$$

Using Eq. (3.33) and Eq. (3.34) in Eq. (3.31), one can write:

$$\int d\mathbf{r}_3 \int dt_3 \left[\mathcal{G}_0^{-1}(\mathbf{r}_1, t_1, \mathbf{r}_3, t_3) - \Sigma(\mathbf{r}_1, t_1, \mathbf{r}_3, t_3)\right] \mathcal{G}(\mathbf{r}_3, t_3, \mathbf{r}_2, t_2) = \delta(t_1 - t_2)\delta(\mathbf{r}_1 - \mathbf{r}_2) \tag{3.35}$$

After multiplying from the left by $\int dr \int t\mathcal{G}_0(\mathbf{r}_4, t_4, \mathbf{r}, t)$ and renaming the indices for clarity, one obtains:

$$\mathcal{G}(\mathbf{r}_1, t_1, \mathbf{r}_2, t_2) = \mathcal{G}_0(\mathbf{r}_1, t_1, \mathbf{r}_2, t_2) + \int d\mathbf{r}_3 \int dt_3 \int dr_4 \int dt_4 \mathcal{G}_0(\mathbf{r}_1, t_1, \mathbf{r}_3, t_3)\Sigma(\mathbf{r}_3, t_3, \mathbf{r}_4, t_4)\mathcal{G}(\mathbf{r}_4, t_4, \mathbf{r}_2, t_2) \tag{3.36}$$

This equation is called the Dyson equation [76]. For clarity and convenience as it will be developed in the next part for the implementation of the quantum transport, these equations will also be written in a matrix form of Green's function for a discretized real space basis. The Dyson equation for the Green's function matrix reads:

$$\mathbf{G} = \mathbf{G}_0 + \mathbf{G}_0\Sigma\mathbf{G} \tag{3.37}$$

where $\mathbf{G}_0$ is the contour-ordered Green's function matrix of the unperturbed system. We can also define the same functions for the self-energy as the lesser than, greater than, retarded and advanced Green's function. These self-energies respect a relation similar to Eq. (3.24). The kinetic equations for all these quantities are derived from Eq. (3.31) by using the Langreth's rules [100]. Langreth's rules allow one to express the contour integral of a product into an integral over real time. Then for a contour integral of a product of two terms which write $A = \int_C BC$, the lesser-than and retarded quantities of A write as:

$$\begin{aligned} A^<(t_1, t_2) &= \int_{t_1}^{t_2} dt_3 \left[B^R(t_1, t_3)C^<(t_3, t_2) + B^<(t_1, t_3)C^A(t_3, t_2)\right] \\ A^R(t_1, t_2) &= \int_{t_1}^{t_2} dt_3 B^R(t_1, t_3)C^R(t_3, t_2) \end{aligned} \tag{3.38}$$

From Eq. (3.31), we apply Langreth's rules for the contour integral with the self-energy. These equations can be put under the same form as Eq. (3.37) and read as for the retarded/advanced Green's functions [119]:

$$\begin{aligned} \mathbf{G}^r &= \mathbf{G}_0^r + \mathbf{G}_0^r\Sigma^r\mathbf{G}^r = \left[{\mathbf{G}_0^r}^{-1} - \Sigma^r\right]^{-1} \\ \mathbf{G}^a &= \mathbf{G}_0^a + \mathbf{G}_0^a\Sigma^a\mathbf{G}^a = \left[{\mathbf{G}_0^a}^{-1} - \Sigma^a\right]^{-1} \end{aligned} \tag{3.39}$$

And for the lesser-than and greater-than Green's functions, we have:

$$\begin{aligned} \mathbf{G}^< &= \mathbf{G}_0^< + \mathbf{G}_0^r\Sigma^r\mathbf{G}^< + \mathbf{G}_0^r\Sigma^<\mathbf{G}^a + \mathbf{G}_0^<\Sigma^a\mathbf{G}^a \\ \mathbf{G}^> &= \mathbf{G}_0^> + \mathbf{G}_0^r\Sigma^r\mathbf{G}^> + \mathbf{G}_0^r\Sigma^>\mathbf{G}^a + \mathbf{G}_0^>\Sigma^a\mathbf{G}^a \end{aligned} \tag{3.40}$$

In this case, it can be shown by iterating Eq. (3.40) inside itself that the final form is [119]:

$$\begin{aligned}\mathbf{G}^{<} &= (\mathbf{I}+\mathbf{G}^{r}\Sigma^{r})\mathbf{G}_{0}^{<}(\mathbf{I}+\mathbf{G}^{a}\Sigma^{a})+\mathbf{G}^{r}\Sigma^{<}\mathbf{G}^{a}\\ \mathbf{G}^{>} &= (\mathbf{I}+\mathbf{G}^{r}\Sigma^{r})\mathbf{G}_{0}^{>}(\mathbf{I}+\mathbf{G}^{a}\Sigma^{a})+\mathbf{G}^{r}\Sigma^{>}\mathbf{G}^{a}\end{aligned} \tag{3.41}$$

3.1.B.5 Steady-state kinetic equations

Now that all the kinetic equations for the Green's functions have been derived, we express them under steady-state conditions. Under these conditions, the Green's functions depend only on the difference between the two times t and t' that is $\tau = t - t'$. Then we use the Fourier transform to express the Green's functions as well as the self-energies in the energy domain E as:

$$G(\mathbf{r},\mathbf{r}',E) = \int d\tau \mathcal{G}(\mathbf{r},\mathbf{r}',\tau)e^{iE\tau} \tag{3.42}$$

The kinetic equations read:

$$\begin{aligned}&[E-H(\mathbf{r})]G^{r/a}(\mathbf{r},\mathbf{r}',E) =\\ &\qquad \delta(\mathbf{r}-\mathbf{r}') + \int d\mathbf{r}''\Sigma^{r/a}(\mathbf{r},\mathbf{r}'',E)\,G^{r/a}(\mathbf{r}'',\mathbf{r},E)\\ &[E-H(\mathbf{r})]G^{</>}(\mathbf{r},\mathbf{r}',E) =\\ &\qquad \delta(\mathbf{r}-\mathbf{r}') + \int d\mathbf{r}''\left[\Sigma^{r/a}(\mathbf{r},\mathbf{r}'',E)\,G^{</>}(\mathbf{r}'',\mathbf{r},E)+\Sigma^{</>}(\mathbf{r},\mathbf{r}'',E)\,G^{a/r}(\mathbf{r}'',\mathbf{r},E)\right]\end{aligned} \tag{3.43}$$

We move back to the discretized real space basis, where the energy dependent Green's function is represented by a matrix. The equations remain as the Eq. (3.39) for $\mathbf{G}^{r/a}$. For lesser-than and greater-than Green's functions, it can be shown that, in steady state conditions [5], the first term of the right side of Eq. (3.41) vanishes as it corresponds to the initial condition of a non-interacting system. Thus, they become:

$$\begin{aligned}\mathbf{G}^{r} &= [(E-i\eta)\mathbf{I}-\mathbf{H}-\Sigma^{r}]^{-1}\\ \mathbf{G}^{a} &= [(E+i\eta)\mathbf{I}-\mathbf{H}-\Sigma^{a}]^{-1}\\ \mathbf{G}^{<} &= \mathbf{G}^{r}\Sigma^{<}\mathbf{G}^{a}\\ \mathbf{G}^{>} &= \mathbf{G}^{r}\Sigma^{>}\mathbf{G}^{a}\end{aligned} \tag{3.44}$$

where η is an infinitesimal term added to ensure the existence of the Green's functions.

3.1.C Electrical and transport properties from NEGF

As mentioned in the previous parts, the Green's functions that we derived are directly linked to the electrical and transport properties of the system.

3.1.C.1 Density of states

The density of states describes the energy level of a system and is directly linked to the Hamiltonian. This information is given by the spectral function, noted A, that we define as [5]:

$$A(\mathbf{r},\mathbf{r}',E) = i\left[G^r(\mathbf{r},\mathbf{r}',E) - G^a(\mathbf{r},\mathbf{r}',E)\right] \tag{3.45}$$

And the density of states corresponds to the diagonal terms of the spectral function:

$$DOS(E) = \frac{1}{2\pi}\int d\mathbf{r} A(\mathbf{r},\mathbf{r},E) \tag{3.46}$$

3.1.C.2 Spectral carrier density

The spectral carrier density corresponds to the electron and hole concentration. It is a statistical quantity of the system and depends on the energy level occupied by the electrons. As such, it is described by the lesser than and greater than Green's functions for the electrons n and holes p respectively [5].

$$\begin{aligned} n(\mathbf{r},E) &= -\frac{1}{2\pi}\mathrm{Im}(G^<(\mathbf{r},\mathbf{r},E)) \\ p(\mathbf{r},E) &= \frac{1}{2\pi}\mathrm{Im}(G^>(\mathbf{r},\mathbf{r},E)) \end{aligned} \tag{3.47}$$

3.1.C.3 Spectral current density

The spectral current density j is directly derived from its expression given with the Schrödinger equation from elementary quantum physics where it reads [76]:

$$j(\mathbf{r},E) = -\frac{e\hbar}{2m_0}\lim_{\mathbf{r}'\to r}(\nabla_{\mathbf{r}} - \nabla_{\mathbf{r}'})G^<(\mathbf{r},\mathbf{r}',E) \tag{3.48}$$

Then to compute the current in the device, we consider the device region with the open-boundary condition. In this regards, one has to connect the device with two semi-infinite leads that are represented by the self-energies Σ_L and Σ_R for the left and right contact self-energies. From this, we define the transmission as [5]:

$$T(E) = \mathrm{tr}(\mathrm{G^{r\backprime}_L G^{a\backprime}_R}) \tag{3.49}$$

where tr represents the trace and $\Gamma_{L(R)}$ represents the broadening of the energy state due to the coupling with semi-infinite leads which is computed from the self-energies as:

$$\Gamma = i\left[\Sigma^r - \Sigma^a\right] \tag{3.50}$$

From the transmission defined in Eq. (3.49), the current I, at the leads, can be calculated with the Landauer's formula [5]:

$$I = \frac{2e}{\hbar} \int dE \;\; T(E) \left[f(E - \mu_S) - f(E - \mu_D) \right] \tag{3.51}$$

where f is the Fermi-Dirac distribution and μ_S and μ_D are the Fermi levels in the source and the drain, respectively.

3.2 Device simulation using plane-waves Hamiltonians

In this section, we focus on the description of quantum transport in nanoscale devices and physical systems, based on a plane-wave Hamiltonian from EPM [135, 136] or DFT [137] calculation. We choose these two approaches as they provide a full band description of the material and an atomistic wave function in the real space needed to study nanoscale devices. A generic plane-wave Hamiltonian takes the following form:

$$H_{\mathbf{k}}(\mathbf{G}, \mathbf{G}') = \underbrace{\frac{\hbar^2}{2m}(\mathbf{k} + \mathbf{G})^2 \delta_{\mathbf{G},\mathbf{G}'}}_{T} + \underbrace{V_L(\mathbf{G} - \mathbf{G}') + V_{NL}(\mathbf{k} + \mathbf{G}, \mathbf{k} + \mathbf{G}')}_{V} \tag{3.52}$$

where $\mathbf{G}$ is the reciprocal lattice vector, T is the kinetic term, V is the potential that we can separate in local part V_L and non-local part V_{NL}. The latter is null in the case of local EPM, but not for the DFT Hamiltonian due to the pseudopotential reproducing the interaction of valence and core electrons with the core.

For the transport calculation, we choose an expansion volume for $\mathbf{G}$ vectors that is given by the cube set by the condition:

$$|\mathbf{G}_s|^2 \leq \frac{N_d}{2} \left(\frac{2\pi}{a_0} \right) \tag{3.53}$$

with $s = x, y, z$. This definition allows us to link the discretized real space used in the device simulation with the $\mathbf{G}$ vectors used in the calculations of the Hamiltonian.

By assuming the periodicity of the potential V due to the periodic structure of the crystal studied, according to Bloch theorem, the wave function and eigenvalues are obtained by solving the following secular equation:

$$\sum_{\mathbf{G}'} H_{\mathbf{k}}(\mathbf{G}, \mathbf{G}') B_{\mathbf{k}}(\mathbf{G}') = E_{\mathbf{k}} B_{\mathbf{k}}(\mathbf{G}) \tag{3.54}$$

where $E_{\mathbf{k}}$ is the energy of the corresponding state and the eigenvectors $B_{\mathbf{k}}$ represent the component of the periodic part of the Bloch waves functions.

$$\Phi_{\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} u_{n,\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} \sum_{\mathbf{G}} B_{\mathbf{k}}(\mathbf{G}) e^{i\mathbf{G}\cdot\mathbf{r}} \tag{3.55}$$

where $u_{n,\mathbf{k}}(\mathbf{r})$ is the periodic part of the Bloch function. Concerning the calculation of the kinetic energy operator,

within a continuous real-space model, in the reciprocal space, it is expressed as:

$$T(\mathbf{k}+\mathbf{G}) = \frac{\hbar^2(\mathbf{k}+\mathbf{G})^2}{2m_0} \tag{3.56}$$

However, in the formalism for electronic transport, we need to discretize the real space in order to use an hybrid basis detailed in the next section. If one describes the second derivative with a centered difference approximation of order $2p$, the kinetic energy operator in the reciprocal space takes the following form:

$$\begin{gathered} T(\mathbf{k}+\mathbf{G}) = -t_0 \sum_{s=x,y,z} \left[C_0 + 2\sum_{r=1}^{p} C_r \cos\left(r(k_s+G_s)\,d\right) \right] \\ \text{with} \quad t_0 = \frac{\hbar^2}{2m_0 d^2} \end{gathered} \tag{3.57}$$

where the coefficients C_r are the convolution mask for the second derivative and d is the same mesh size for all the space directions. In general, pseudopotential Hamiltonian calculations require a higher than second-order discretization for the kinetic energy operator [35]. If not stated otherwise, $2p$=8 is chosen as for the general discretization step chosen. It is high enough for the discretited kinetic energy values to converge with the continuous expression [135].

Solving the Green's functions equation for an Hamiltonian obtained with the method presented in Chapter 2 is computationally expensive due to the matrix product and matrix inversion from Eq. (3.39) and Eq. (3.44). The bottle neck of this problem is the size of the matrix given in a plane wave basis where in the case of DFT Hamiltonian, it can reach tens of thousands elements. In order to make the problem tractable, it is mandatory to reduce the size of the Hamiltonian. The first step is to remark that only the diagonal elements from the Green's functions are necessary to derive the physical quantities of interest. Knowing that, it is possible to use a recursive algorithm presented in Appendix A [163]. To apply the recursive algorithm, one needs to have a tridiagonal by block Hamiltonian. In order to simplify the Hamiltonian, two transformations are required. A first on to make the Hamiltonian tridiagonal by block and the second one allow us to reduce the size of these block.

3.2.A Use of a hybrid basis

The first transformation into a hybrid basis is related to solve different problems concerning device simulation. Indeed, to simulate a device, it is necessary to considerate semi-infinite leads which will connect the simulated device region. In order to connect the system to the leads without including the leads into the system, we choose to transform the transport direction component of plane waves into real space. This is possible thanks to the separability property of the plane waves. The real space along the transport direction, x axis, is thus discretized according to the number of vectors G_x that we are considering in the Hamiltonian calculation as $d_x = a_x/N_{G_x}$ where a_x is the lattice parameter of the unit cell considered along the transport direction.

This transformation from the plane wave to a hybrid basis made of a real space component along the transport

direction and plane waves for the two other transverse directions presents the advantage to accurately describe the Hamiltonian with a tridiagonal by block matrix even with the non-local term of the DFT Hamiltonian as they present a relatively short-range (a few Ångstroms) nature which makes them negligible after the first neighbour. The Hamiltonian presents now the following form:

$$[\mathbf{H}]_{x,\mathbf{k}_{yz}+G_{yz}} = \begin{bmatrix} \ddots & \mathbf{H}_{0,1} & 0 & \dots \\ \mathbf{H}_{0,1}^{\dagger} & \mathbf{H}_{i,i} & \mathbf{H}_{0,1} & 0 \\ 0 & \mathbf{H}_{0,1}^{\dagger} & \mathbf{H}_{i+1,i+1} & \mathbf{H}_{0,1} \\ \vdots & 0 & \mathbf{H}_{0,1}^{\dagger} & \ddots \end{bmatrix} \tag{3.58}$$

where each block describes an a_x long region consisting of N_{d_x} discretization points $x_j = 0, d_x, 2d_x ... (a_x - d_x)$, so that blocks $\mathbf{H}_{i,i}$ and $\mathbf{H}_{0,1}$ have a rank $N_G = N_{d_x} N_{G_y} N_{G_z}$ (with $N_{d_x} = N_{G_x}$). We obtain this value for N_G since we use a volume expansion for $\mathbf{G}$ vectors that is given by a cube in the reciprocal space for the transport calculation.

In this form, all the off-diagonal blocks $\mathbf{H}_{0,1}$ are identical for an homogeneous system and the diagonal blocks $\mathbf{H}_{i,i}$, that represent the i-th unit-cell along the transport direction, only differ from one another by the external potential, due to an electric field or doping, present at their positions. If we are in the case of an homogeneous material, all the diagonal blocks are identical and will be denoted $\mathbf{H}_{0,0}$. These blocks can be evaluated numerically similarly for EPM and DFT, from a unitary transformation, that transform the x component of the plane waves into real-space, with the matrix $[\mathbf{U}]_{k_x+G_x,x}$ that reads:

$$[\mathbf{U}_{k_x+G_x}] = \frac{1}{\sqrt{N_{G_x}}} \begin{bmatrix} \mathbf{I} & e^{i(k_x+G_{x1})d_x}\mathbf{I} & \dots & e^{i(k_x+G_{x1})(a_x-d_x)}\mathbf{I} \\ \mathbf{I} & e^{i(k_x+G_{x2})d_x)}\mathbf{I} & \dots & e^{i(k_x+G_{x2})(a_x-d_x)}\mathbf{I} \\ \dots & \dots & \ddots & \vdots \\ \mathbf{I} & e^{i(k_x+G_{xN_{G_x}})d_x}\mathbf{I} & \dots & e^{i(k_x+G_{xN_{G_x}})(a_x-d_x)}\mathbf{I} \end{bmatrix} \tag{3.59}$$

where $\mathbf{I}$ is the identity matrix with a rank of $N_{G_y} N_{G_z}$. In the case of a local Hamiltonian like for EPM, it is enough to transform the plane waves Hamiltonian with $[\mathbf{U}]_{k_x+G_x,x}$ as:

$$[\mathbf{H}]_{x,\mathbf{k}_{yz}+\mathbf{G}_{yz}} = [\mathbf{U}_{k_x+G_x}]^{\dagger}[H_{\mathbf{k}}][\mathbf{U}_{k_x+G_x}] \tag{3.60}$$

where $[H_{\mathbf{k}}]$ is defined in Eq. (3.52). Then by identification, it is possible to get $\mathbf{H}_{i,i}$ and $\mathbf{H}_{0,1}$. In the case of DFT Hamiltonian, these blocks are estimated by considering the Hamiltonian from neighbouring unit cells along the transport direction. If we consider N_{cx} neighbouring unit cells, the unitary transformation is performed with the matrix:

$$[\mathbf{U}^{N_{cx}}_{k_x+G_x}] = \frac{1}{\sqrt{N_{cx}}} \left[[\mathbf{U}_{k_x+G_x}], [\mathbf{U}_{k_x+G_x}] e^{ik_x a_x}, ..., [\mathbf{U}_{k_x+G_x}] e^{ik_x (N_{cx}-1) a_x} \right] \tag{3.61}$$

Then, the Hamiltonian is transformed into the hybrid basis with:

$$[\mathbf{H}]_{x,\mathbf{k}_{yz}+\mathbf{G}_{yz}} = \sum_{k_x} [\mathbf{U}^{N_{cx}}_{k_x+G_x}]^\dagger [H_\mathbf{k}][\mathbf{U}^{N_{cx}}_{k_x+G_x}] \tag{3.62}$$

With k_x running over the N_{cx} and taking values in the interval $[-\frac{\pi}{a_x}, \frac{\pi}{a_x}[$. Then $\mathbf{H}_{0,0}$ and $\mathbf{H}_{0,1}$ can be identified as:

$$\begin{aligned} \mathbf{H}_{0,0} &= \frac{1}{N_{cx}} \sum_{k_x} [\mathbf{U}_{k_x+G_x}]^\dagger [H_\mathbf{k}][\mathbf{U}_{k_x+G_x}] \\ \mathbf{H}_{0,1} &= \frac{1}{N_{cx}} \sum_{k_x} [\mathbf{U}_{k_x+G_x}]^\dagger [H_\mathbf{k}][\mathbf{U}_{k_x+G_x}] e^{ik_x a_x} \end{aligned} \tag{3.63}$$

3.2.B Reduced basis

The second transformation is now related to reduce the size of the Hamiltonian to achieve efficient calculations. The Hamiltonian in the hybrid basis is still too large as the number of plane waves along y and z directions remains important. For DFT Hamiltonian, this is critical as the cut-off energy necessary to reach convergence is generally of the order of 40 Ry which results in tens of thousands plane waves for the y and z directions. For the EPM Hamiltonian, even if the number of plane waves is relatively low due to the smoothness of the wave function in the core region of the atoms, the block size in the hybrid basis may still be reduced to a lower number. For the choice of the reduced basis, it is important to remark that only the carriers of energy typically close to the Fermi level contribute to the current. An energy window of some eV around the Fermi level is enough to compute the current. The idea is to project the transverse plane waves components of the Hamiltonian into a basis that reproduces only the band inside an energy window of interest. For the EPM Hamiltonian, it is possible to do that with the mode-space transformation.

3.2.B.1 Mode-space transformation

The principle of the mode-space transformation is to form a reduced basis made of the lowest energy eigenmodes on the cross-section of the structure. These modes are obtained by solving the following secular equation at $k_x = 0$ and for each single x discretization point sections along the transport direction of this work:

$$\left[\mathbf{H}^\dagger_{0,1} + \mathbf{H}_{0,0} + \mathbf{H}_{0,1}\right] \Phi^{\mathrm{MS}}_n = E_n \Phi^{\mathrm{MS}}_n \tag{3.64}$$

where $\mathbf{H}_{0,0}$ and $\mathbf{H}_{0,1}$ correspond to the Hamiltonian blocks in the Hybrid basis for a single x point. Φ^{MS}_n is the eigenvector corresponding to the n-th mode. By selecting enough modes to reconstruct the band structure in our energy window, the transformation matrix is defined from the previously determined eigenmodes as:

$$\mathbf{U}_{\mathrm{MS}} = \left[\Phi^{\mathrm{MS}}_1, \Phi^{\mathrm{MS}}_2, .., \Phi^{\mathrm{MS}}_{N_{mod}}\right] \tag{3.65}$$

Then the matrix is used on the Hamiltonian blocks represented in the hybrid basis. To apply this method, we are forced to assume a local form of the Hamiltonian. Indeed, for a non-local Hamiltonian, using mode-space transformation results in unphysical bandstructure. Indeed, the mode space cannot describe the full band because it uses a projection of the Bloch states only for $k_x = 0$, which means that we neglect interaction with the nearest neighbour section to describe only the transverse direction modes and the coupling between them. In non-local Hamiltonian, there are interactions between unit-cells along all directions. It results that the modes defined at one unit-cell may be coupled to the modes associated with the other unit-cells even along the transport direction. However, this information is lost in the mode-space basis if we do not take all the modes. It is important to mention that it could be possible to use the mode-space basis with non-local Hamiltonian. However, it requires a post-processing treatment on the reduced basis where designed states are constructed and added to the basis in order to put the spurious states outside of the energy window of interest [127, 173].

3.2.B.2 Unit-cell restricted Bloch functions

A general extension of the mode-space basis to non-local Hamiltonian is needed to be able to treat efficiently the non-local Hamiltonians. The idea is to use unit-cell restricted Bloch functions to construct a basis for the lowest energy transverse modes similarly to the mode-space basis. The difference with the mode-space basis is that the Bloch's functions are obtained by solving the secular equation over the unit-cell instead of a section of the unit-cell as it is sketched in the Fig. 3.2.

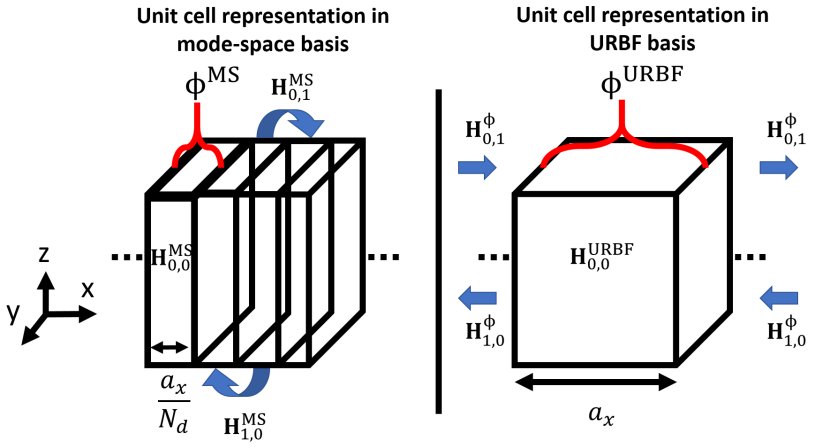

Figure 3.2: Schematic representation of the Hamiltonian for a unit cell in the mode space basis and the URBF basis.

The Bloch functions are taken at different values of k_x, which means that we consider interactions with the neighbouring unit-cells along the transport direction. It is convenient with a view to reproducing non-local Hamiltonians as interactions between adjacent unit-cells are considered.

To express the Bloch functions in the n-th unit cell from the home unit-cell Bloch functions, we make use of the

periodicity of the Bloch function. It writes as:

$$\Psi^n_{\mathbf{k}}(x + na_x, \mathbf{G}_{yz}) = \Psi^0_{\mathbf{k}}(x, \mathbf{G}_{yz}) e^{ik_x n a_x} \tag{3.66}$$

where $x \in [0 : a_x]$ and Ψ^n is the Bloch function of the n-th unit-cell away from the home unit-cell which has the corresponding Bloch function Ψ^0. The latter Bloch functions is the solution to the secular equation:

$$\left[\mathbf{H}^{\dagger}_{0,1} e^{-ik_x a_x} + \mathbf{H}_{0,0} + \mathbf{H}_{0,1} e^{ik_x a_x}\right] \Psi^0_{\mathbf{k}} = E_{\mathbf{k}} \Psi^0_{\mathbf{k}} \tag{3.67}$$

where the only requirement of this equation is the boundary condition of the Bloch function on x for the selected k_x:

$$\Psi^0_{\mathbf{k}}(a_x, \mathbf{G}_{yz}) = \Psi^0_{\mathbf{k}}(0, \mathbf{G}_{yz}) e^{ik_x a_x} \tag{3.68}$$

The Bloch functions from the principal unit cell are obtained by solving Eq. (3.67). They are chosen to sample the first Brillouin zone and by choosing the appropriate number of bands N_{B} to reconstruct the bands in the energy window needed to describe the transport properties.

The Bloch functions $\Psi_{m,\mathbf{k}}$, $\Psi_{n,\mathbf{k}'}$ restricted to the same unit cell along the transport direction x are orthogonal for $k'_x = k_x$ and $\mathbf{k}'_{yz} = \mathbf{k}_{yz}$, and are also orthogonal for $\mathbf{k}'_{yz} \neq \mathbf{k}_{yz}$(independently of k'_x and k_x) due to the periodic boundary conditions along y and z. Moreover, Bloch functions restricted to different unit cells along x are by definition orthogonal. The $\Psi_{m,\mathbf{k}}$, $\Psi_{n,\mathbf{k}'}$ restricted to the same unit cell along x are instead not orthogonal for $\mathbf{k}'_{yz} \neq \mathbf{k}_{yz}$ and $k'_x \neq k_x$, and we denote with $S_{n,k_x,m,k'_x}(\mathbf{k}_{yz}) = \left\langle \Psi_{m,\mathbf{k}} \middle| \Psi_{n,\mathbf{k}'} \right\rangle$ the corresponding matrix of the overlap integrals. The basis set of orthogonal URBFs for a given $\mathbf{k}_{yz}$ can be concisely written as:

$$\Phi_{i,k_x,\mathbf{k}_{yz}}(\mathbf{r}) = \begin{cases} \sum_{n,k_x} \left[S^{-\frac{1}{2}}(\mathbf{k}_{yz})\right]_{n,k_x,i} \frac{u_{n,\mathbf{k}}(\mathbf{r})}{\sqrt{N_{cyz}}} e^{i\mathbf{k}.\mathbf{r}} \\ 0 \text{ otherwise} \end{cases} \tag{3.69}$$

where $u_{n,\mathbf{k}}(\mathbf{r})$ is the periodic part of the Bloch function of mode n which is normalized over a single unit-cell, and N_{cyz} is the number of unit cells along the periodic directions y and z. The index i refers to a basis function corresponding to a given $\mathbf{k}_{yz}$, and each $\Phi_{i,k_x,\mathbf{k}_{yz}}$ is defined as an appropriate linear combination of N_B Bloch functions $\Psi_{m,\mathbf{k}}$ (restricted to the unit cell). The orthonormalisation of the basis is not necessary, but it allows the transformation from the reduced basis to real space basis which is necessary for quantities like charge density. Then the Hamiltonian for a given $\mathbf{k}_{yz}$ is projected on this basis as presented in Eq. (3.70).

$$[\mathbf{H}]^{\Phi}_{x,\mathbf{k}_{yz}} = {\chi^{\Phi}_{yz}}^{\dagger} [\mathbf{H}]_{x,\mathbf{k}_{yz}+G_{yz}} \chi^{\Phi}_{yz} \tag{3.70}$$

where $\mathbf{H}^{\Phi}$ is the Hamiltonian in the URBF basis and χ^{Φ}_{yz} is:

$$\chi^{\Phi}_{yz} = \left[\Phi_{n_1,k_{x_1},\mathbf{k}_{yz}}, .., \Phi_{n_1,k_{xN},\mathbf{k}_{yz}}, \Phi_{n_2,k_{x_1},\mathbf{k}_{yz}}, .., \Phi_{n_B,k_{xN},\mathbf{k}_{yz}}\right] \tag{3.71}$$

where N corresponds to the number of k_x points taken and N_B is the number of bands chosen.

3.2.C The concept of self-energy

As mentioned above when giving its definition in Sec. 3.1.B.4, the concept of self-energy allows us to include external perturbations to the system and to take into account many-body interactions. The general form of a system of interest is sketched in Fig. 3.3. It represents our unperturbed system, described by the Hamiltonian $\mathcal{H}_{device}$, coupled to different sources of interactions. For device simulations, we have to consider an open boundary system. The electrical contacts are idealized as a semi-infinite leads on the left and right sides through two contact self-energies, one at the left Σ_L and one at the right Σ_R. Moreover, interaction with a phonon reservoir is also taken into account with a self-energy.

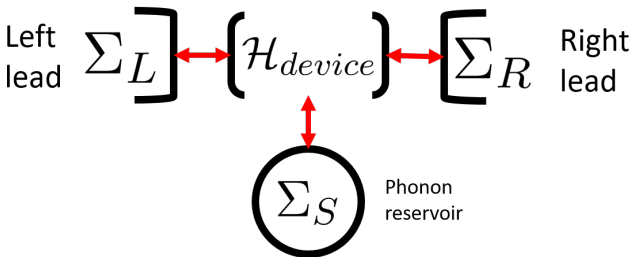

Figure 3.3: Schematic representation of the device studied with NEGF method. The system is coupled with to semi-infinite leads whose interaction with the device is expressed through the two self energies Σ_L and Σ_R. Carriers in the device are also scattered due to the interaction with phonons which is represented by the self energy Σ_S.

3.2.C.1 Contact self-energy

In order to define the contact self-energy, we have to separate the system into three regions: the left lead, the device and the right lead. The leads are the contacts by which the electron are injected and collected. The device and lead regions are divided by layers stacked along the transport direction.

A sketch of the Hamiltonian of the total system is presented in Fig. 3.4. In this figure, we identify the three distinct regions. For clarity, all the regions are made of the same material. Each layer is described by a section Hamiltonian $\mathbf{H}_{i,i}$, where i is the index of the section, which is connected to the first neighbour sections by the coupling term $\mathbf{H}_{0,1}$. The two leads are assumed to be semi-infinite. The coupling between the leads and the device is described by the Hamiltonian matrices $\mathbf{H}_{L,D}$, where we have $\mathbf{H}_{L,D} = \mathbf{H}_{0,1}$) since the lead and the device region are made of the same material and $\mathbf{H}_{D,R}$ (also with $\mathbf{H}_{D,R} = \mathbf{H}_{0,1}$) for the left and right leads, respectively.

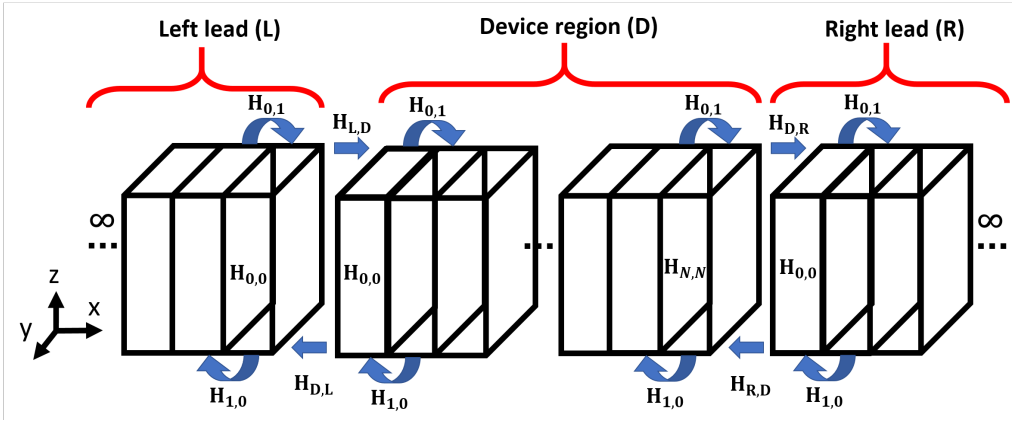

Figure 3.4: Schematic representation of the Hamiltonian of the device connected to two semi-infinite leads.

As we consider the coupling between sections to be limited to its first neighbours, the left lead couples to the Green's function of the first section of the device and the right lead couples to the Green's function of the last section of the device region. Then by using the tridiagonal by block nature of the Hamiltonian, one can demonstrate that the self-energies from the leads write [7]:

$$\begin{aligned}\Sigma_L &= \mathbf{H}_{D,L}\mathbf{g}^S_{L(R)}\mathbf{H}_{L,D}\\ \Sigma_R &= \mathbf{H}_{D,R}\mathbf{g}^S_{L(R)}\mathbf{H}_{R,D}\end{aligned} \tag{3.72}$$

where $\mathbf{g}^S_{L(R)}$ is the surface Green's function of the left (right) lead. The surface Green's function is the unknown quantity that we need to derive. In order to compute it, a recursive algorithm was proposed by Sancho et al. [163]. First, we have to assume that a lead is a semi-infinite chain of periodic section. Each section is described by a local Hamiltonian $\mathbf{H}_{0,0}$ and is connected to its first neighbor sections with a constant coupling $\mathbf{H}_{0,1}$ and $\mathbf{H}_{1,0} = \mathbf{H}^\dagger_{0,1}$. The equations to be iterated at the i-th step until convergence is reached are as follows:

$$\begin{aligned}\mathbf{H}^B_i &= \mathbf{H}^B_{i-1} + \mathbf{H}^R_{i-1}\left[(E+i\eta)\mathbf{I} - \mathbf{H}^B_{i-1}\right]^{-1}\mathbf{H}^L_{i-1} + \mathbf{H}^L_{i-1}\left[(E+i\eta)\mathbf{I} - \mathbf{H}^B_{i-1}\right]^{-1}\mathbf{H}^R_{i-1}\\ \mathbf{H}^S_i &= \mathbf{H}^S_{i-1} + \mathbf{H}^R_{i-1}\left[(E+i\eta)\mathbf{I} - \mathbf{H}^B_{i-1}\right]^{-1}\mathbf{H}^L_{i-1}\\ \mathbf{H}^R_i &= \mathbf{H}^R_{i-1}\left[(E+i\eta)\mathbf{I} - \mathbf{H}^B_{i-1}\right]^{-1}\mathbf{H}^R_{i-1}\\ \mathbf{H}^L_i &= \mathbf{H}^L_{i-1}\left[(E+i\eta)\mathbf{I} - \mathbf{H}^B_{i-1}\right]^{-1}\mathbf{H}^L_{i-1}\end{aligned} \tag{3.73}$$

where i is the index of the iteration and we define different effective Hamiltonians at each iteration. $\mathbf{H}^B$ represents the 'bulk' Hamiltonian as it corresponds to the Hamiltonian of a section far from the surface that is connected at the right and left side by a group of section that double at each iteration, while $\mathbf{H}^S$ represents the 'surface' Hamiltonian. It corresponds to the Hamiltonian of the first section that is connected at the left with the group of sections included in the bulk Hamiltonian. We also define two coupling Hamiltonian, $\mathbf{H}^R$ and $\mathbf{H}^L$, that represent the right and left

coupling, respectively, of a section with the second neighbor sections. At the first step, the set of equations is:

$$\begin{aligned}
\mathbf{H}_1^B &= \mathbf{H}_{0,0} + \mathbf{H}_{0,1}\left[(E+i\eta)\mathbf{I} - \mathbf{H}_{0,0}\right]^{-1}\mathbf{H}_{1,0} + \mathbf{H}_{1,0}\left[(E+i\eta)\mathbf{I} - \mathbf{H}_{0,0}\right]^{-1}\mathbf{H}_{0,1} \\
\mathbf{H}_1^S &= \mathbf{H}_{0,0} + \mathbf{H}_{0,1}\left[(E+i\eta)\mathbf{I} - \mathbf{H}_{0,0}\right]^{-1}\mathbf{H}_{1,0} \\
\mathbf{H}_1^R &= \mathbf{H}_{0,1}\left[(E+i\eta)\mathbf{I} - \mathbf{H}_{0,0}\right]^{-1}\mathbf{H}_{0,1} \\
\mathbf{H}_1^L &= \mathbf{H}_{1,0}\left[(E+i\eta)\mathbf{I} - \mathbf{H}_{0,0}\right]^{-1}\mathbf{H}_{1,0}
\end{aligned} \tag{3.74}$$

The set of equations for the recursive calculation are obtained by using the Dyson equation [119] (see Eq. (3.37)) to couple the unperturbed system to the neighbor sections.

Once the convergence of the iterative loop is attained, from the surface Hamiltonian, the surface Green's function writes:

$$\mathbf{g}_{L(R)}^S = \left[(E+i\eta)\mathbf{I} - \mathbf{H}^S\right]^{-1} \tag{3.75}$$

This algorithm allows a quick convergence of the self-energy since after n iterations, 2^n sections are taken into account in the surface Green's function.

3.2.C.2 Electron-phonon self-energy

Electron-phonon scattering represents one of the major sources scattering present in nanoscale semiconductor devices. It is important to take this interaction into account, but we have to assume some approximations. The first is to consider the phonon bath at equilibrium, which implies that the phonon Green's function reduces to the Bose-Einstein distribution and writes as:

$$\mathbf{D}_0^<(\mathbf{q}, \hbar\omega_\mathbf{q}) \propto N_\mathbf{q} + 1/2 \pm 1/2 \tag{3.76}$$

where $\mathbf{D}_0^<$ is the unperturbed lesser-than Green's function for the phonon bath. Then the self-energy can be derived within the first-order self-consistent Born approximation (SCBA) [77], which is necessary to achieve current conservation in steady state conditions. Within this approximation the lesser-than self-energy reads:

$$\Sigma^< = \mathbf{D}_0^<\mathbf{G}^< \tag{3.77}$$

where $\mathbf{G}^<$ is the lesser-than Green's function for the electrons. It is represented by the Feynmann diagram of Fig. 3.5.

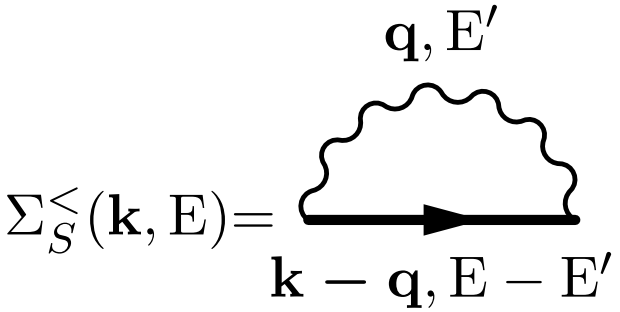

Figure 3.5: Diagrammatic representation of the self-consistent Born approximation for the electron–phonon interaction (see Ref. [76])

In order to fulfill the requirements of the recursive Green's functions algorithms, we need to consider electron-phonon interaction to be local in real space. It allows us to have a block tridiagonal Hamiltonian. The next approximation concerns the electron-phonon matrix element. In the following, we distinguish acoustic and optical phonons. For acoustic phonon scattering at room temperature, the elastic approximation can be used, namely considering $\mathrm{E} \pm \hbar\omega_{\mathrm{q}} \approx \mathrm{E}$, where ω_{q} is the phonon frequency at wave vector $\mathbf{q}$. Optical phonon modes are considered dispersion-less. Moreover, by using the scalar assumption for the acoustic deformation potential D_{ac} the electron-acoustic phonon matrix element reads out

$$\left|M_{\mathbf{q}}\right|^2_{\mathrm{ac}} = \frac{D^2_{\mathrm{ac}} \hbar \mathrm{q}}{2\rho v_s} \tag{3.78}$$

where $\mathrm{M}_{\mathbf{q}}$ is the matrix element for the el-ph interaction and $\mathbf{q}$ is the phonon wave-vector, $\mathrm{q} = |\mathbf{q}|$, ρ is the mass density, and v_s is the sound velocity. Due to the elastic approximation, the Bose-Einstein distribution becomes: $N_{\nu,\mathbf{q}} \approx N_{\nu,\mathbf{q}} + 1 \approx k_B\, T/(\hbar v_s\, q)$ and the term $\left|\mathrm{M}_{\mathbf{q}}\right|^2 \mathbf{D}^{<}_0(\mathbf{q}, E')$ in Eq. (3.77) for the acoustic phonon contribution becomes q-independent and it reads:

$$\left|M_{\mathbf{q}}\right|^2_{\mathrm{ac}} \mathbf{D}^{<}_0(\mathbf{q}, E') \approx \frac{D^2_{\mathrm{ac}}}{\rho v_s^2} \delta(E'). \tag{3.79}$$

For the optical phonons, by assuming a mode ν with a constant energy transfer $\hbar\omega_\nu$, a formulation of the deformation potential independent of q is given by

$$\left|\mathrm{M}_{\mathbf{q}}\right|^2_{\mathrm{op},\nu} = \frac{\hbar D^2_{\mathrm{op},\nu}}{2\rho\omega_\nu} \tag{3.80}$$

and the term $\left|\mathrm{M}_{\mathbf{q}}\right|^2 \mathbf{D}^{<}_0(\mathbf{q}, E')$ in Eq. (3.77) for the optical phonon contribution reads out for phonon emission process:

$$\left|\mathrm{M}_{\mathbf{q}}\right|^2_{\mathrm{op}} \mathbf{D}^{<}_0(\mathbf{q}, E') \approx \sum_\nu \frac{\hbar D^2_{\mathrm{op},\nu}}{2\rho\omega_\nu} \left[N_{\nu,\mathbf{q}} + 1\right] \delta(E' + \hbar\omega_\nu). \tag{3.81}$$

and for phonon absorption process:

$$\left|\mathrm{M}_{\mathbf{q}}\right|^2_{\mathrm{op}} \mathbf{D}^{<}_0(\mathbf{q}, E') \approx \sum_\nu \frac{\hbar D^2_{\mathrm{op},\nu}}{2\rho\omega_\nu} \left[N_{\nu,\mathbf{q}}\right] \delta(E' - \hbar\omega_\nu). \tag{3.82}$$

By using Eqs. (3.79-3.82), the self-energies for acoustic and optical phonon scattering in the real space read

$$\Sigma^{<,(\mathrm{ac})}(\mathbf{r},\mathbf{r}',\mathbf{k}_{yz},E)=\frac{D_{\mathrm{ac}}^2 k_B\,T}{\rho v_s^2}\sum_{\mathbf{q}_{yz}} G^{<}(\mathbf{r},\mathbf{r}',\mathbf{k}_{yz}-\mathbf{q}_{yz},E)\,\delta(\mathbf{r}-\mathbf{r}') \tag{3.83}$$

and

$$\Sigma^{<,(\mathrm{op})}(\mathbf{r},\mathbf{r}',\mathbf{k}_{yz},E)=\sum_{\nu}\frac{\hbar D_{\mathrm{op},\nu}^2}{2\rho\omega_\nu}\left[N_\nu+\frac{1}{2}\pm\frac{1}{2}\right]\sum_{\mathbf{q}_{yz}} G^{<}(\mathbf{r},\mathbf{r}',\mathbf{k}_{yz}-\mathbf{q}_{yz},E\pm\hbar\omega_\nu)\,\delta(\mathbf{r}-\mathbf{r}'). \tag{3.84}$$

Then, the self-energy for a given $\mathbf{k}_{yz}$ in the reduced space is obtained by performing an transformation using the rectangular matrix χ_{yz} defined in Eq. (3.71). The generic component of self-energy matrix in the reduced space reads:

$$\begin{aligned}
\Sigma^{<}_{n,n'}(\mathbf{k}_{yz},E)&=\frac{1}{\Omega}\int d\mathbf{r}\int d\mathbf{r}'\,\Phi^{\dagger}_{n,\mathbf{k}_{yz}}(\mathbf{r})\,\Sigma^{<}(\mathbf{r},\mathbf{r}',\mathbf{k}_{yz},E)\,\Phi_{n',\mathbf{k}_{yz}}(\mathbf{r}')\\
\Sigma^{<}_{n,n'}(\mathbf{k}_{yz},E)&=\frac{D}{\Omega}\sum_{\mathbf{q}_{yz}}\int d\mathbf{r}\int d\mathbf{r}'\,\Phi^{\dagger}_{n,\mathbf{k}_{yz}}(\mathbf{r})\;G^{<}(\mathbf{r},\mathbf{r}',\mathbf{k}_{yz},E)\,\Phi_{n',\mathbf{k}_{yz}}(\mathbf{r}')\delta(\mathbf{r}-\mathbf{r}')\\
\Sigma^{<}_{n,n'}(\mathbf{k}_{yz},E)&=\frac{D}{\Omega}\sum_{\mathbf{q}_{yz}}\int d\mathbf{r}\,\Phi^{\dagger}_{n,\mathbf{k}_{yz}}(\mathbf{r})\;G^{<}(\mathbf{r},\mathbf{r},\mathbf{k}_{yz},E)\,\Phi_{n',\mathbf{k}_{yz}}(\mathbf{r})
\end{aligned} \tag{3.85}$$

where D is the constant coefficient in the deformation potential formulation and Ω is the volume of the unit cell. The lesser-then Green's function in the real-space can be developed in terms of the matrix element of the Green's function in the reduced space according to:

$$G^{<}(\mathbf{r},\mathbf{r},\mathbf{k}_{yz}-\mathbf{q}_{yz})=\sum_{m,m'}G^{<}_{m,m'}(\mathbf{k}_{yz}-\mathbf{q}_{yz})\,\Phi_{m,\mathbf{k}_{yz}-\mathbf{q}_{yz}}(\mathbf{r})\,\Phi^{\dagger}_{m',\mathbf{k}_{yz}-\mathbf{q}_{yz}}(\mathbf{r}). \tag{3.86}$$

By using such a simplified notation, from Eqs. (3.85) and (3.86) the lesser-than self-energy for a given $\mathbf{k}_{yz}$ in the reduced space reads out

$$\begin{aligned}
\Sigma^{<}_{n,n'}(\mathbf{k}_{yz})&=\frac{D}{\Omega}\sum_{\mathbf{q}_{yz}}\sum_{m,m'}G^{<}_{m,m'}(\mathbf{k}_{yz}-\mathbf{q}_{yz})\int d\mathbf{r}\,\Phi^{\dagger}_{n,\mathbf{k}_{yz}}(\mathbf{r})\,\Phi_{m,\mathbf{k}_{yz}-\mathbf{q}_{yz}}(\mathbf{r})\,\Phi^{\dagger}_{m',\mathbf{k}_{yz}-\mathbf{q}_{yz}}(\mathbf{r})\,\Phi_{n',\mathbf{k}_{yz}}(\mathbf{r})\\
&=D\sum_{\mathbf{q}_{yz}}\sum_{m,m'}G^{<}_{m,m'}(\mathbf{k}_{yz}-\mathbf{q}_{yz})\;\mathcal{F}^{(m',\mathbf{k}_{yz}-\mathbf{q}_{yz}),(n,\mathbf{k}_{yz})}_{(m,\mathbf{k}_{yz}-\mathbf{q}_{yz}),(n',\mathbf{k}_{yz})},
\end{aligned} \tag{3.87}$$

where we have defined the generalized form factor as

$$\mathcal{F}^{(m',\mathbf{k}_{yz}-\mathbf{q}_{yz}),(n,\mathbf{k}_{yz})}_{(m,\mathbf{k}_{yz}-\mathbf{q}_{yz}),(n',\mathbf{k}_{yz})}=\int\frac{d\mathbf{r}}{\Omega}\,\Phi^{\dagger}_{n,\mathbf{k}_{yz}}(\mathbf{r})\,\Phi_{m,\mathbf{k}_{yz}-\mathbf{q}_{yz}}(\mathbf{r})\,\Phi^{\dagger}_{m',\mathbf{k}_{yz}-\mathbf{q}_{yz}}(\mathbf{r})\,\Phi_{n',\mathbf{k}_{yz}}(\mathbf{r}) \tag{3.88}$$

For the last step, we consider the case where the off-diagonal elements of the lesser(greater)-than Green's functions and self-energies can be neglected with respect to the diagonal ones. In this case, the form factor depends only on

two indexes and the previous equations simply read:

$$\Sigma^{<,\mathrm{ac}}_{n,n}(\mathbf{k}_{yz},E)=\frac{D^2_{\mathrm{ac}}k_B\,T}{\rho v_s^2}\sum_{\mathbf{q}_{yz}}\sum_{m}G^{<}_{m,m}(\mathbf{k}_{yz}-\mathbf{q}_{yz},E)\,\mathcal{F}_{(m,\mathbf{k}_{yz}-\mathbf{q}_{yz}),(n,\mathbf{k}_{yz})}. \tag{3.89}$$

and

$$\Sigma^{<,\mathrm{op}}_{n,n}(\mathbf{k}_{yz},E)=\sum_{\nu}\frac{\hbar D^2_{\mathrm{op},\nu}}{2\rho\omega_\nu}\left[N_\nu+\frac{1}{2}\pm\frac{1}{2}\right]\sum_{\mathbf{q}_{yz}}\sum_{m}G^{<}_{m,m}(\mathbf{k}_{yz}-\mathbf{q}_{yz},E\pm\hbar\omega_\nu)\,\mathcal{F}_{(m,\mathbf{k}_{yz}-\mathbf{q}_{yz}),(n,\mathbf{k}_{yz})}. \tag{3.90}$$

3.2.D Self-consistent calculations

Finally, solving the kinetic equations is not enough to model the device accurately. Indeed, in order to describe properly the electrostatic potential of the device, the Green's function equations have to be coupled with Poisson's equation and solved self-consistently. The Poisson's equation reads:

$$\nabla_r\cdot\epsilon(\mathbf{r})\nabla_r\phi(\mathbf{r})=-\rho(\mathbf{r}) \tag{3.91}$$

where $\epsilon(\mathbf{r})$ is the dielectric permittivity, $\rho(\mathbf{r})$ is the charge density and $\phi(\mathbf{r})$ is the electrostatic potential. The electrostatic potential intervenes in the calculation of the retarded/advanced Green's functions of Eq. (3.39) since the potential is added to the Hamiltonian as:

$$\mathbf{H}=\mathbf{H}_0-e\tilde{\phi} \tag{3.92}$$

where $\mathbf{H}_0$ is the Hamiltonian without the external electric field and $\tilde{\phi}$ is the potential of Eq. (3.91) transformed in the basis where $\mathbf{H}$ is defined.

From the Green's functions, it is possible to derive the charge density with Eq. (3.47). However, it means that the carrier density, in the device, is potential dependent. This make the Eq. (3.91) a non-linear Poisson's equation. The non-linear equation must be solved self-consistently in order to describe the electrostatics of the device. Each iteration will produce a new potential that will redistribute the carrier density in the device. The new carrier density will then produce a new potential computed with the Poisson's equation and so on until the convergence of the potential is reached. The flowchart of the simulation of the device is illustrated in Fig.(3.6).

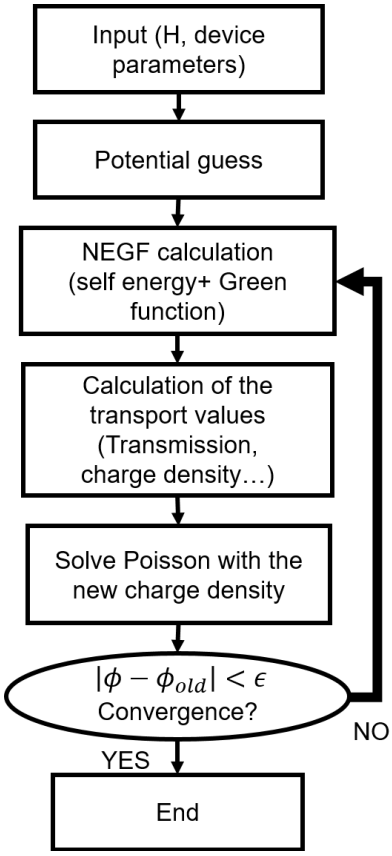

Figure 3.6: Flowchart of the quantum transport simulation with the self-consistent NEGF method.

Chapter 4

Empirical pseudopotential Hamiltonians for quantum transport in heterostructure tunnel devices

In this chapter, we report the study related to EPM based quantum transport. A methodology to address the transport through lattice matched heterojunctions with a gradual transition at the interface is presented. For this approach, we extended the previously written NEGF code [135, 136] based on an EPM Hamiltonian to treat heterostructures and we added a way to compute the gradual chemical composition change at the interface between the two materials.

The use of heterojunctions is an important technological option in, among other possible examples, bipolar transistors [75, 166], as well as in MODFETs and Tunnel FETs [12, 25]. Tunnelling devices form a promising class of devices for the next-generation electronics. As mentioned in Sec. 1.4.B, the tunnel FET is a candidate to replace the MOSFET as a low power transistor. The Esaki diode, is a PN junction with highly doped regions and, from the point of view of the fabrication process, it can be considered as a first step before the realization of a TFET. It is also of interest for applications in HF circuits as an oscillator by exploiting the negative differential resistance (NDR) observed in its I-V characteristics [48, 198]. In both devices, the current originates from the band-to-band tunnelling between the valence and conduction bands in P and N regions, respectively. An heterostructure with the favourable band alignment, corresponding to type III heterostructure, i.e. with broken gap interface (see chapter 1), is the most efficient system for TFET operation, since it provides a sharp energy transition between the source and the channel region resulting in a high ON-current due to the reduced tunnelling path. The InAs/GaSb heterostructure is a promising option for this application as the two bulk materials present a band alignment corresponding to such a broken gap configuration. As most of III-V materials, both InAs and GaSb have a low effective mass in the lowest

conduction band ($0.023m_0$ for InAs). The GaSb has also a relatively small valence band effective mass ($0.4m_0$ for the heavy holes and $0.05m_0$ for the light holes in GaSb) which is an advantage to enhance the tunneling and thus the ON-current. Moreover, the two materials are nearly lattice matched, which should make it possible to achieve a low density of defects at the interface.

However, there are clear experimental evidences that fabricated heterojunctions are not abrupt, but may instead feature a transition region extending at least over a few atomic layers [29, 72, 209], which, from the computational point of view, requires an atomistic description of the interfaces. Thus, we have chosen to develop a methodology to study the heterostructure made of InAs and GaSb with EPM.

The empirical pseudopotential method has been successfully developed for many III-V materials, which ensures an accurate reproduction of the band structure obtained from experimental data. Moreover, in the presence of heterojunctions, a gradual transition between the two materials can be conveniently modelled with EPM by using the virtual crystal approximation (VCA) [19], whose soundness has been proved also for quaternary alloys [196]. The use of the VCA in the context of plane-wave pseudopotentials is justified, even for non-local pseudopotentials, when the atoms composing the alloy belong to the same group and when the pseudopotentials use the same number of valence orbitals for each angular momentum. We argue that the same methodology could be applied directly to a DFT Hamiltonian as the VCA can also be applied to the *ab-initio* pseudopotentials within a plane-wave basis, which makes this approach more general.

The methodology presented in this chapter provides a description of either abrupt or gradual heterojunctions, where the EPM parameters are changed according to the position so as to ensure a correct alignment of conduction and valence band edges along the heterojunction. To take into account the confinement effect in ultra-scaled devices, we also present a model for a local quantum confinement operator with the EPM from Ref. [136]. The NEGF based formalism for quantum transport, which was previously developed for homojunctions [135, 136] is also extended in order to duly account for a non-homogeneous system. Two exemplary and yet technologically relevant case studies are then addressed, namely a III-V semiconductor-based Esaki diode and both n- and p-type heterojunction Tunnel FETs.

4.1 Model description

Here, we briefly describe the full-band model employed in this work for the calculation of the band-structure of GaSb and InAs and quantum transport, based on the EPM [32, 41]. The general methodology has been already described in Chapter 2.2. It is an extension of previous works in the literature [135, 136], so that we here focus on the aspects related to the treatment of heterojunctions that have been specifically developed for this work. The quantum transport model used here relies on the NEGF method formulated in a hybrid basis, $x\mathbf{K}_{yz}$, consisting of a discretized real space in the transport direction x and of plane waves in the (y,z) directions [136].

The Hamiltonian matrix inside a given semiconductor computed from EPM is written into the hybrid basis as described in section 3.2.A and takes the block tridiagonal form

$$[\mathbf{H}]_{\mathbf{k}_{yz}} = \begin{bmatrix} \mathbf{H}(x_1,x_1) & \mathbf{H}_{0,1} & 0 & 0 & \cdots & & \\ \mathbf{H}_{0,1}^{\dagger} & \mathbf{H}(x_2,x_2) & \mathbf{H}_{0,1} & 0 & \cdots & & \\ \cdots & & \ddots & & & & \\ 0 & 0 & \mathbf{H}_{0,1}^{\dagger} & \mathbf{H}(x_j,x_j) & \mathbf{H}_{0,1} & 0 & \cdots \\ & & \cdots & \cdots & \cdots & & \end{bmatrix} \tag{4.1}$$

where x_j is a discretization point in the transport direction. This Hamiltonian is parameterized with the wave vector $\mathbf{k}_{yz}$. For a 3D electron gas (i.e. with no quantum confinement in directions y, z) the size of each block in Eq. (4.1) is $M_{3D} = 2N_G/N_d$, where N_G is the number of $\mathbf{G}$ vectors in EPM calculations (which is set by the EPM cutoff energy), and N_d is the number of discretization points inside a unit cell. For a 2D or a 1D electron gas the Hamiltonian matrix still has the tridiagonal form in Eq. (4.1), but the size of the blocks is increased [136]. We used a simple second order, centered difference discretization for the kinetic energy operator, described in Chapter 3, to be able to use a mode-space basis to reduce the size of the non-local term of our Hamiltonian. In this order, the kinetic energy operator reads:

$$T(\mathbf{k}+\mathbf{G}) = -t_0 \sum_{s=x,y,z} \left[1+\cos\left(r(k_s+G_s)d\right)\right] \quad \text{with} \quad t_0 = \frac{\hbar^2}{2m_0 d^2} \tag{4.2}$$

where $d = a_0/N_d$ is the discretization step along x. We choose $N_d = 32$ which corresponds to the discretization along the transport direction needed to ensure that the bands obtained with the discretized kinetic energy operator converge to the same bands as that described with the continuous formulation. For the pseudopotential term, a cut-off energy $\mathrm{E_{cut}}$= 4.5 Ry corresponding to $N_G \approx 800$ has been chosen, which corresponds to the value needed to achieve the convergence of the band structure, for both materials, with respect to the bandgap and the different energy separations between X, L and Γ valleys from reference [42].

4.1.A Modelling of heterojunctions with the EPM method

For a homogeneous semiconductor the diagonal blocks $\mathbf{H}(x_j,x_j)$ in Eq. (4.1) depend on x_j only through the local pseudopotential $V_L(x_j, \mathbf{G}_{yz}\text{–}\mathbf{G}'_{yz})$, which is obtained via a discrete Fourier transform to move from the plane-wave basis to real space for the x direction and it is periodic of a_0 [136]. In this chapter, we studied lattice matched heterojunctions by using a virtual crystal approximation [19], whose soundness has been proved also for quaternary alloys [196]. For the alloy $(\mathrm{GaSb})_{1-x}(\mathrm{InAs})_x$, the lattice parameter is assumed to be the same as that of GaSb and

InAs and its form factor is evaluated as:

$$U_{(\mathrm{GaSb})_{1-x}(\mathrm{InAs})_x}(x_i, |\mathbf{G}|) = (1-x)\,U_{\mathrm{GaSb}}(x_i, |\mathbf{G}|) + xU_{\mathrm{InAs}}(x_i, |\mathbf{G}|) \tag{4.3}$$

where $0 \leq x \leq 1$ and U_{GaSb} is the form factor of GaSb and U_{InAs} is the form factor of InAs. In this approach, the EPM parameters $U_S(x_i, |\mathbf{G}|)$, $U_A(x_i, |\mathbf{G}|)$ change along the transport direction so as to describe either an abrupt heterojunction, or a gradual heterojunction having a transition width, W_{HJ}, of length of a few lattice constants [24]. More precisely, the Hamiltonian blocks of the two materials are first computed in the hybrid space $x\mathbf{K}_{yz}$, then they are patched at adjacent positions to form the heterojunction. The coupling Hamiltonian matrix between adjacent blocks is simply given by the spatial discretization of the kinetic energy operator [136], hence it is material independent.

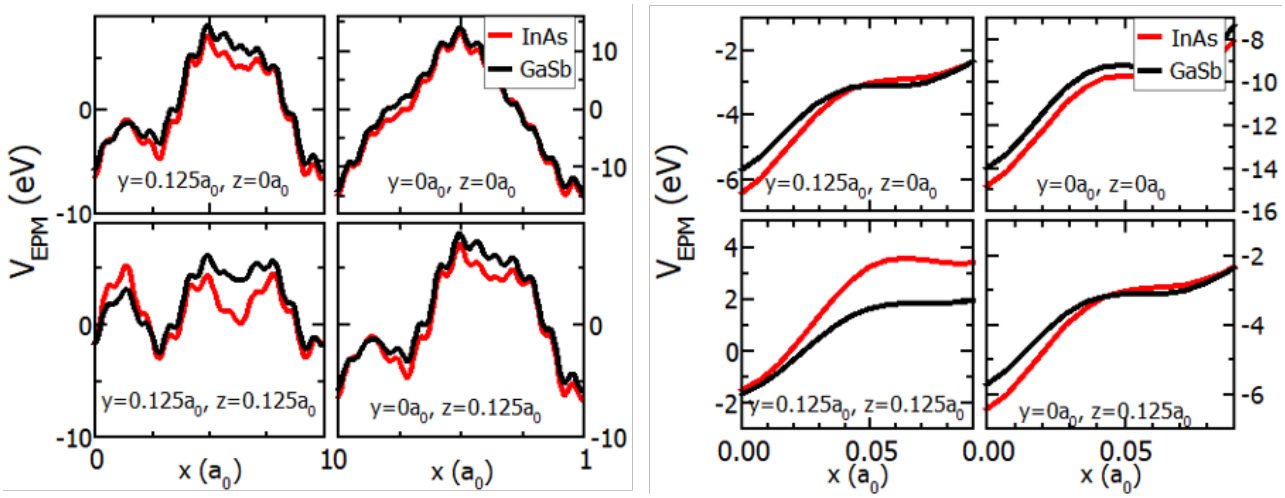

Figure 4.1: Left: real space pseudopotential plotted along the transport direction, x, inside a unit cell and for some relevant point of the transverse plane (y,z): InAs (red lines) and GaSb (black lines). Right: zoom at the edge at the edge of the unit cell.

A particular attention must be given to this approach since we started by considering the pseudopotential from a plane wave representation which induces periodic replica of the material in all three directions of space. The potential, moved to the hybrid basis, from one unit cell of one material is then patched at the interface to the potential of the second material. It implies that at the heterojunction interface a discontinuity can exist in the empirical pseudopotentials. In order to be more quantitative, we have illustrated in Fig.4.1 the pseudopotential profile along x inside a unit cell for GaSb and InAs pseudopotentials and for a few different points in the (y,z) plane. The unit cell used corresponds to the Zinc blende unit-cell of InAs discussed in Chapter 2. The left plots in Fig.4.1 report the pseudopotential profile throughout the unit cell, whereas the right plots show a zoom at the edge of the cell to better quantify the pseudopotential discontinuity. For the case studied here, however, such a discontinuity is small because GaSb and InAs have similar EPM spectral components (see Tab. 4.1). We are aware

that a more accurate description is possible by defining a supercell of the interface. It would require to compute a large Hamiltonian matrix block describing several unit cells at the interface which will end up to increase severely the computational burden. However, we argue that the pseudopotential discontinuity at the interface can be absorbed in the offset discontinuity defined with $U_S(|\mathbf{G}| = 0)$.

Going back to the description of heterojunctions via the EPM method, the correct alignment of conduction and valence band edges between the different materials must be ensured. In this regard, in Chapter 2.2, Eq. (2.10) shows that a non zero $U_S(|\mathbf{G}| = 0)$ simply results in a rigid shift of the band-structure for a given material. Consequently, in the analysis of a homogeneous system one typically sets $U_S(0) = 0$, because $U_S(0)$ can be absorbed in the choice of the energy reference.

In our modelling of heterojunctions, however, $U_S(0)$ has been changed in order to ensure the correct alignment of the conduction and valence band edges of the constituent materials, which is an approach already discussed and validated in Ref. [59]. For devices including heterojunctions, the position dependence of the local EPM parameters contributes to the x_j dependence of the diagonal Hamiltonian blocks $\mathbf{H}(x_j, x_j)$ in Eq. (4.1), as it is illustrated in the sketch of Fig. 4.2, which represents the Hamiltonian of the heterostructure in the hybrid basis.

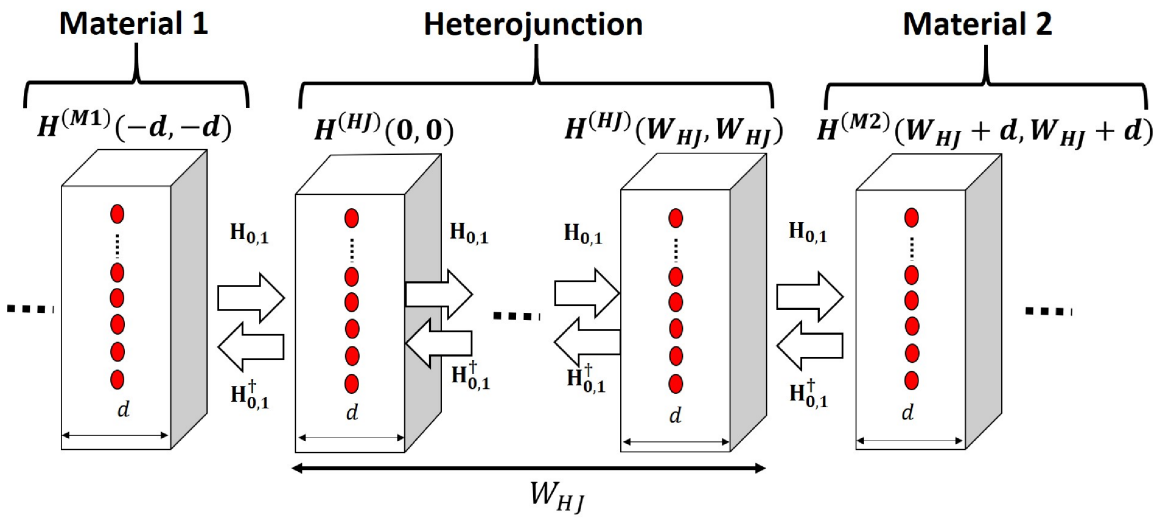

Figure 4.2: Schematic illustration of the blocks of the Hamiltonian matrix across a gradual heterojunction where, in order to simplify the notation, one end of the heterojunction has been placed at x=0. W_{HJ} is the width of the heterojunction transition region. The notation is consistent with Eq. (4.1). In the heterojunction region the EPM parameters change with the position and thus contribute to the position dependence of the diagonal Hamiltonian blocks, $H^{HJ}(x_i, x_i)$.

Table 4.1 reports the EPM parameters for GaSb and InAs used in all the calculations of this study, and Fig. 4.3 displays the bandstructure of the two materials, as well as the corresponding density of states. As it can be seen, the GaSb-InAs heterojunction provides a broken bandgap interface, in the sense that the bottom of the InAs conduction band is slightly below the top of the GaSb valence band. By varying the molar fraction x in $(\mathrm{GaSb})_{1-x}(\mathrm{InAs})_x$ the corresponding effective mass at the Γ minimum of the conduction band changes as reported in Fig. 4.4.

Table 4.1: EPM parameters (in Ry) for GaSb, InAs (Ref.[42]) and pseudo oxide (Ref.[136]) employed in this work. $U_S(0)$ is set to 0 for InAs and to 0.84 eV for GaSb.

	$U_S(3)$	$U_S(8)$	$U_S(11)$	$U_A(3)$	$U_A(4)$	$U_A(11)$
GaSb	-0.22	0	0.05	0.06	0.05	0.01
InAs	-0.22	0	0.05	0.08	0.05	0.03
Pseudo oxide	-0.64	0	0.14	0.225	0.14	0.08

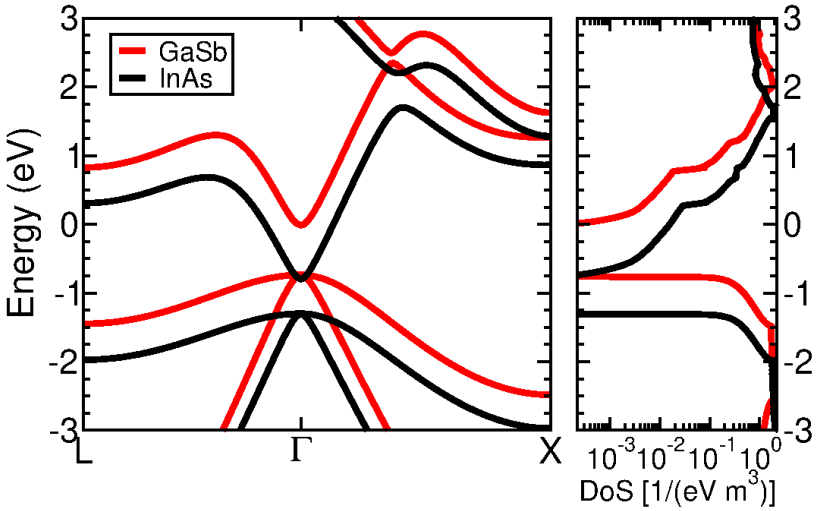

Figure 4.3: (left) Highest valance bands and lowest conduction bands for GaSb and InAs along the high-symmetry directions inside the first Brillouin zone and (right) the corresponding density of states.

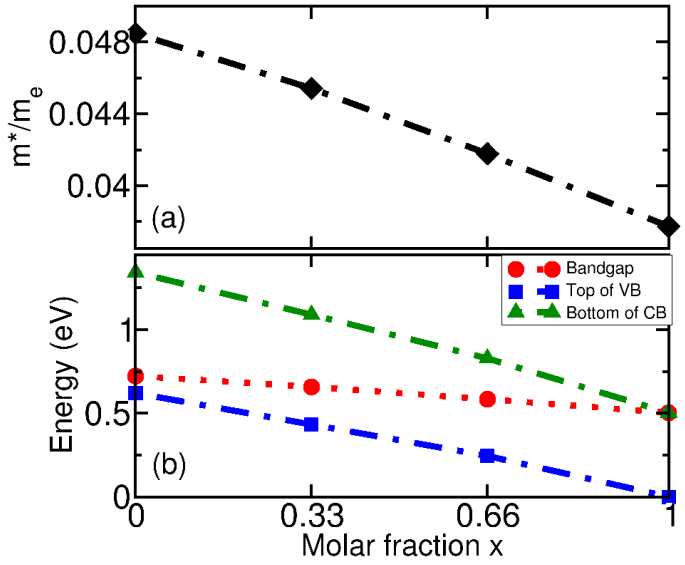

Figure 4.4: (a) Electron effective mass m^* at the Γ point of the conduction band versus the molar fraction x of the alloy $GaSb_{1-x}InAs_x$. (b) Top of valence band, bottom of conduction band and band-gap versus the molar fraction. The energy reference is the top of the InAs valence band.

4.1.B Local quantum confinement operator

For ultra-thin devices, confinement effects are not negligible. The confinement induces quantization of energy levels, which modifies the band structure of the material with respect to the bulk one. In order to model it in our description, a confinement operator has been defined. It has been developed by Marco Pala and David Esseni for the EPM Hamiltonian [135, 136].

Two definitions of this operator have been developed for EPM. First, a non-local description based on a discontinuity of the bands is possible. In this model, the valence bands are shifted to lower energies and the conduction bands are shifted to higher energies in the regions outside the semiconductor. However, as mentioned in the Chapter 3, the non-locality of the operator would prevent the use of a mode-space basis.

The second approach developed for the use of a mode-space basis, consists in defining a quantum confinement operator, which is local in real space and it is non null only outside of the confined material where a pseudo-oxide material is defined. The EPM parameters of the pseudo-oxide presented in Tab. 4.1 are tailored so as to obtain the appropriate discontinuity at the semiconductor-oxide interface of the conduction band minimum and valence band maximum, which confines the carriers in the semiconductor region. This pseudo-oxide can be seen as a computational tool to introduce appropriate band discontinuity outside of the confined material. For simplicity, it is computed with the same atomic lattice and structure. Accordingly, the pseudopotiential of a confined system in one direction(i.e. z axis), V_{2D}, is defined in real space as:

$$V_{2D}(\mathbf{r}) = V_{sc}(\mathbf{r}) + V_{cnf}(\mathbf{r})\theta_{2D}(z) \tag{4.4}$$

where V_{sc} is the pseudopotential from the bulk semiconductor used for the confined material and we define V_{cnf} from the pseudopotential of the pseudo-oxide V_{ox} as:

$$V_{cnf}(\mathbf{r}) = V_{ox}(\mathbf{r}) - V_{sc}(\mathbf{r}) \tag{4.5}$$

We also define θ_{2D} as step function that separates the regions of the oxide and the semiconductor and it reads

$$\theta_{2D}(z) \begin{cases} 0 \text{ for } |z| < \frac{t_{sc}}{2} \\ 1 \text{ otherwise} \end{cases} \tag{4.6}$$

where t_{sc} is the thickness of the semiconductor layer, i.e. of the quantum well. The pseudopotential V_{2D} is then moved to the reciprocal space with a Fourier transform and by using the local nature of this potential, it can be

written as:

$$V_{2D}(\mathbf{K}-\mathbf{K}') = V_{sc}(\mathbf{G}-\mathbf{G}')\delta_{\mathbf{k},\mathbf{k}'} + \sum_{G_z''} V_{cnf}(\mathbf{G}_{xy}-\mathbf{G}'_{xy}, G_z - G_z' - G_z'')\theta_{2D}(K_z - K_z' + G_z'')\delta_{\mathbf{k}_{xy},\mathbf{k}'_{xy}} \tag{4.7}$$

where $\mathbf{G} = (\mathbf{G}_{xy}, G_z)$ and $\mathbf{K} = \mathbf{k} + \mathbf{G}$. This potential replaces the previous pseudopotential in the EPM Hamiltanian which writes:

$$H_{\mathbf{k}_{xy}}(\mathbf{K}-\mathbf{K}') = T(\mathbf{k}+\mathbf{G})\delta_{k_z,k_z'}\delta_{\mathbf{G},\mathbf{G}'} + V_{2D}(\mathbf{K}-\mathbf{K}') \tag{4.8}$$

4.1.C Quantum transport

Since the current in the studied devices originates from band-to-band tunneling, a formalism that rigorously incorporates quantum phenomena is required. Quantum transport calculations are thus addressed by means of the NEGF method in the ballistic approximation, described in Chapter 3, and solved self-consistently with the Poisson's equation to account for the overall device electrostatics.

In order to efficiently compute the set of kinetic equations, one obtains a substantial reduction of the size of the blocks in Eq. (4.1) by employing a mode-space transformation [195], as described in Chapter 3. Differently from the homogeneous case [136], the modes along the heterojunction direction change with the position x_j, which is duly accounted for in the transformation of the Hamiltonian into the mode space, where each block is transformed into the mode-space basis constructed from the corresponding material.

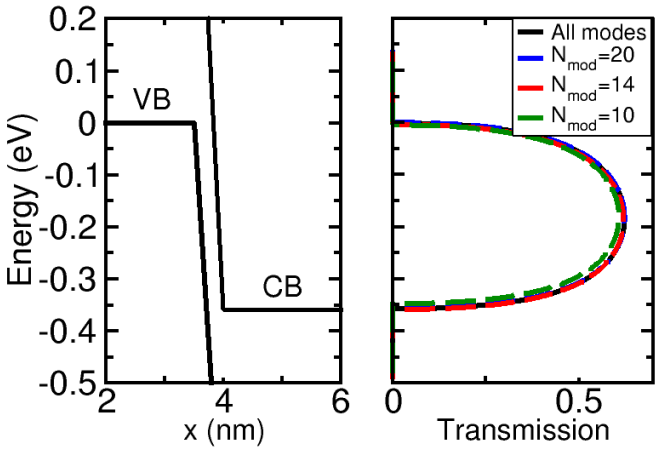

Figure 4.5: (Left) Conduction and valence band profile along the Esaki diode for a flat band calculation. (Right) Corresponding the transmission at $k_y = k_z = 0$, versus the energy and for different numbers of modes used for the reduced basis. The bias is $V_D = 0.35$ V.

For the III-V semiconductor-based Esaki diode and heterojunction Tunnel FETs of this study, it has been found that, by selecting only the lowest energy transverse modes in each device section (here, the unit cell of the material

is divided in $N_d = 32$ sections), the size of the Hamiltonian blocks could be vastly reduced compared to the blocks with a size of 50×50 in Eq. (4.1) [136]. Actually, even in heterojunction devices, 14 modes were sufficient to achieve results that agree within a few percent with those obtained without introducing the mode-space reduction. We prove it by plotting in Fig. 4.5 the transmission obtained in an Esaki diode for different numbers of modes. It can be emphasized that with 14 modes, the results fit accurately the transmission computed by taking all the modes into account. However, if we reduce the size of the basis to ten modes, the transmission begins to be slightly underestimated. The cases $N_{mod} = 8$ and less are not shown in this figure since the first eight modes correspond to the valence bands. It results that the conduction bands are not described with such a basis.

4.2 Self-consistent device simulations

In this section we report exemplary simulation results obtained with the EPM-based methodology developed for two tunnel devices employing a GaSb-InAs heterojunction. All simulations were carried at room temperature and for the transport direction along [100].

4.2.A Esaki tunnelling diode

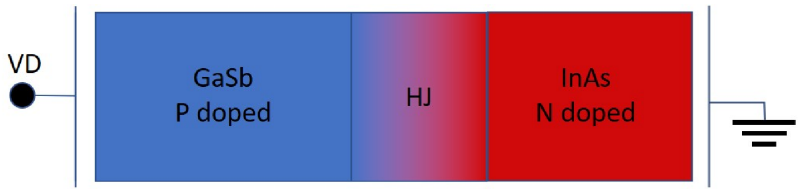

Figure 4.6: Sketch of the Esaki diode where red is used for InAs and blue is used for GaSb

The simulated Esaki diode is sketched in Fig. 4.6 and consists of a p-type GaSb region and an n-type InAs region. The doping of the p- and n-type regions was set to N_A=10^{20} cm^{-3} and N_D=3×10^{19} cm^{-3}, respectively, which results in a degenerate carrier gas in both regions. We assumed periodic boundary conditions along the transverse directions and sampled the wave vectors k_y and k_z with a constant step of $\Delta k_y = \Delta k_z = 0.05 \times 2\pi/a_0$ and with $a_0 = 6.068$ Å.

The current per unit area, J_D, versus the applied bias V_D of the Esaki diode is shown in Fig. 4.7 for different widths, W_{HJ}, of the heterojunction transition region. In the forward bias an NDR is observed, which is more clearly visible in the right panel of the figure that focusses on the forward bias region. As it can be seen the peak current in the NDR region decreases for more gradual heterojunctions (i.e. for larger W_{HJ}), and it is maximum for the abrupt case. A similar dependence of J_D on W_{HJ} is observed also in the reverse bias region, where the current is still completely dominated by BTBT in the bias range explored in Fig. 4.7. In contrast, Fig. 4.8 presents the valley current of the NDR region in log scale and shows a small increase for more gradual heterojunctions. However, it is

important to note that our values for the valley current are significantly underestimated due to the absence of the electron-phonon interaction and traps that are known to be responsible for the recombination current [142].

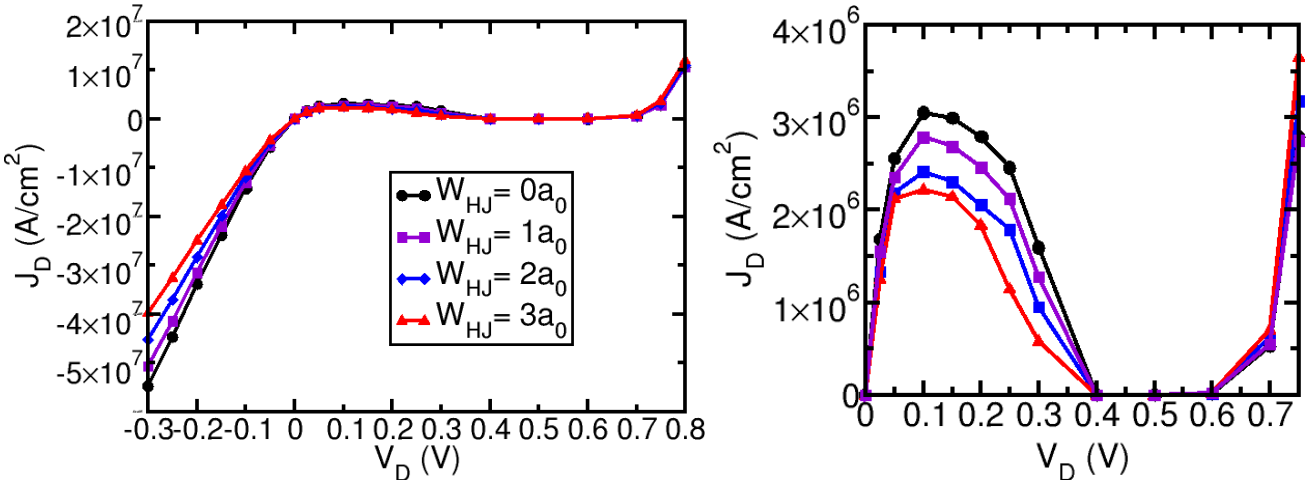

Figure 4.7: (Left) Current density, J_D, versus the applied bias V_D, for different widths, W_{HJ}, of the heterojunction transition region. (Right) shows a focus on the negative differential resistance region.

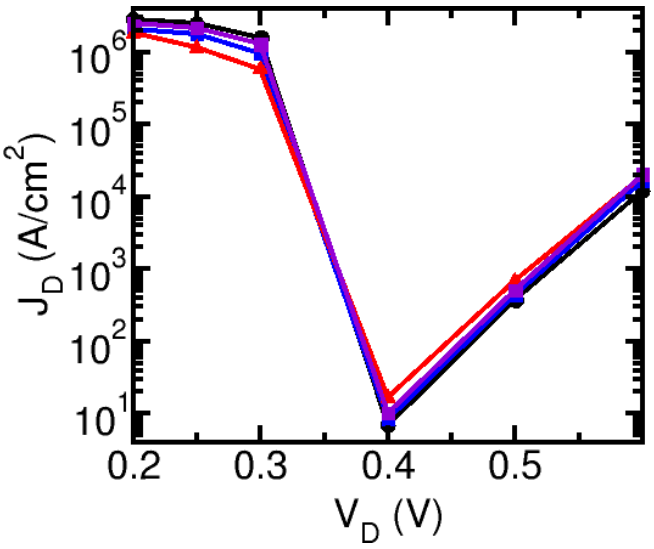

Figure 4.8: Zoom on the valley current density for different widths, W_{HJ}, of the heterojunction transition region.

Figure 4.9 shows the conduction and valence band profiles along the device for V_D=0.1 V (i.e. approximately at the peak current in the NDR region), as well as the energy resolved current density J_D. As expected, J_D takes the highest values in the narrow Fermi energy window defined as the interval limited by the Fermi levels $E_{f,n}$ and $E_{f,p}$ of the n- and p-region, respectively. Moreover, the narrower the transition region W_{HJ} is, the shorter is the BTBT path, which leads to the corresponding J_D increase in Fig. 4.7.

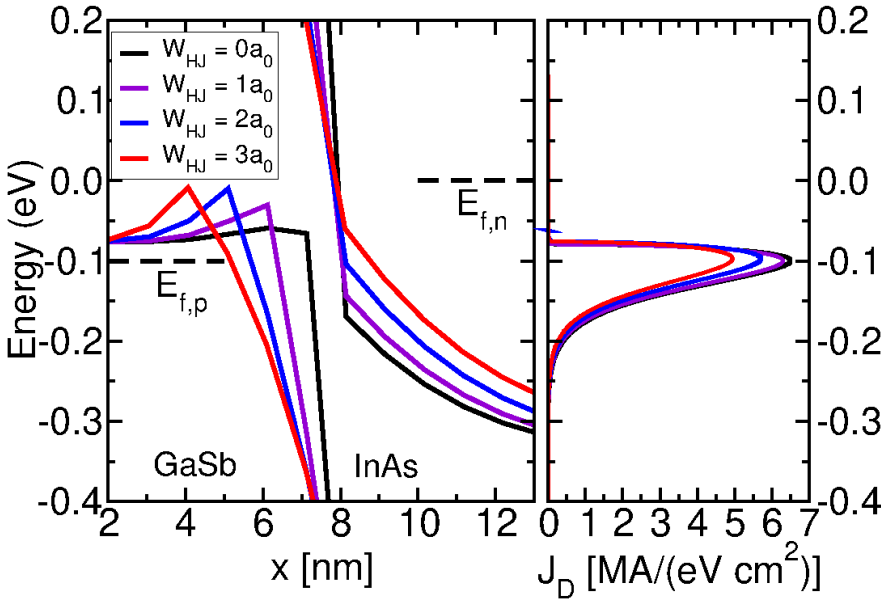

Figure 4.9: (Left) Conduction (in InAs) and valence (in GaSb) band profile along the Esaki diode for different widths W_{HJ} of heterojunction transition region. The source Fermi level, $E_{f,n}$ in the n region is taken as the energy reference. (Right) Corresponding current spectrum at $k_y = k_z = 0$, J_D, versus the energy and for different W_{HJ}. The bias is $V_D = 0.1$V

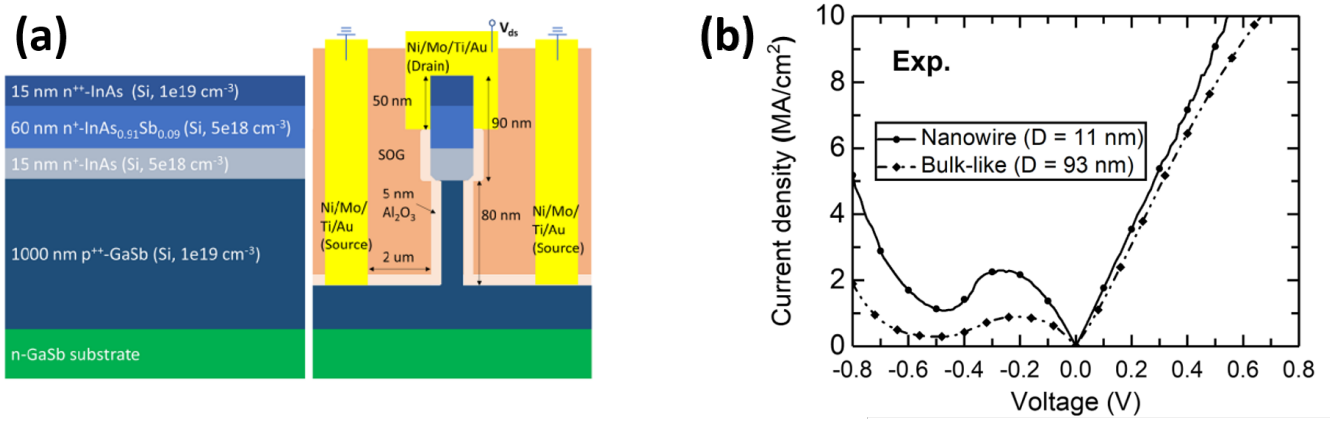

Figure 4.10: (a) Sketch of an experimental Esaki diode. (b) Current density versus the applied bias. Experimental data taken from Ref. [169]

. Note that in this Figure the reverse bias regime corresponds to positive voltage values.

For comparison, Shao et al. [169] reported experimental measurements on a III-V Esaki diode made of a vertical nanowire and its bulk-like counterpart made of InAs/GaSb as schematized in Fig. 4.10.(a). Quite interestingly, their experimental conditions are close to our simulation with a slightly lower doping of $N_D = 10^{19}$ cm^{-3} and $N_A = 5 \times 10^{18}$ cm^{-3} than the one we used. The current densities measured for the bulk and the nanowire devices are shown in Fig. 4.10.(b). Our simulation results of current density fit quite well these experimental data in the reverse bias region with a current density slope of 17 MA/(Vcm^2) for the experimental curve, while our results give a current

density slope in the range 14-20 MA/(Vcm^2). In the forward bias, Shao et al. measured a peak current density of 2.3 MA/cm^2 for the nanowire and 1 MA/cm^2 for the bulk-like device, which is in the range of our simulation result (2-3 MA/cm^2). However, the peak-to-valley current ratio cannot be compared due to our ballistic approximation and principally to the absence of trap assisted tunnelling [165] in our calculations which is expected to reduce it to a more realistic value.

4.2.B Ultra-thin body, heterojunction Tunnel FETs

Nanoscale FETs are often fabricated by using ultra-thin semiconductor films or nanowires, which can induce significant quantum confinement effects in the resulting low dimensional electron gas, as it is the case for the TFETs simulated in this section. The modelling of quantum confinement in this EPM based formalism has been described previously and has been verified in Ref. [136] with the non-local method for confinement described in Sec. 4.1.B. For both the n-type and p-type TFETs, we used a physical oxide thickness of 1.5 nm and a relative dielectric constant $\kappa = 10$, corresponding to an equivalent oxide thickness of 0.6 nm. The transverse wave vector k_y was sampled with a constant step of $\Delta k_y = 0.05 \times 2\pi/a_0$. In this work we used the pseudo-oxide EPM parameters proposed in Ref. [136] and reported in table 4.1, which results in a conduction and valence band discontinuity between InAs and the pseudo-oxide of respectively 3.8 eV and 4.5 eV.

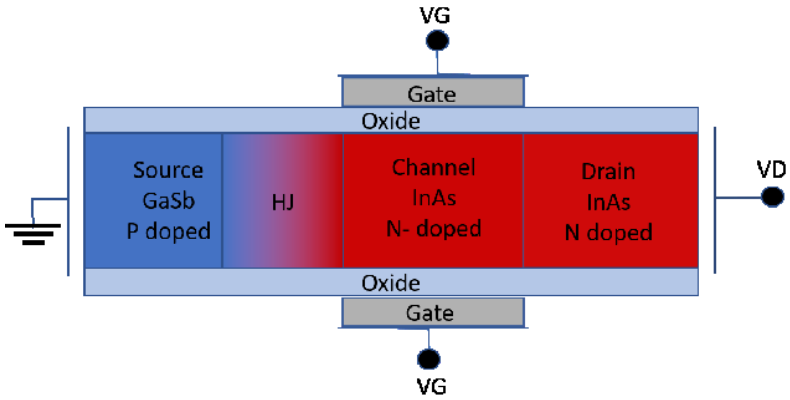

Figure 4.11: Sketch of the simulated n-type HJ TFET.

Table 4.2: Device parameters for the heterojunction n-type Tunnel FET schematized in Fig. 4.11. T_{ch}, L_G, $L_{S/D}$, N_A, N_D denote the channel thickness, the gate length, the source/drain length and the dopant concentrations in the source and drain region, respectively.

	T_{ch} [nm]	L_G [nm]	$L_{S/D}$ [nm]	N_A [cm^{-3}]	N_D [cm^{-3}]
GaSb/InAs	3.0	16	14	10^{20}	10^{19}

A sketch of the n-type heterojunction Tunnel FET is shown in Fig. 4.11 and the main device parameters are

reported in Tab. 4.2. Figure 4.12 reports the current versus gate voltage characteristics of the TFET for different widths, W_{HJ}, of the GaSb-InAs heterojunction and for $V_{DS} = 0.3$ V. For both the n- and p-type TFETs, the metal gate workfunction was adjusted so that the OFF-current (i.e. the I_D at $V_{GS} = 0$ V and $V_{DS} = 0.3$ V) is approximately 1 nA/μm. The transistor exhibits a steep sub-threshold swing around 30 mV/dec, which is an expected result for a device with no defects and with a semiconductor-oxide interface free of traps and microscopic roughness [46, 45].

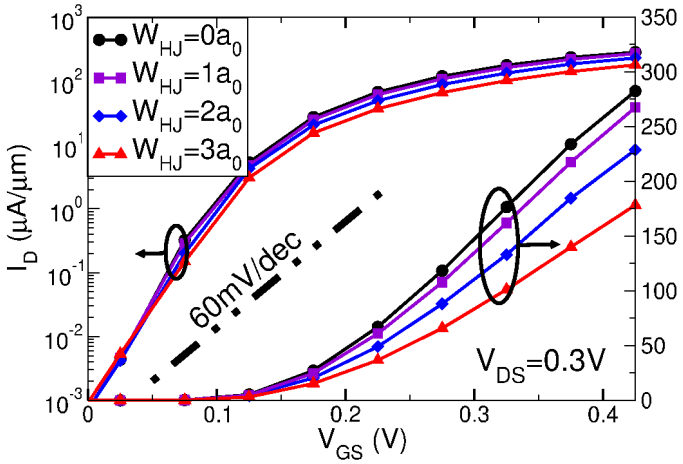

Figure 4.12: Drain-current versus gate-voltage characteristics at V_{DS}=0.3 V for different widths, W_{HJ}, of the heterojunction transition region. The metal gate workfunction was set to 4.625 eV.

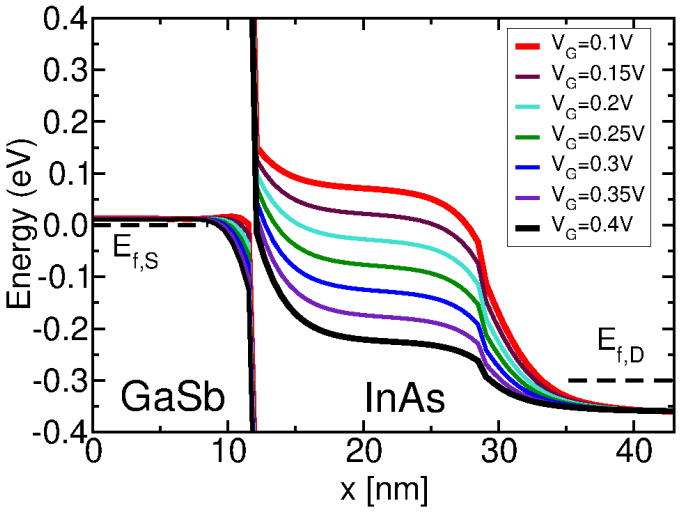

Figure 4.13: Conduction (in InAs) and valence (in GaSb) band profile at different V_{GS} and Fermi level in the source region $E_{f,S}$ and drain region $E_{f,D}$ for an abrupt heterojunction GaSb-InAs n-type tunnel FET.

The ON-current reported here is relatively high for a lateral TFET. Figure 4.13 shows the conduction and valence band profile along the device for different V_{GS}. It presents a sharp energy transition at the interface that is mainly due to the broken-gap configuration of this heterostructure. It explains the high value of current achieved as it reduces the tunnelling path between the valence and conduction bands. The low effective mass of the materials

also helps in achieving a high current. Quite interestingly, the width W_{HJ} of the GaSb-InAs heterojunction has no sizeable influence on the subthreshold current and swing, but it does affect the ON-current, that is in fact degraded when the transition region is wider.

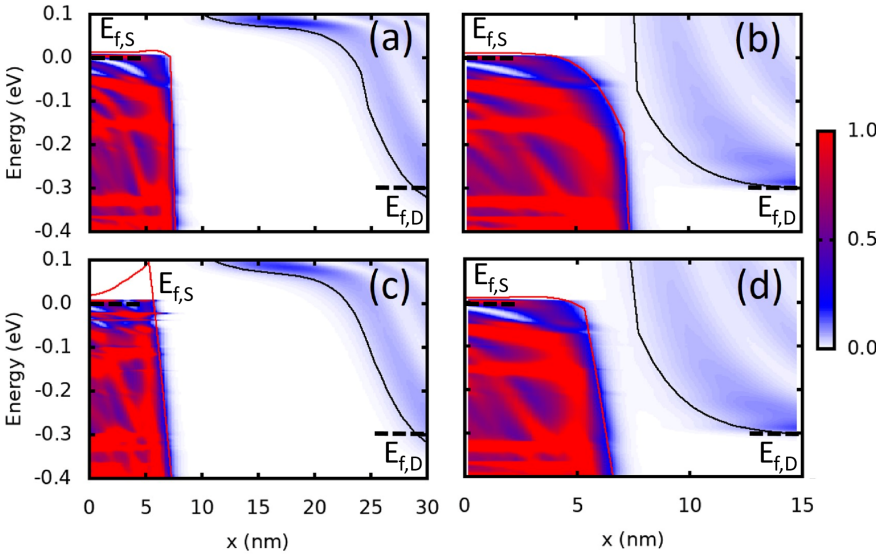

Figure 4.14: Local density of states, LDOS, for the T-FET simulated in Fig.4.11. (a),(c) report the LDOS in the OFF-state for V_{GS}=0.0 V and V_{DS}=0.3 V; (b),(d) report the LDOS in the ON-state for V_{GS}=0.4 V and V_{DS}=0.3 V. Plots in (a) and (b) correspond to an abrupt heterojunction (i.e. W_{HJ}=0), while plots in (c) and (d) correspond to W_{HJ}=3a_0.

Figure 4.14 presents the LDOS for the TFET in the ON-state and the OFF-state and for two different transition lengths. It shows that the different influence of W_{HJ} on the OFF and ON-currents is again related to the modulation of the band-to-band-tunnelling path. In the OFF-state, actually, the impact of W_{HJ} on the relatively long BTBT path is modest, whereas in the ON-state the relative importance of W_{HJ} is larger because the tunnelling path is much shorter.

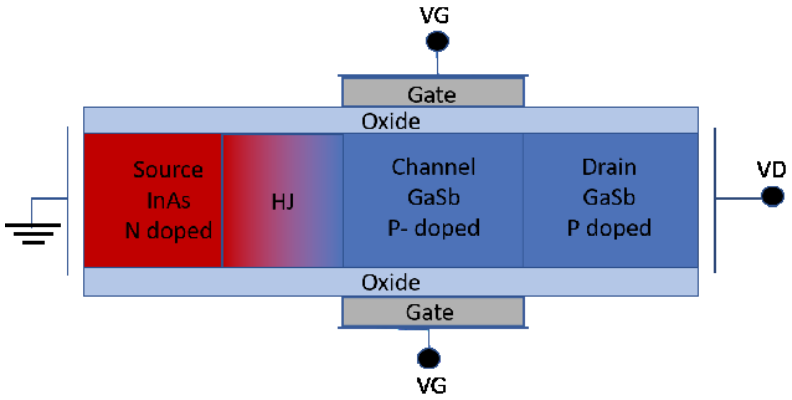

Figure 4.15: Sketch of the simulated p-type HJ TFET.

Table 4.3: Device parameters for the heterojunction p-type Tunnel FET simulated in Fig. 4.15.

	T_{ch} [nm]	L_G [nm]	$L_{S/D}$ [nm]	N_A [cm^{-3}]	N_D [cm^{-3}]
InAs/GaSb	3.7	16	14	5×10^{19}	2×10^{19}

By leveraging the inherently full-band nature of our EPM based modelling approach, we also investigated a p-type heterojunction Tunnel FET, whose sketch is shown in Fig. 4.15 and whose parameters are reported in Tab. 4.3. Due to the fairly low conduction band density of states of InAs, the source doping N_D=2×10^{19} cm^{-3} is large enough to induce a strongly degenerate electron gas in the source. Such a doping concentration is fairly consistent with the value recommended in Ref. [194], which was the outcome of an optimization of the tunnelling junction operation.

However, the current-voltage characteristics in Fig. 4.16 show that the p-type Tunnel FET can hardly achieve a sub-threshold swing below 60 mV/dec, with a degraded minimum swing compared to the n-type counterpart in Fig. 4.12.

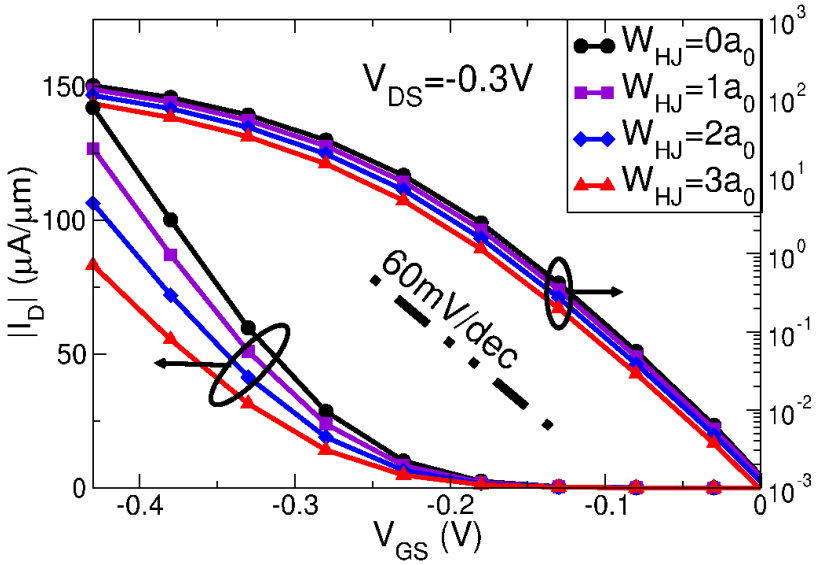

Figure 4.16: Drain-current versus gate-voltage characteristics for different widths, W_{HJ}, of the heterojunction transition region. The metal gate workfunction was set to 4.665 eV.

To explain this behaviour, in Fig. 4.17 we show the conduction and valence band along the p-type TFET for different V_{GS}. In the source region, we observe a relatively large screening length of the electric field produced by the gate bias (compared to the p-type source of the n-type Tunnel FET), which in turn degrades the sub-threshold swing [114, 56]. The relatively low value of 148 μA/μm for the ON-current is not unexpected because of the low density of states in the source region. A difference of channel thickness between the n-type and p-type transistors may be responsible of the lower p-type performances due to a degraded electrostatic control. In Fig. 4.17, It can also be observed that the high doping results in a strong degeneracy in the source. It is responsible of the lowering of the ON-current as the filtering action of the TFET is not very efficient because electrons belonging to the Fermi

tail can tunnel from the source to the channel. Hence, reducing the doping in the source to achieve a subband profile close to the Fermi level is expected to increase the sub-threshold swing. As for the influence of the width W_{HJ} of the heterojunction transition region on the current-voltage characteristic of the device, the main features observed for the n-type Tunnel FET are confirmed. Namely, W_{HJ} weakly affects the sub-threshold region, where the tunnelling distance is substantially larger than W_{HJ}, whereas the ON-current of the device is progressively reduced when increasing W_{HJ}. In the ON-state, actually, a gradual transition region for the heterojunction results in a sizeable lengthening of the tunnelling path.

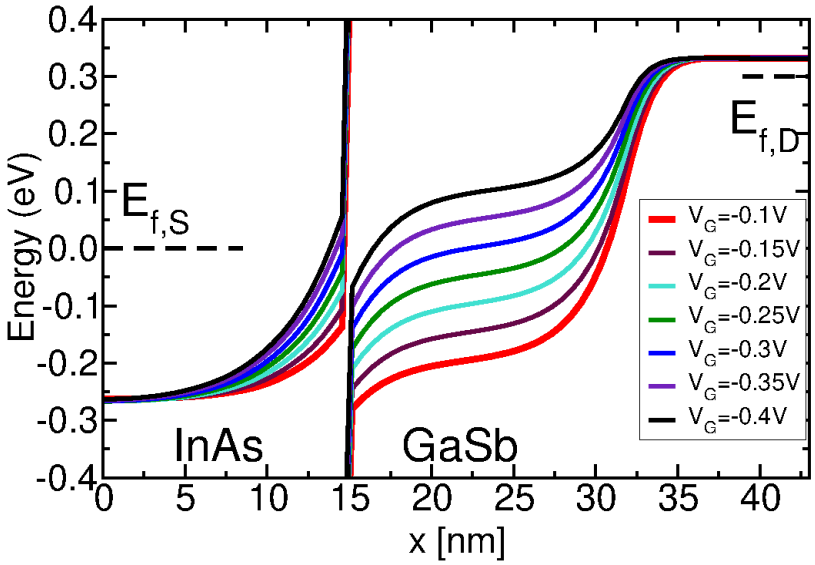

Figure 4.17: Conduction (in InAs) and valence (in GaSb) band profile at different V_{GS} for an abrupt HJ GaSb-InAs p-type tunnel FET.

Similarly to the Esaki diode, Shao et al. reported results obtained for a vertical nanowire TFET made of GaSb/InAsSb [168]. The device showed poor performances in comparison to our simulation with an ON-current of 47 μA/μm and a minimum sub-threshold swing of 252 mV/dec at room temperature. The measurements were done at different temperatures and a SS of 80 mV/dec was found at 77 K which supports the hypothesis that the degradation comes from trap-assisted tunnelling. The absence of defects and scattering phenomena in our simulation may probably explain the discrepancy observed with experiments.

4.3 Summary

We have presented the implementation and application of a full-band quantum transport methodology for nanoscale electron devices with heterojunctions. The model was developed in the framework of the empirical pseudopotentials method, and the intermediate region between the two materials was described according to a virtual crystal approximation, which allowed us to investigate the influence on the device characteristics of a gradual interface between the constituent materials. We then leveraged the inherently full-band nature of the model and reported device simula-

tion results for an n- and a p-type, ultra-thin body heterojunction Tunnel FET, as well as for an Esaki diode. The interface at the heterojunction was varied from the abrupt case to a more realistic transition width extended up to a few unit cells. The tunnelling current in the Esaki diode was found to be sensitive to the width of the heterojunction both in reverse bias and in the negative differential resistance region. In the GaSb-InAs Tunnel FETs the transition width affects the current only in the on-state. In the sub-threshold region, in fact, the effective tunnelling path is much larger than the transition length and thus it is not appreciably modulated by this length.

The results of this work qualify our methodology based on the empirical pseudopotentials and the NEGF formalism as a viable approach for the modelling and design of homojunction and heterojunction nanoscale electron devices. The comparison with experimental measurements emphasizes the accurate description of the BTBT. However, it shows also the need to implement the appropriate sources of carrier scattering in the model to properly simulate the device performance in some particular regimes, such as the valley current in Esaki diodes and the on-current in transistors. The implementation of electron-phonon scattering in our transport model will be addressed in Chapters 6 and 7.

Chapter 5

Lateral heterostructure treatment for quantum transport with *ab-initio* Hamiltonians

In chapter 4, we presented a model to simulate electron transport through heterostructures with a local empirical pseudopotential Hamiltonian. The EPM Hamiltonian was conveniently described in a hybrid basis considering real space along the transport direction and plane waves along transverse directions. By considering a DFT Hamiltonian described in an analogous basis, the similarity between both formalisms makes it possible to extend the transport simulation method developed in Chapter 4 to DFT Hamiltonians. The main difference between the two Hamiltonians comes from the non-locality of DFT Hamiltonians, which makes the coupling between two unit-cells material-dependent. The determination of this coupling term for the interface in heterostructures has to be addressed rigorously. This chapter focuses on quantum transport calculations in the ballistic approximation for heterostructures based on 2D materials and more precisely on the assessment of the coupling Hamiltonian describing the interface between two materials.

2D materials, as a class of new recently discovered materials, are not as well-known as the conventional bulk semiconductors such as Si, Ge or GaAs. As mentioned in Chapter 1, 2D materials comprise numerous promising candidates for the next generation of electronic devices. For what concerns electronic applications, devices based on these materials present an optimal electrostatic control of the channel due to their thinness. Some of them even possess highly anisotropic effective masses like black phosphorus [37, 113] with a low effective mass along one direction($0.17m_0$) and a large one along the transverse directions ($1.12m_0$), which allows large ON-current values. It is thus natural that lateral heterostructures based on lattice-matched 2D materials are considered as a promising option to design efficient electron devices such as MOSFETs [86], tunnel-FETs [39] and energy-filtering FETs [121].

Experimentally, lateral heterostructures of 2D materials have already been achieved by several groups validating the global growing interest for them [110, 207, 36, 67]. To scan the huge number of combinations achievable with 2D materials, an efficient predictive framework to select the best candidates is highly desirable to help experimental research. In this regard, a high precision in the description at the atomistic level of the materials is needed. Thus, in order to rigorously describe the transport through a 2D material heterostructure, due to the lack of experimental data of physical characterization of the interfaces, a full *ab-initio* approach based on DFT and NEGF methods is required. The methodology to simulate a heterostructure from density-functional theory will be applied to describe 2D material heterostructures, but we note that it can easily be extended to 1D and 3D materials.

As mentioned previously, the framework to construct the Hamiltonian of a heterostructure from the Hamiltonian of the isolated material that we used with the EPM one presented in Chapter 4 can be extended to the DFT Hamiltonians. The only difference is in the coupling matrix at the interface between the two materials that remains unknown. In this chapter, we present two methodologies to assess this coupling matrix between the materials composing the heterostructure. The first method approximates such a matrix with the coupling matrix obtained from one of the isolated materials. The second, more rigorous, consists in extracting the matrix from the Hamiltonian of a supercell containing the heterostructure. To assess these two methods, we focus on a lateral heterostructure MOSFET based on a bi-layer/mono-layer $PtSe_2$ structure and a tunnel-FET based on a $PdSe_2/SnS_2$ heterostructure. These two cases will be presented and discussed in Sec. 5.2.

5.1 Methodology

In this section, we present the DFT calculations for materials and the two methods used for the description of the coupling matrix between two materials for the NEGF transport in a heterostructure.

5.1.A DFT calculations

The Hamiltonians used in this study come from DFT calculation realised with the Quantum ESPRESSO suite [65]. The functional for the exchange-correlation term is chosen as the generalized gradient approximation by using the Perdew-Burke-Ernzerhof (PBE) functional. For each material, the cut-off energy is chosen large enough for the convergence of the band structures with $E_{cut} = 60$ Ry and all the atomic positions are relaxed until the forces acting on the atoms are below only 10^{-3} eV/Å. Simulating 2D materials on a plane-wave basis is cumbersome due to the periodic repetition of the unit cell along all directions induced by the basis. The strategy to avoid the coupling between the 2D materials with its replica in the vertical direction is to add enough vacuum in the unit cell along this direction. For layered materials, all layers interact with each other through Van der Waals (vdW) forces and it is thus important to describe this interaction accurately, which is not easy since it comes from nonlocal and long-range electron correlations. Standard DFT is not able to accurately describe this dispersive interaction because of the

approximation on the exchange correlation functional. The introduction of non-local exchange-correlation for vdW forces allows to provide a better description of this interaction. Consequently, we have considered the non-local van der Waals functional vdW-DF3 [31] for vertical heterostructures of 2D materials.

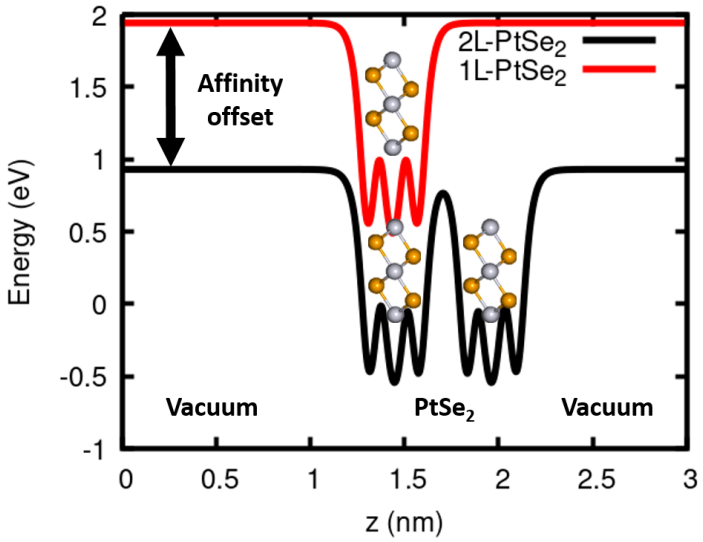

Figure 5.1: Energy difference between the ions potential + Hartree potential along the out of plane direction of the unit-cell simulated with DFT minus the lowest unoccupied energy level for monolayer(1L) $PtSe_2$ and bilayer (2L) $PtSe_2$.

In order to describe the heterostructures Hamiltonian, since we use DFT within a plane-wave basis, a procedure similar to the one adopted in Chapter 4 for the EPM case can be employed. The periodic replica of the unit cell along all directions prevents us, for computational reason, from using directly the whole Hamiltonian of a supercell with the interface for the heterostructure. The supercell must be large enough to ensure the vanishing of the interaction of the interface in the last unit cells of the the supercell. It is required to connect the supercell with the isolated material present in the rest of the device. Thus we propose to construct the heterostructure Hamiltonian from the DFT Hamiltonian of the isolated materials constituting the heterostructure. The band alignment is accounted by including a rigid energy shift between the two materials. It is added directly to the diagonal part of their Hamiltonian. The band offset between the two materials is estimated as the affinity difference between the two materials. The affinity is computed as the difference between the energies of the vacuum and of the lowest unoccupied energy level. The lowest unoccupied energy level is easily determined from the band structure calculation. To estimate the vacuum energy, we take advantage of the presence of vacuum in our unit-cell to plot the energy of the ions plus the Hartree potential in the unit-cell. Then in the vacuum region of the unit-cell, the value found corresponds to the vacuum energy as illustrated by Fig. 5.1 showing the ion energy plus the Hartree potential energy for the unit-cell of a monolayer $PtSe_2$ and a bilayer $PtSe_2$. For each material, we can identify the potentials generated by the three ions for each layer of $PtSe_2$ simulated. The vacuum energy is identified far from the material potential in the flat-potential region. It is then possible, from the lowest conduction band energy obtained with the band structure calculation, to

define the affinity offset between the two materials by taking the vacuum energy as reference.

For the transport calculation, we first compute the reduced basis Hamiltonian for the two isolated materials (in the case of Fig. 5.1 the 1L $PtSe_2$ and the 2L $PtSe_2$) with the method explained in Sec. 3, then we determine the coupling term. We start from DFT Hamiltonians that are transformed from the plane-wave basis into the hybrid basis that considers the real-space along x and plane waves in the transverse directions. This transformation has been detailed in Sec. 3.2.A.

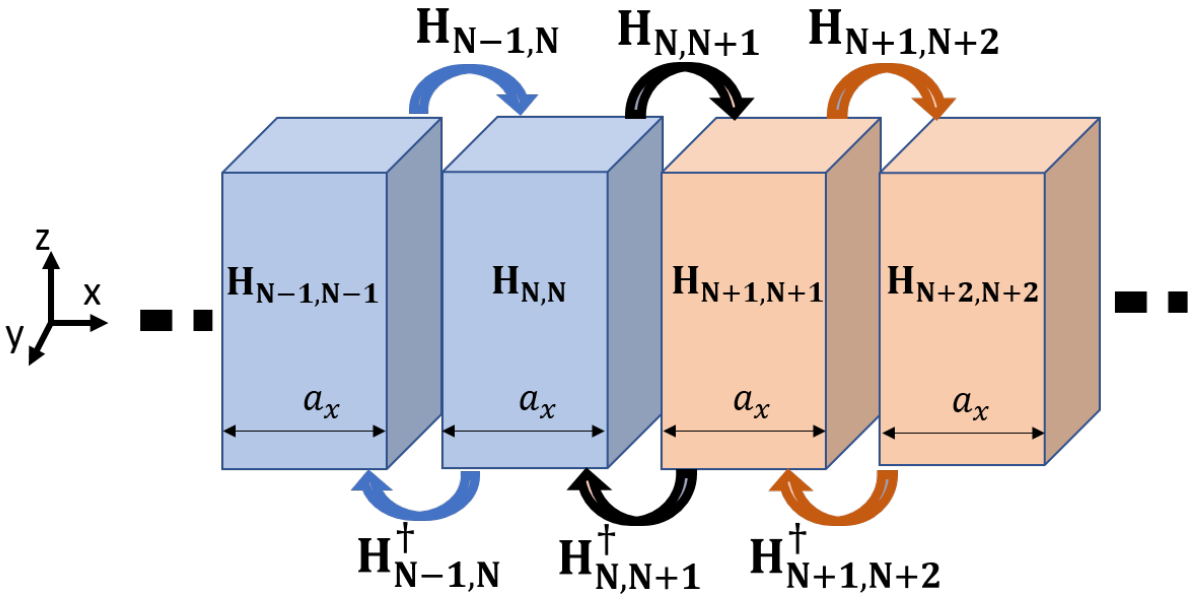

Figure 5.2: Sketch of the representation of the heterostructure Hamiltonian, into the hybrid basis, at the interface between the two materials

Figure 5.2 represents the Hamiltonian of the heterojunction, in the hybrid basis, close to the interface. Each block represents the diagonal block matrix and the arrows represent the coupling matrix between two unit cells. This Hamiltonian is block tridiagonal as shown in the matrix in Eq. (5.1) with off-diagonal blocks describing the coupling with first-neighbour unit-cell and each diagonal block representing the local interaction inside one unit-cell. It is important to remark that the non-locality of the DFT Hamiltonian in the hybrid basis is limited to the adjacent unit cell. Physically, this is due to the relatively short range of the non-local component of the pseudopotential.

$$\mathbf{H}([x, \mathbf{k}_{yz} + \mathbf{G}_{yz}], [x, \mathbf{k}_{yz} + \mathbf{G}'_{yz}]) = \begin{bmatrix} \ddots & \mathbf{H}_{N-1,N} & 0 & \dots \\ \mathbf{H}^{\dagger}_{N-1,N} & \mathbf{H}_{N,N} & \mathbf{H}_{N,N+1} & 0 \\ 0 & \mathbf{H}^{\dagger}_{N,N+1} & \mathbf{H}_{N+1,N+1} & \mathbf{H}_{N+1,N+2} \\ \vdots & 0 & \mathbf{H}^{\dagger}_{N+1,N+2} & \ddots \end{bmatrix} \tag{5.1}$$

We can see that to construct this Hamiltonian, it is required to determine two diagonal block matrices, one for each material, and three coupling matrices, i.e. two for the coupling of each material with itself and one for the coupling

between the two materials. When the Hamiltonians for the single materials are obtained, it is possible to use them to reconstruct the heterostructure as presented by the sketch in Fig. 5.2. By assuming that the non-local terms of the DFT Hamiltonian for the heterostructure present a short-range nature, which makes them negligible beyond the first neighbour, all the blocks needed to construct this Hamiltonian may be obtained from the isolated material calculation with the exception of the coupling matrix between the two materials.

Then, in order to be able to solve efficiently the kinetic equations for the Green's functions, the size of these blocks is reduced by projecting them into the unit-cell restricted Block function basis which is defined in Sec. 3.2.B.2. We represent the matrix composed of URBF as χ and the block matrix of the Hamiltonian is transformed into the reduced basis by:

$$\begin{aligned} \chi &= [\Phi^{n1}_{k_{x1},\mathbf{k}_{yz}}, .., \Phi^{n1}_{k_{xN},\mathbf{k}_{yz}}, \Phi^{n2}_{k_{x1},\mathbf{k}_{yz}}, .., \Phi^{n_B}_{k_{xN},\mathbf{k}_{yz}}] \\ \mathbf{H}^{\Phi} &= \chi^{\dagger}\mathbf{H}\chi \end{aligned} \tag{5.2}$$

where $\Phi^{n}_{k_x,\mathbf{k}_{yz}}$ is the Bloch function for the n-th band and k_x and $\mathbf{k}_{yz}$ correspond to the wave vectors that belong to the first Brillouin zone.

In the following parts of the chapter, we will describe the two methods to compute this the coupling matrix between the two materials, to be used in the URBF basis for transport calculations.

5.1.B Single material approximation

A simple and computationally efficient way to approximate the coupling matrix between the two regions consists in using the coupling matrix from one of the isolated material. This approach has already been adopted in previous works using maximally localized Wannier functions basis [121]. It consists in approximating the unknown coupling matrix by one of the two isolated materials. Since the URBF basis is different for the two materials, a transformation on the reduced basis has to be done with the following Eq. (5.3), if we keep the notation from the sketch of Fig. 5.2 to identify the blocks.

$$\mathbf{H}^{\Phi}_{N,N+1} = \chi^{\dagger}_{B}\mathbf{H}_{N-1,N}\chi_{A} \tag{5.3}$$

where χ_A and χ_B are the reduced basis Bloch functions matrices of materials A and B, respectively, and $\mathbf{H}_{N-1,N}$ corresponds to the coupling matrix of material A, but we could have chosen $\mathbf{H}_{N+1,N+2}$ which is the coupling matrix of material B. The exchange between materials A and B for the coupling matrix is expected to result in small differences [121], but this point will be investigated later in this chapter.

The main advantage of this strategy resides in its low computational cost as all elements needed have been previously calculated for an isolated material. However, *a priori*, nothing may justify that the connection with another material is the same as the connection with the same material. Thus it may be inaccurate, especially if, additionally, it does not take into account the relaxation of the atomic positions at the interface. This should be

carefully examined.

5.1.C Supercell extraction

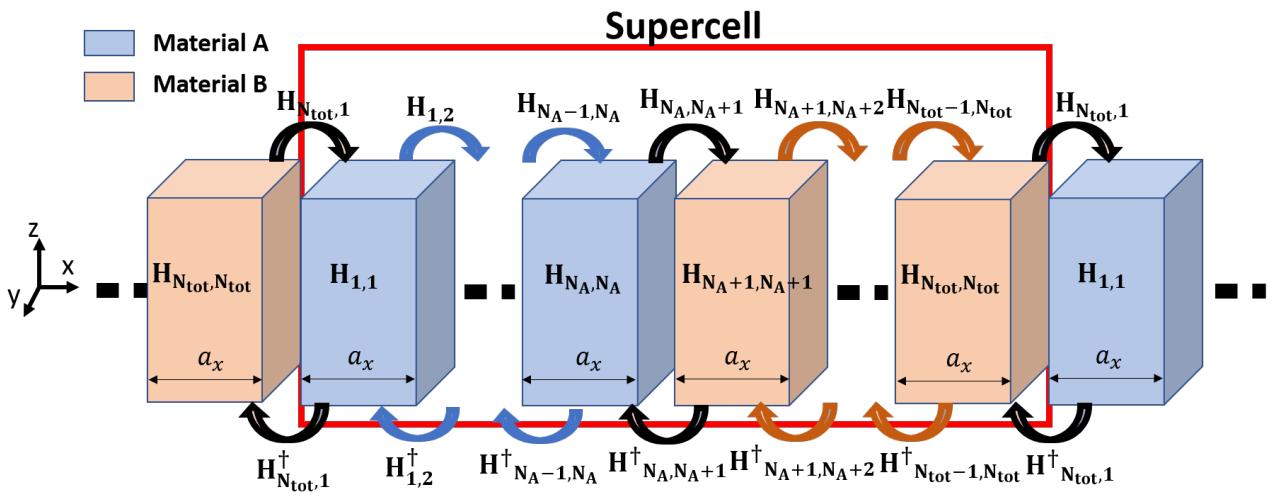

Figure 5.3: Sketch of the supercell heterostructure DFT Hamiltonian simulated on a hybrid basis. Each block corresponds to a unit-cell which are coupled to their first neighbour unit-cell by the coupling matrix

A more accurate way to obtain this coupling matrix from plane waves DFT would be to simulate a huge supercell of the heterostructure including the interface. It can also include a more accurate relaxation of the atom positions at the interface. To suppress all interactions from the supercell replica along the x direction into the interface region, it should be defined large enough. Then, from the Hamiltonian of the system in the hybrid basis, it is easy to identify the coupling matrix at the interface of the two materials by assuming a tridiagonal by block matrix. The block matrix extracted can then be transformed into the reduced basis in a process similar to the previous method in Sec. 5.1.B. The supercell is illustrated with the sketch of Fig. 5.3 where it is possible to observed the periodicity induced by DFT and the presence of the coupling matrix at the interface.

In order to accurately describe the interface in the simulated supercell, all atomic positions are relaxed and if there are dangling bonds at the interface, we passivate them with hydrogen atoms. For a large supercell, only the relaxation of the unit-cell close to the interface is sufficient, which makes the calculation tractable. In fact, simulating a supercell large enough to be sure of the convergence of the charge density at the interface is computationally challenging due to the large memory required to perform the transformation into the hybrid basis and then into the reduced one. We illustrate this procedure with the example of a supercell made of monolayer $PtSe_2$ and bi layer $PtSe_2$ heterostructure as shown in Fig. 5.4. In this figure, we consider a supercell composed of three unit-cell, i.e. two unit cells of monolayer $PtSe_2$ and one of bilayer $PtSe_2$. In the sketch of the supercell in Fig. 5.4 (a), the bilayer presents dangling bonds for the layer that is not connected to the monolayer. These dangling bonds are passivated with hydrogen atoms to avoid spurious states. In this figure, we also represented the block matrix for each unit-cell with the coupling matrix

used in our equations for clarity.

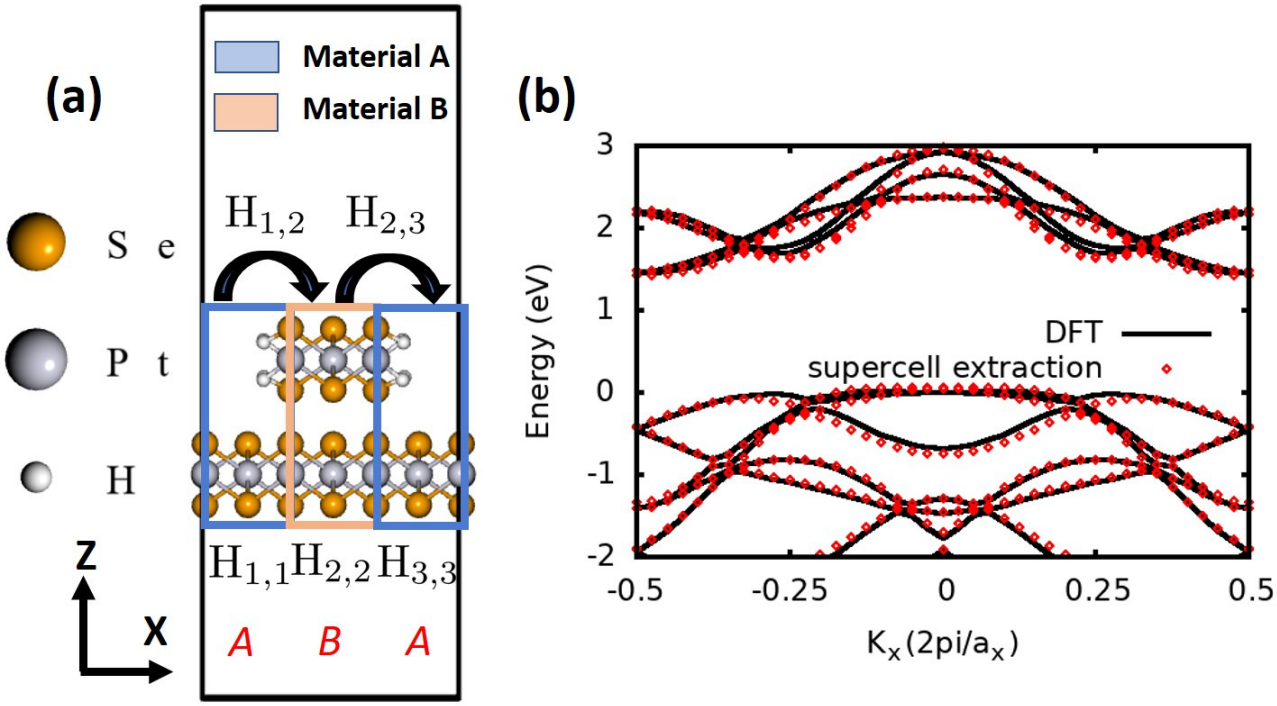

Figure 5.4: (a) Supercell of monolayer $PtSe_2$ and bi layer $PtSe_2$ heterostructure, (b) Band structure of the monolayer $PtSe_2$ computed with the DFT Hamiltonian of the isolated materials (black line) and with the Hamiltonian extracted from the supercell (red diamond).

For the determination of the coupling matrix, another approximation is necessary. We used the URBF basis from the isolated materials because the construction of the URBF basis for the supercell would be difficult to manipulate since the Bloch functions computed by QE correspond to the whole supercell and not only the interface region. It means that it would be much more difficult to find a reduced basis with this Bloch functions for the coupling matrix. The approximation made by using the isolated materials URBF basis amounts to assuming that the Bloch functions for one material used are not different from the wave function present at the interface in this region. The validity of such an approximation has to be verified. This is performed by reconstructing the band structure of one of the materials at the interface by using the Hamiltonian blocks extracted from the supercell Hamiltonian which are transformed into the Bloch function basis of the corresponding isolated material. The condition to perform that is to have at least two connected unit cells of this material into the supercell. In the hybrid basis, the block matrices are easily associated with the corresponding unit cells. It is thus possible to extract and project them on the associated basis. From the Hamiltonian block extracted in the hybrid basis and projected in the reduced basis, we calculated the band structure shown in Fig. 5.4 (b). It has been obtained with the diagonal and coupling block Hamiltonian of

monolayer $PtSe_2$ with Eq. (5.4).

$$\begin{aligned}\mathbf{H}^{\Phi}_{3,3} =& \chi^{\dagger}_{A}\mathbf{H}_{3,3}\chi_{A} \\ \mathbf{H}^{\Phi}_{3,1} =& \chi^{\dagger}_{A}\mathbf{H}_{3,1}\chi_{A}\end{aligned} \tag{5.4}$$

where the notation corresponds to that of Fig. 5.4 (a) with $\mathbf{H}_{3,3}$ being the diagonal block of the supercell Hamiltonian in the hybrid basis. It corresponds to the third unit-cell made of monolayer $PtSe_2$. $\mathbf{H}_{3,1}$ is the off-diagonal block of the supercell Hamiltonian which represents the coupling of the third unit-cell with the first unit-cell also made of monolayer $PtSe_2$.

In this case, we can note some differences between the bands from the isolated material and the ones from the supercell. However, these small discrepancies are not surprising and may be due to the new atomic positions (after relaxation) at the interface and the presence of the hydrogen atoms. The result remains satisfactory as there are no spurious states in the band gap and the maximal error remains smaller then 0.1 eV.

Then, by identifying the coupling block matrix at the interface, the last step consists in projecting it into the reduced basis as

$$\mathbf{H}^{\Phi}_{1,2} = \chi^{\dagger}_{B}\mathbf{H}_{1,2}\chi_{A} \tag{5.5}$$

where $\mathbf{H}_{1,2}$ is the off-diagonal block of the supercell Hamiltonian at the interface in the $(x, \mathbf{K}_{yz})$ basis.

5.2 Simulation of heterostructure-based FETs

To assess the two methods presented, we propose to study two different cases of hetero-junction devices. The first application concerns the case of a MOSFET and the second one is a Tunnel FET. It is important to note that the first situation is a system simpler than the second one regarding the complexity of the coupling between the two materials. Indeed, in the first system, the bilayer $PtSe_2$ is a stacking of two monolayer $PtSe_2$ that is used as the second material. The purpose of those applications is to test both methods for two different transport regimes to emphasize their differences and limitations. Ballistic transport properties are obtained with our home-made self-consistent DFT-NEGF solver coupled with Poisson's equation with the formalism presented in Chapter 3. For each material, the full band description is done by sampling the Brillouin zone along k_y with a step of $0.1 \times 2\pi/a_y$.

5.2.A Multi-layer $PtSe_2$ MOSFET

Monolayer $PtSe_2$ is a 2D material semiconductor with a thickness of 0.5 nm. As for any other 2D materials, doping of $PtSe_2$ is a challenge that still has to be solved [104, 68] with a view to reduce the contact resistance. Substantial interest has been recently raised for $PtSe_2$, because of a neat experimental observation of the semimetal-to-semiconductor transition in ultra-thin $PtSe_2$ [40, 8]. It is possible to reduce its bandgap by adding more layers

until it becomes semi-metallic for a thickness of about 5 layers ($\approx$ 6.5 nm) as reported by Ansari et al. [8]. Figure 5.5 shows their angle-resolved photoemission spectroscopy (ARPES) images that confirm this phenomenon. A relatively good agreement between the DFT calculation and the ARPES measurements can be observed. A clear reduction of the bandgap is observed from the monolayer to the bilayer $PtSe_2$.

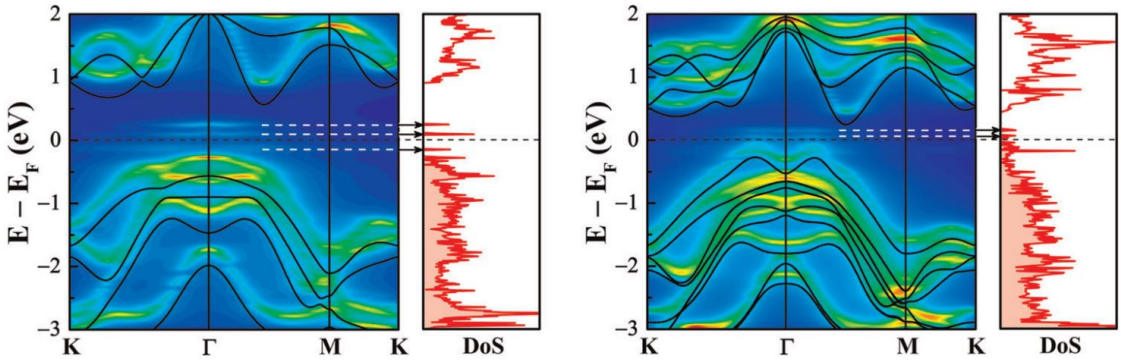

Figure 5.5: Band structure of $PtSe_2$ from DFT calculation in black solid lines on band structure experimentally obtained by ARPES measurements for monolayer (left) and bilayer $PtSe_2$ (right). Figure from [8]

This phenomenon being linked to the interaction between layers, it is important to accurately address the vdW interaction in this material. A comparison of the bandgap evolution versus the thickness of the $PtSe_2$ film is presented for different vdW corrections with spin-orbit coupling and GW calculation in Fig. 5.6.

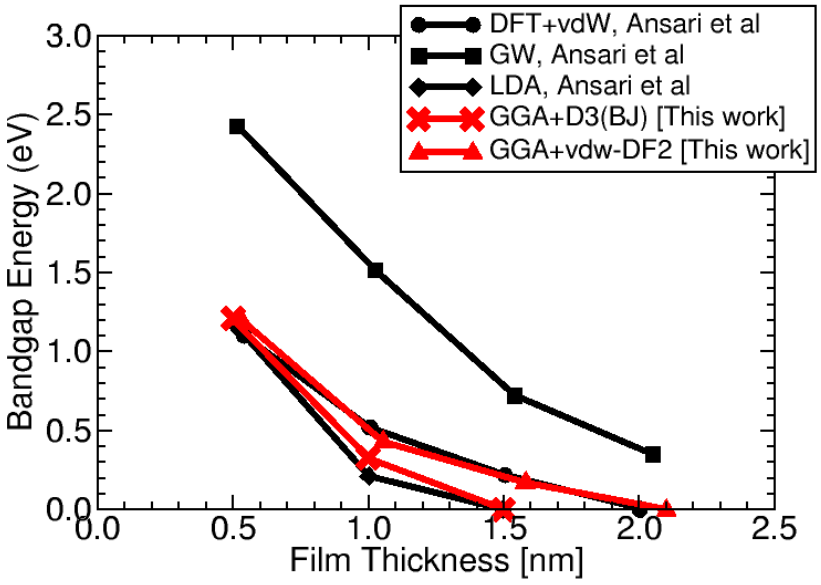

Figure 5.6: Bandgap energy versus PtSe2 film thickness from our DFT calculations: GGA+D3(BJ) (red crosses); GGA+vdw-DF2 (red triangles). A comparison with data from Ref. [8] is also shown: LDA (black diamond); DFT+vdW (black circle); GW (black square).

In these results, we can remark that the choice of the functional has a large influence on the bandgap evolution. As

expected from GW calculation that corrects the bandgap underestimation, a larger bandgap is obtained compared to that resulting from the other DFT calculations. However, the bandgap still converges toward a metallic phase. The semi-empirical correction (GGA+D3(BJ) represented by red crosses in the Fig. 5.6) due to Grimme[70] presents a smoother transition than the non-local vdW correction (GGA+vdw-DF2 and DFT+vdW curves in Fig. 5.6) [74, 31], which results in a semiconductor-to-metal transition for the four-layer film instead of the three-layer one for the non-local correction. Since the non-local correction should result in a more accurate description of the vdW interaction, we have decided to choose this functional for this work. We can also mention that the spin-orbit coupling has almost no effect on the conduction bands of monolayer $PtSe_2$ [98]. Accordingly, since we study an n-type MOSFET, the spin-orbit coupling has been neglected.

The band structure obtained from our calculation is shown in Fig. 5.7 with a bandgap of 1.4 eV for the monolayer and 0.31 eV for the bilayer. Now, we take advantage of this property of $PtSe_2$ to design a MOSFET device with a heterojunction made of bilayer $PtSe_2$, in the drain and source regions for a better electron injection and of a monolayer $PtSe_2$ in the channel to act as a gate-controlled barrier . The sketch of the simulated device is shown in Fig. 5.8. This kind of device has already been proposed by Marin et al. [121] with the $PdSe_2$ that presents the similiar properties as $PtSe_2$.

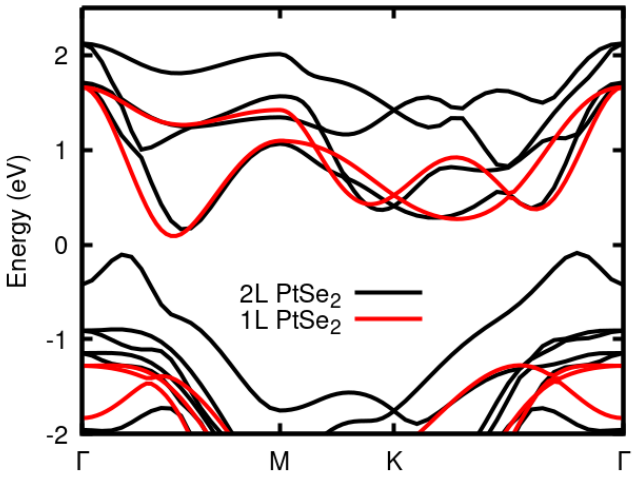

Figure 5.7: Bandstructure of bi-layer and monolayer $PtSe_2$ computed from DFT with non-local vdW correction.

The architecture of the device is presented in Fig. 5.8. A moderately short gate length of $L_G = 15$ nm was chosen. For the electrostatic control of the device, we have chosen an oxide thickness of $t_{ox} = 2$ nm for an EOT = 0.5 nm. For the source and the drain, both regions have been n-doped with a density of $N_D = 5 \times 10^{13}$ cm^{-2}. For the band alignment, the affinity offset has been estimated from DFT to 1.01 eV which results in the band alignment observed in Fig. 5.7.

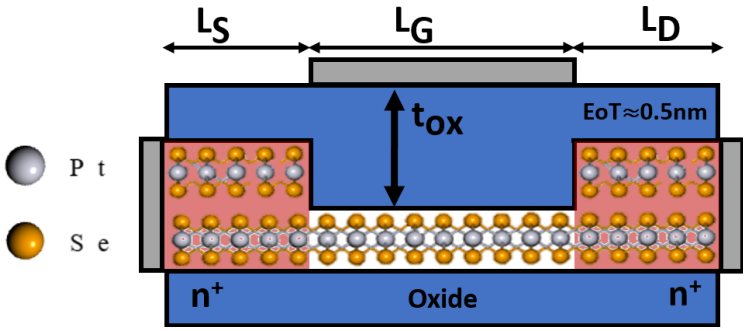

Figure 5.8: Sketch of the bi-layer/monolayer $PtSe_2$ single-gate MOSFET. Geometrical parameters: $L_S = L_D =$ 11 nm, $L_G = 15$ nm and $t_{ox} = 2$ nm.

The transfer characteristics are shown in Fig. 5.9. Similarly to Ref. [95], it is found that the good electrostatic control of the 2D material, due to its thinness, makes it possible to approach the limit of 60 mV/dec for the sub-threshold swing. The ON-current is also excellent with a value of 1.5 mA/μm. This result shows that $PtSe_2$ is a promising candidate as a channel material for MOSFET. However, as we have neglected all the scattering mechanisms, it represents an upper limit of the potential of this device. It means that further studies are needed to validate the promising results achieved here.

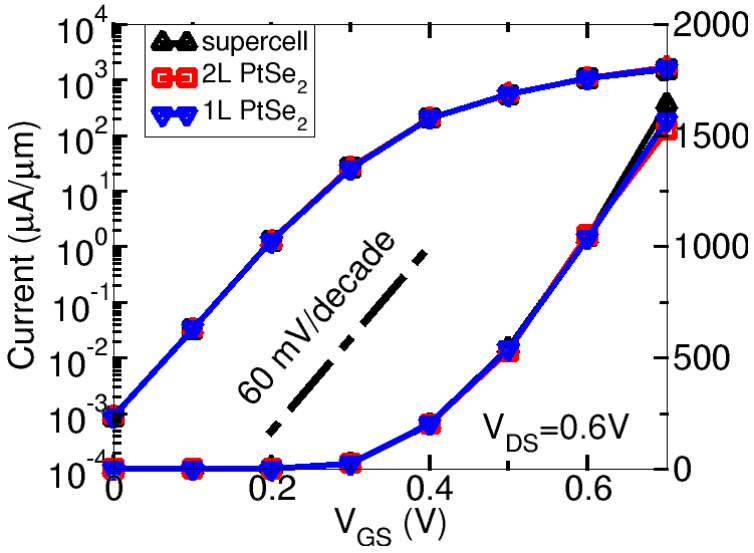

Figure 5.9: Transfer characteristics of the device in Fig. 5.8 computed at $V_{DS} = 0.6$ V by using the coupling matrix of the single material (monolayer $PtSe_2$) and the one extracted from the supercell in Fig. 5.4 (a).

The I-V curves have been obtained by using both methods described in Sec. 5.1 to estimate the influence of the Hamiltonian coupling term describing the coupling between the two materials. The coupling matrix is estimated from the single material approximation by using the coupling matrix of monolayer $PtSe_2$ in one case and of bilayer $PtSe_2$ in a second case.

The third curve corresponds to the calculation with the coupling matrix obtained with the supercell approach. In Fig. 5.9, we see only a small difference at high V_{GS} with a slightly better current for the supercell method

(black curve). Similarly, the band profile along the channel is the same for the three calculations. The transmission probability in Fig. 5.10 confirms that a slightly higher transmission is achieved in the case of the supercell coupling matrix. It can be understood that this method provides a more efficient coupling between the two materials and consequently a larger transmission, since it takes into account the coupling from the monolayer to the monolayer plus the second layer, which is not taken into account with the simplified method.

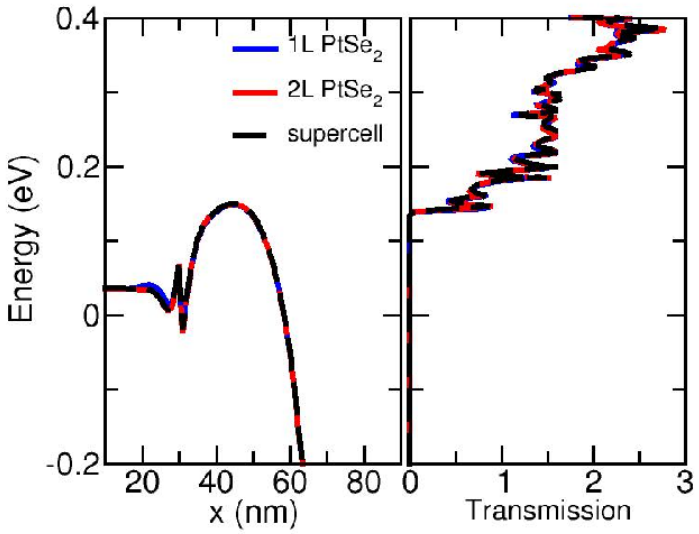

Figure 5.10: Conduction band profile and transmission probability computed at $V_{GS} = 0.5$ V by using the coupling matrix of the single material and the one extracted from the supercell in Fig. 5.4 (a).

In the LDOS presented at Fig. 5.11, almost no difference is observed between the two methods. Only a slightly higher barrier can be detected at the interface between bi- and mono-layer ($x = 7$ nm) in the case of the single material approximation, which results in a reduced transmission.

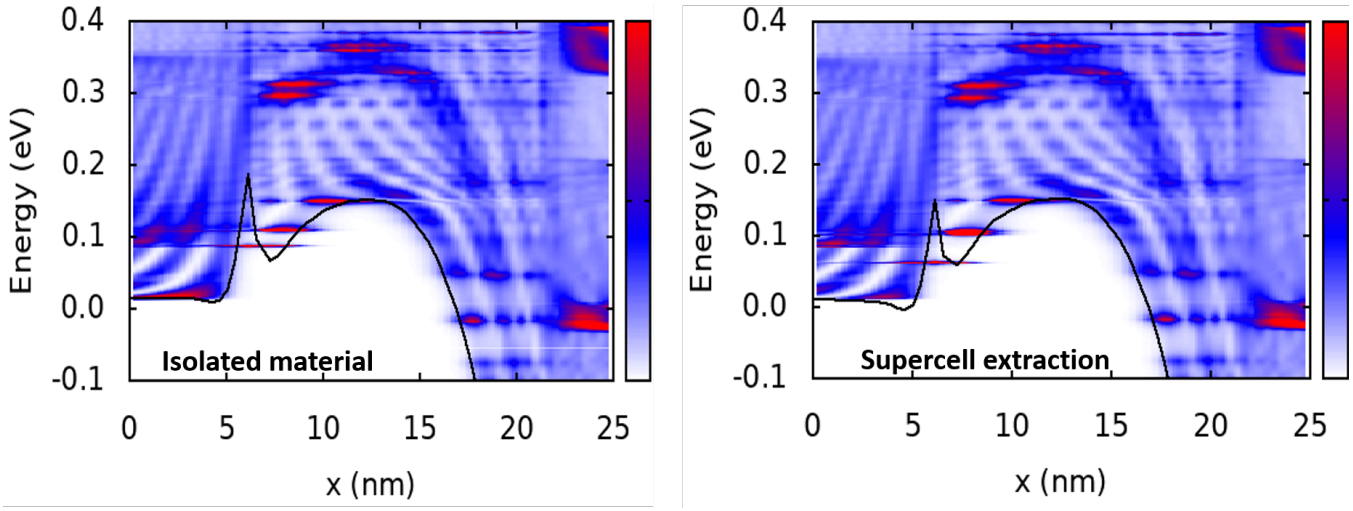

Figure 5.11: LDOS of $PtSe_2$ MOSFET with single material approximation (left) and the supercell method (right) at $V_{GS} = 0.5$ V and $V_{DS} = 0.6$ V

The results for the $PtSe_2$ heterostructure MOSFET indicate that regardless of the method or material used for the off-diagonal coupling matrix, the transfer characteristics remain similar. This could be related to the nature of the heterostructure which is made of two similar materials such as the monolayer $PtSe_2$ and in the bilayer unit-cell, which means that the coupling changes only because of the interaction with the other layer. Consequently, its influence is weak in this case. Moreover, it suggests that the transport through the heterojunctoin is essentially lateral and involves only atoms directly connected between the two materials.

5.2.B Monolayer $PdSe_2/SnS_2$ TFET

Monolayer $PdSe_2$ and monolayer SnS_2 are both TMD semiconductors that satisfy the very selective conditions necessary to make a lateral tunnel-FET. First, they are lattice matched with a lattice parameter close to 0.37 nm by considering a strain smaller than 1% . They also share the same 1T phase which is expected to produce a cleaner interface between the two materials since the atoms arrangement is the same compared to a heterostructure with two different phases. For the TFET part, a specific band alignment is required in order to achieve high ON-current. For 1ML $PdSe_2$ and 1ML SnS_2, the band alignment is favourable for a band-to-band tunnelling device, as the heterostructure corresponds to a type II heterostructure close to type III heterostructure. The band alignment is represented in Fig. 5.13 with an affinity offset of 1.55 eV computed from DFT. For the heterostructure, the highest valance band is located in the 1ML $PdSe_2$ and the lowest conduction band in the 1ML SnS_2.

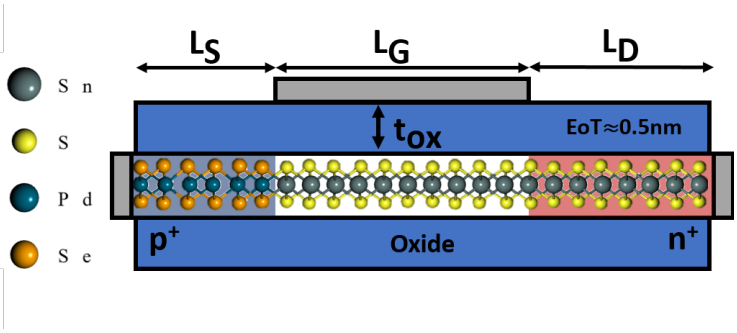

Figure 5.12: Sketch of the monolayer $PdSe_2/SnS_2$ single-gate TFET. Geometrical parameters: $L_S = L_D = 6$ nm, $L_G = 15$ nm, $t_{ox} = 2$ nm. The $PdSe_2$ source is p-doped with a density of $N_A = 1.5 \times 10^{13}$ cm^{-2} and the SnS_2 drain is n-doped with a density of $N_D = 1.5 \times 10^{13}$ cm^{-2}.

The simulated device is shown in Fig. 5.12. As for the MOSFET, a gate length of $L_G = 15$ nm and an EOT of 0.5 nm are chosen. Regarding the doping, the source has been p-doped with a density of $N_A = 1.5 \times 10^{13}$ cm^{-2} and the drain has been n-doped with a density of $N_D = 1.5 \times 10^{13}$ cm^{-2} in order to degenerate both regions, which is appropriate to obtain high current in the ON-state.

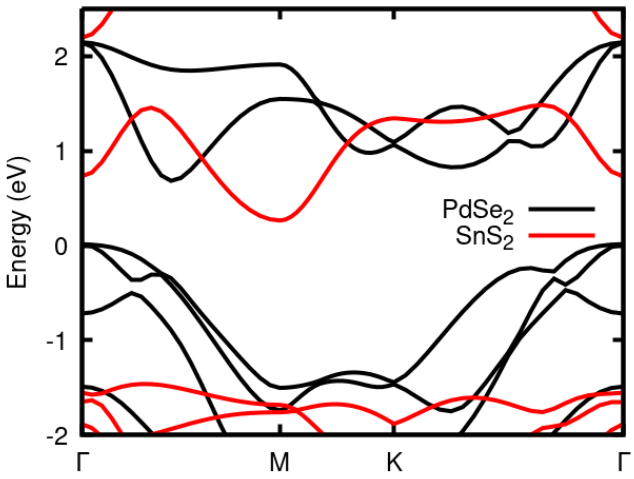

Figure 5.13: Bands of the $PdSe_2$ and SnS_2 obtain with DFT.

The transfer characteristics obtained with the two methods for the coupling matrix at the interface are presented in Fig. 5.14. The typical behaviour of TFET is observed, namely a subthreshold swing of 30 mV/dec, smaller than the MOSFET limit of 60 mV/dec. The ON-current is around 5 μA/μm at 0.3 V. It is relatively low, about two orders of magnitude smaller than the IDRS target of 500 μA/μm. However, this result is somehow expected for this kind of materials since the effective masses along the transport direction of these materials, which correspond to $0.17m_0$ for SnS_2 conduction band and $0.6m_0$ for the valence band of $PdSe_2$, are larger than in most III-V materials, for example. The ON-current is much smaller than the one obtained in Chapter 4, where a ON-current around 300 μA/μm was obtained for a similar III-V TFET. The larger effective mass in a TFET greatly reduces the current due to the tunnelling barrier that the carriers have to go through.

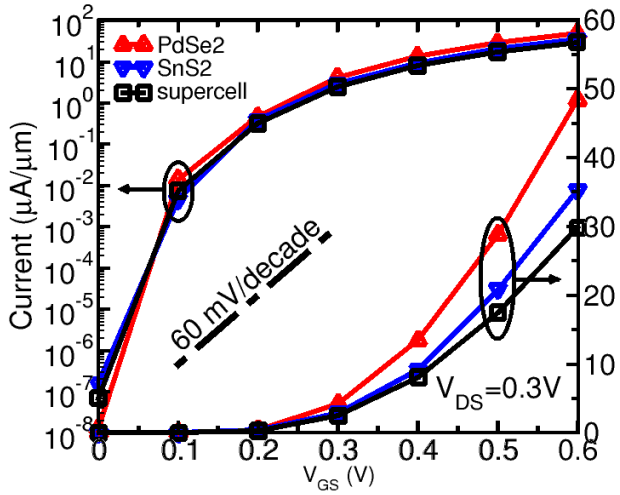

Figure 5.14: Transfer characteristics of the device in Fig. 5.12 computed by using the coupling matrix of the single material $PdSe_2$, SnS_2 and the one extracted from the supercell made of $PdSe_2/SnS_2$

The transfer characteristics have been computed with three different coupling matrices between the two materials. Two were computed with the single material approximation, i.e. one with the coupling matrix corresponding that of the isolated 1ML $PdSe_2$ and the other one to the coupling matrix of the isolated 1ML SnS_2. The last case corresponds to the coupling matrix of the interface extracted from the supercell. Interestingly, the three cases show different I-V curves. Similarly to the MOSFET in the previous section, the differences between the two methods are significant mainly at high V_{GS} but in this case, the opposite tendency between the methods is observed. Indeed, the lowest current comes from the calculation performed with the supercell approach. Moreover, the method with the coupling matrix from the isolated material also provides different results depending on the material chosen for the coupling matrix. This difference may be explained as we are describing the transport through two different materials, whereas previously, the two materials were strongly similar with close band structure.

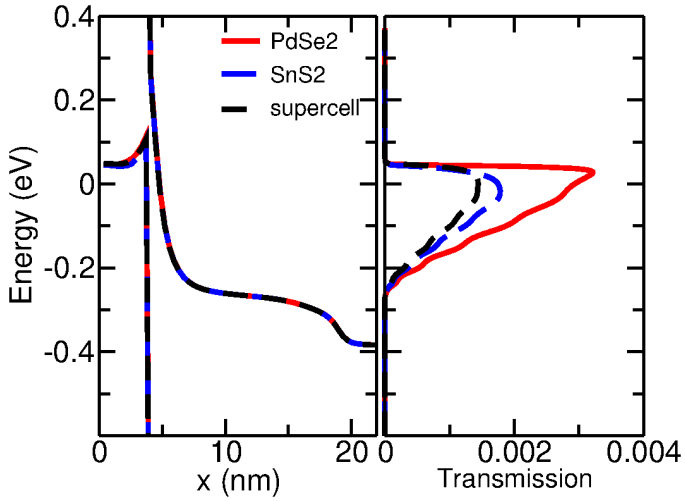

Figure 5.15: Conduction band profile and transmission probability computed at V_{GS}=0.4V by using the coupling matrix of the single material $PdSe_2$, SnS_2 and the one extracted from the supercell made of $PdSe_2/SnS_2$ like in Fig. 5.10

Figure 5.15 shows the bands profiles along the device and the transmission obtained for the TFET and the corresponding transmission probabilities. It can be emphasized that the bands obtained are strictly the same and only the transmission changes. In this device where the transport mechanism is the band-to-band tunnelling, the propagation of evanescent states in the bandgap is significantly dependent on the coupling matrix and the single material approximation provides results that may be highly different depending on the choice of the material for the coupling matrix.

5.3 Summary

In this chapter, we have presented the development of a methodology for the calculation of the coupling Hamiltonian matrix necessary for transport calculations in lattice matched longitudinal heterostructures from first-principles. Two

devices based on lattice-matched heterostructures and paradigms have been simulated to validate the methods. One where we used the coupling matrix of the isolated materials and the second one where we extracted the coupling matrix from the Hamiltonian of a supercell of the heterostructure. The results suggest that for heterostructures based on the assembling of mono- and multi-layers of the same material, both methods show close results, which is a good indication of the validity of the two methods. However, in the case of heterostructures made of different materials, some precautions should be taken for the isolated material method as the results could differ according to the material chosen. If possible, the supercell method should be preferred though it is computationally more demanding.

All of the previous calculations remained in the ballistic regime of transport, which is known to provide an upper limit of the device performances since different scattering events are responsible for many device degradation. In the comparison performed between experiment and our simulation in Chapter 4, we saw that the simulated Esaki diode provides a highly overestimated NDR ratio compared to the experimental measurement. It is predicted to originate from trap assisted tunneling, which illustrate the importance to take into account the possible defects and interactions. Among the interactions that impact the device performance, the electron-phonon interaction is considered as one of the most important as it can be responsible for several scattering events. Hence, the following chapters will focus on the implementation of scattering and more precisely, the electron-phonon interaction.

Chapter 6

Phonon-assisted transport in lateral and van der Waals heterostructure tunnel devices

The previous chapters were dedicated to the development of a simulation framework to investigate the transport properties of materials and the performance of electron devices by means of pseudopotential Hamiltonians (EPM, DFT) with a special focus on heterostructures. For 2D materials, we developed a methodology to model lateral heterostructures based on the DFT. For a lateral tunnel FET (TFET), the ON-current remains generally small compared to the low power device projection made by IRDS [3], for most TMDs [188], due to their large bandgap which increases the tunnelling path and the large effective mass that degrades the transmission through the tunnel barrier. A possible solution to this difficulty is to use van der Waals (VdW) heterostructures to design a vertical TFET. In these heterostructures, a band inversion between the conduction and the valence bands in the channel can occur at high gate voltages, which enables a higher ON-current in the device compared to the lateral TFET since the tunnelling barrier is suppressed and also the tunneling region is larger (line-tunnelling vs point-tunneling) [188, 27, 6]. However, the Van der Waals region, used in TFET, generally presents a smaller bandgap even in the off state which could degrade the OFF-current because of significant electron-phonon interaction [26, 6], thus requiring a careful device design.

In the case of 2D materials, non-atomistic methods like effective mass approximation are expected to be inaccurate due to the poor description of the complex density of states and charge distribution in these materials. It is especially the case for VdW heterostructures, because of the additional difficulty to estimate the interlayer coupling between the two layers [6]. *Ab-initio* methods thus become crucial to model these devices. It is interesting to investigate on the same footing a lateral and a vertical TFET with our NEGF solver based on a plane-wave first-principles Hamiltonians, which provides the most accurate description of electronic properties of these heterostructures.

All the calculations done until now with our transport model have been performed in the ballistic approximation.

This approximation makes the comparison with experimental measurements difficult for some situations as presented in section 4.2.B for the III-V Esaki diode, where the valley current density was underestimated due to the absense of phonons. Moreover, at room temperature, one of the main sources of scattering comes from the electron-phonon interaction, which usually induces a loss of performance in most devices. Today, most contributions employing a DFT Hamiltonian have been so far limited to ballistic transport [121, 137, 174]. The first attempts to include el-ph interaction have used a simplified el-ph coupling based on deformation potentials [6, 164]. Only a few works have considered the derivation and the use of el-ph matrix elements (ME) by using Hamiltonians based on the maximally localized Wannier functions (MLWF) [187].

To improve the predictive capability of our transport model, it is important to include this scattering mechanism. To achieve this task, we need to develop a full *ab-initio* approach to describe this interaction. As mentioned before, from the computational point of view, an efficient method to include el-ph interaction is to use the deformation potential approximation, which was developed in section 3.2.C.2. The goal of this chapter is to develop a first-principles model of the quantum transport in 2D material heterostructures for tunnel FET devices including el-ph scattering with deformation potentials extracted from DFPT calculations. This extended model is then used to address a comparative study of the two TFET architectures, i.e. based on lateral and vertical heterostructures, respectively.

6.1 Choice of the 2D materials

To perform the comparative study between the lateral and the vertical TFET, the choice of 2D materials is crucial to obtain a high-performance device. The criteria to be fulfilled are the same as mentioned in Chapter 5. For the lateral TFET, in order to minimize dislocations and defects, we need two lattice-matched or nearly lattice-matched materials and, if possible, with the same phase. For the vertical TFET, the previous conditions are not so important since the coupling between the two materials is done through van der Waals forces. Thus, no large strain or defect density should arise in this kind of structure. Regarding the electronic properties, the band alignment is important and need to be of type II or III to reach high ON-current. A gap larger than 0.5 eV is required to achieve a sufficiently low OFF-state current. Ideally, for heterostructure n-type TFET, we would like to choose a material with a large bandgap in the channel and drain to filter the current in the OFF state. For the material in the source region, a smaller bandgap is preferred as it would reduce the tunnelling path in the ON-state regime. For this work, we identified the monolayer 1T-SnS_2 and 1T-$HfSe_2$, as good candidates for TFET devices.

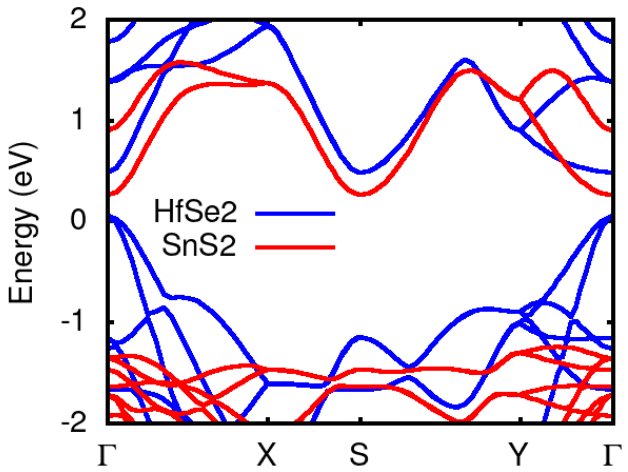

Figure 6.1: Band structure of the orthorhombic unit-cell of monolayer 1T-SnS_2 and 1T-$HfSe_2$ from DFT calculation

In this work, the Hamiltonian and band structure were computed from DFT on a plane wave basis, as discussed chapter 2.3, with norm-conserving pseudopotentials based on the Perdew-Burke Ernzerhof (PBE) [54] exchange-correlation functional. For the monolayer 1T-SnS_2 and 1T-$HfSe_2$, a Γ-centered Monkhorst-Pack k-point grid of dimension $20 \times 20 \times 1$ was used and the plane-wave cutoff energy was set to $E_{cut} = 60$ Ry. From GGA calculation, monolayer 1T-SnS_2 is a semiconductor with a bandgap around 1.64 eV and a unit cell relaxation provides a lattice parameter of 3.7 Å. For monolayer 1T-$HfSe_2$, GGA calculations determined a bandgap of 0.53 eV and an atomic lattice of 3.752 Å. Both band structures are reported in Fig. 6.1 for the orthorhombic unit-cell. Regarding their lattice parameter, they are almost lattice matched (lattice mismatch less than 1%). Moreover, they have the same phase, which makes it possible to consider a stable lateral heterostructure. They have an affinity offset of 0.3 eV determined by DFT giving rise to a type II vdW heterostructure with an energy separation between the lowest conduction band of 1ML SnS_2 and the highest valence band of 1ML $HfSe_2$ of 0.24 eV, which is consistent with literature [184]. The van der Waals interactions were considered in the calculations with non-local van der Waals functional vdW-DF3 [31].

For these materials, the Hamiltonians, in the URBF basis, are constituted of matrix blocks with a rank of about 50 for the monolayer and 100 for the van der Waals heterostructure, which is sufficient to accurately reproduce the band structure as illustrated by Fig. 6.2. In these figures, we report the orthorhombic unit-cell from which we computed the bands of 1ML $HfSe_2$ and SnS_2 obtained at $k_y = 0$ and for $-0.5\frac{2\pi}{a_x} \leq k_x \leq 0.5\frac{2\pi}{a_x}$, where a_x is the lattice parameter along x and has been chosen as 3.727 Å(mean value between the two relaxated materials). The bands were computed with a DFT Hamiltonian in the reduced basis to compare it with the bands obtained with the plane waves basis DFT Hamiltonian. In the reduced basis, for which the rank of the matrix blocks depends on the material, only 42 and 44 Bloch functions are sufficient to reproduce accurately the plane waves DFT Hamiltonian bands of monolayer SnS_2 and $HfSe_2$, respectively.

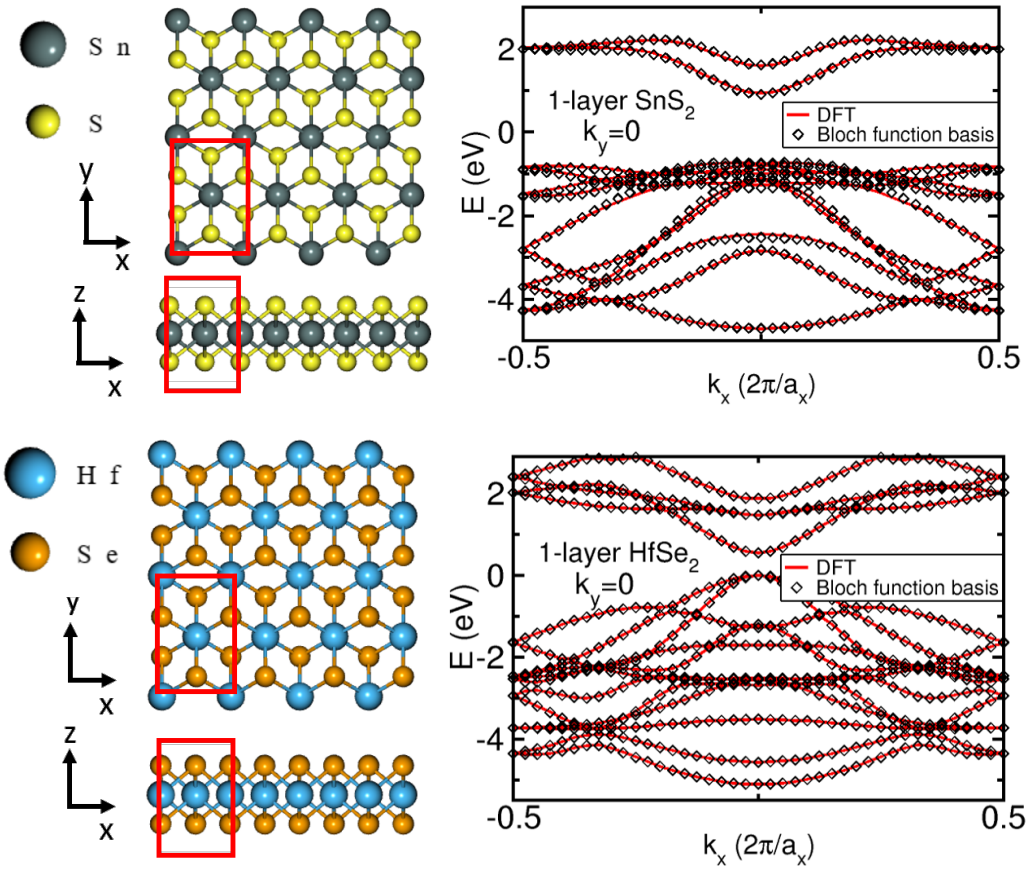

Figure 6.2: Orthorombic unit cell used for transport calculation and band structure on plane waves basis and URBF basis for 1ML SnS_2 (top) and 1ML $HfSe_2$(bottom)

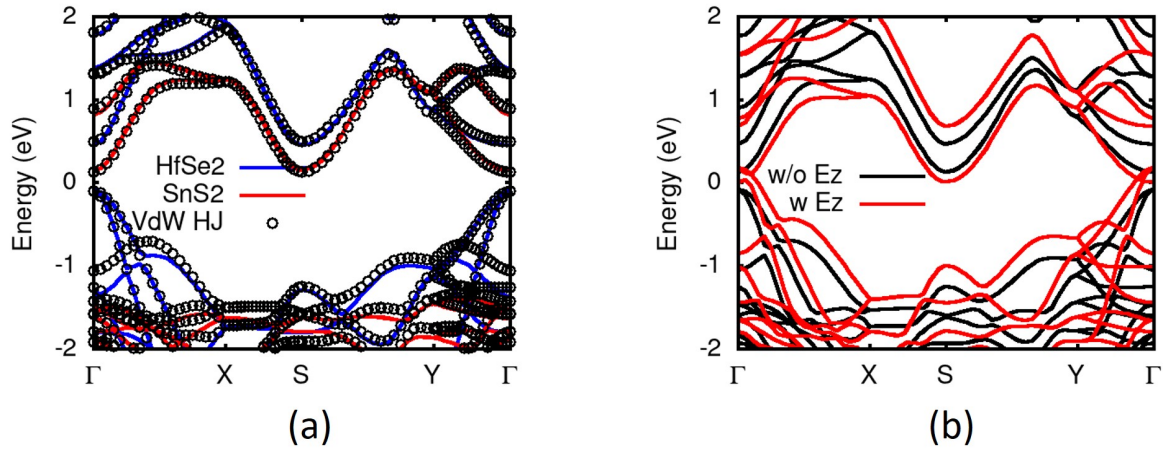

Figure 6.3: Orthorombic unit cell band structure on plane waves without (a) and with an external electric field of 4 V/nm (b)

The band structure of the van der Waals heterostructure obtained by DFT is shown in Fig. 6.3 (a), where the bands of the isolated materials have also been plotted with 1ML 1T-$HfSe_2$ in blue and 1ML 1T-SnS_2 in red. From this figure, we can emphasize that the top of the valence band in the vdW heterostructure accurately fits the bands

from 1ML $HfSe_2$ and similarly for the bottom of the conduction bands with 1ML SnS_2. This is possible thanks to the small degree of hybridization between the two materials in this energy range which arises from the long-range van-der Waals forces and which avoids Fermi level pinning at the interface. Due to the space separation of electrons and holes in the different layers of this Van der Waals heterostructure, a band inversion between the VB of 1ML $HfSe_2$ and the CB of 1ML SnS_2 can be achieved at the Γ point by imposing a vertical electric field as illustrated in Fig. 6.3 (b), showing the band structure at equilibrium and under a vertical electric field for the vdW heterostructure.

To estimate the vertical electric field required, we computed the energy potential inside the unit-cell of the vdW heterostructure with and without an electric field of 4 V/nm fixed in the DFT code, which leads to the band inversion. Fig. 6.4 (a) reports the results for 1T-$HfSe_2$/1T-SnS_2. The electric field is included in the DFT calculation as a constant ramp in the unit-cell surrounded by vacuum on both sides. It can be emphasized that the potential drop is mainly observed outside of the materials which explains the strong electric field required in this simulation to obtain the band inversion. Then, to estimate the electric field inside the vdW heterostructure, we computed the difference of energy with and without the electric field in Fig. 6.4 (b) to remove the ions potential included into the simulation. The remaining potential variation through the vdW heterostructure is that required to obtain the band inversion. This DFT calculation predicts a band inversion between the VB of monolayer $HfSe_2$ and the CB of monolayer SnS_2 under a quite strong electric field of about 0.33 V/nm. This result is achievable in a double gate structure that we propose to investigate in this chapter.

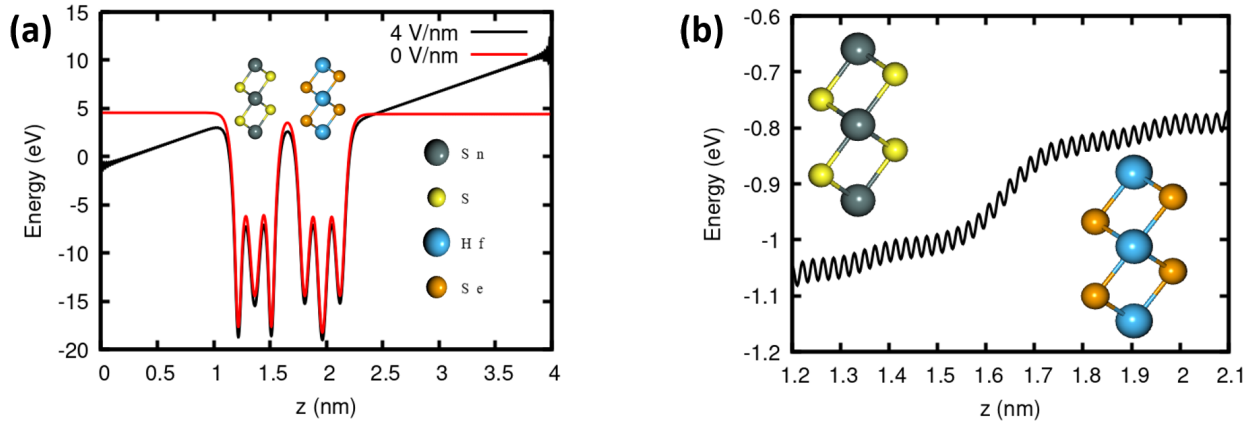

Figure 6.4: (a) Energy of the ions potential + Hartree potential along the out of plane direction of the unit-cell simulated with DFT minus the lowest unoccupied energy level for the vdW heterostructure $HfSe_2/SnS_2$ with and without an external electric field of 4 V/nm and (b) is a zoom, in the vdW heterostructure region, of the difference between previous energies.

6.2 Theoretical model

The transport properties of the lateral and vertical TFETs were computed within the non-equilibrium Green's function formalism with the Green Tea solver [137]. The kinetic equations are solved self-consistently with the 3D

Poisson's equation to obtain an accurate electrostatic description of the system. To take into account the electron scattering with acoustic and optical phonons, we considered DFT and DFPT to derive the deformation potential.

6.2.A Deformation potential from *ab-initio* calculations

To compute the self-energies derived in section 3.2.C.2, it is necessary to model the electron-phonon interaction and phonon properties starting from *ab-initio* calculations. In this Chapter, we have distinguished the acoustic phonons that are considered elastic with a linear dispersion, from the optical phonons, considered to be inelastic with a dispersionless phonon branch. Let us recall the associated self-energies to both types of interaction. For acoustic phonons, the n-th element of the self energy in the URBF basis writes:

$$\Sigma^{<,\mathrm{ac}}_{n,n}(\mathbf{k}_{yz}, E) = \frac{D^2_{\mathrm{ac}} k_B T}{\rho v_s^2} \sum_{\mathbf{q}_{yz}} \sum_{m} G^{<}_{m,m}(\mathbf{k}_{yz} - \mathbf{q}_{yz}, E)\, \mathcal{F}_{(m,\mathbf{k}_{yz}-\mathbf{q}_{yz}),(n,\mathbf{k}_{yz})}. \tag{6.1}$$

where D_{ac} is the deformation potential, $\mathbf{G}^{<}$ is the lesser-than Green's function in the URBF basis, ρ is the mass density of the 2D material, v_s is the sound velocity and $\mathcal{F}$ is a form factor defined in section 3.2.C.2. To compute the self-energy, the unknown parameters are the deformation potential and the sound velocity. For the latter, it is easily estimated from the slope of the acoustic phonon dispersion computed with DFPT. The phonon dispersion for SnS_2 and $HfSe_2$ monolayer is presented in Fig 6.5. The sound velocity is found to be 6120 m.s^{-1} in 1ML 1T-SnS_2 and 4740 m.s^{-1} in 1ML 1T-$HfSe_2$.

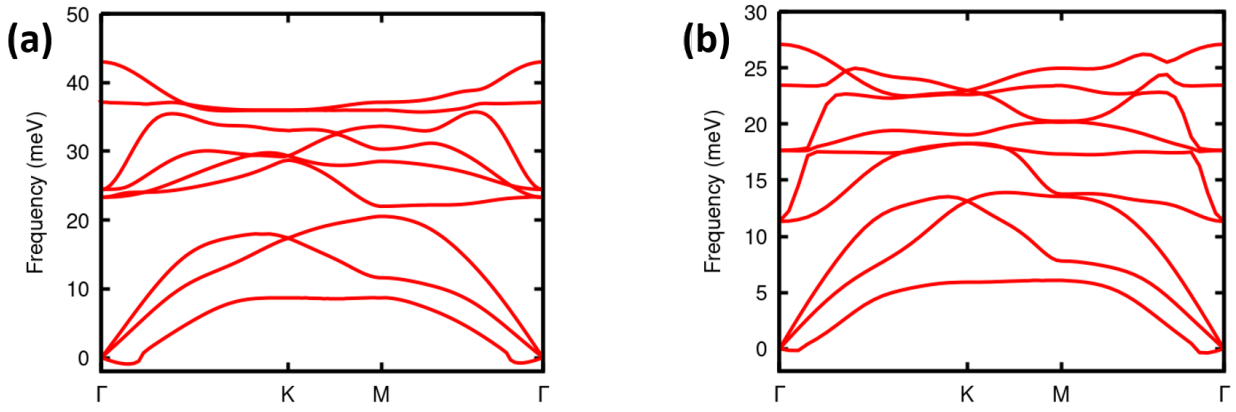

Figure 6.5: DFPT calculation of the phonon dispersion of monolayer SnS_2 (a) and monolayer $HfSe_2$ (b) for the hexagonal unit-cell.

The phonon dispersion of both materials is composed of nine branches since it is computed from a three-atom hexagonal unit-cell. The three lower branches correspond to the acoustic modes. We can remark some negative frequencies, close to Γ, for the lowest branch, which results from a lack of accuracy. Indeed, because of the computational effort required by these calculations, the **q**-grid chosen ($8 \times 8 \times 1$) is too small to provide an accurate description

of the long-range force constants, resulting in errors at the long wavelength limit near the Γ point. The six higher branches correspond to the optical modes presenting a relative dispersion-less behaviour. The deformation potential corresponds by definition to the derivative of the band energy with respect to strain. We used strain values in the range ±2% of the lattice parameter with a step of 1% to evaluate the deformation potential. It can be demonstrated that the deformation potential for acoustic phonons reads [154]:

$$D_{\mathrm{ac}} = \frac{\partial E_{C/V}}{\partial \delta} \tag{6.2}$$

where $E_{C/V}$ is the conduction or valence band energy for electrons and holes respectively, and δ is the biaxial strain applied to the unit cell. From the band structure calculation with different strains, it is possible to estimate the D_{ac} values. The value of the deformation potential for acoustic phonons is found to be 5.47 eV for monolayer 1T-SnS_2 and 2.49 eV for monolayer 1T-$HfSe_2$.

For the polar optical phonons, the self energy derived in Chapter 3 reads:

$$\Sigma^{<,\mathrm{op}}_{n,n}(\mathbf{k}_{yz}, E) = \sum_{\nu} \frac{\hbar D^2_{\mathrm{op,pop}}}{2\rho\omega_{\mathrm{pop}}} \left[N_{\mathrm{pop}} + \frac{1}{2} \pm \frac{1}{2} \right] \sum_{\mathbf{q}_{yz}} \sum_{m} G^{<}_{m,m}(\mathbf{k}_{yz} - \mathbf{q}_{yz}, E \pm \hbar\omega_{\mathrm{pop}})\, \mathcal{F}_{(m,\mathbf{k}_{yz}-\mathbf{q}_{yz}),(n,\mathbf{k}_{yz})} \cdot \qquad \text{with} \quad N_{\mathrm{pop}} = \frac{1}{e^{\frac{\hbar\omega_{\mathrm{pop}}}{k_B T}} - 1} \tag{6.3}$$

where D_{pop} is the deformation potential, N_{pop} is the Bose-Einstein distribution, k_B is the Boltzmann constant, T is the temperature and ω_{pop} is the phonon frequency. The parameters to be determined are the deformation potential and the phonon frequency. Similarly to the sound velocity, the phonon frequency is obtained from the phonon dispersion in Fig. 6.5. For the polar optical phonon mode, the phonon frequency associated is 30 meV for 1ML 1T-SnS_2 and 25 meV for 1ML $HfSe_2$.

Regarding the deformation potentials of optical phonons, a study of the coupling between the electrons and the optical phonon mode shows that the longitudinal optical (LO) mode, which corresponds to the polar optical phonon (POP) mode, is strongly coupled to electrons and dominates with respect to non-polar optical phonons as illustrated by Fig. 6.6. In this figure, the electron-phonon matrix elements are directly extracted from DFPT calculations. The LO phonon coupling exhibits values which are orders of magnitude higher than the other modes. Because of this observation and the fact that in Eq. (6.3) the self-energy is proportional to the square of D_{pop}, in our simulations we will consider only the POP phonon mode.

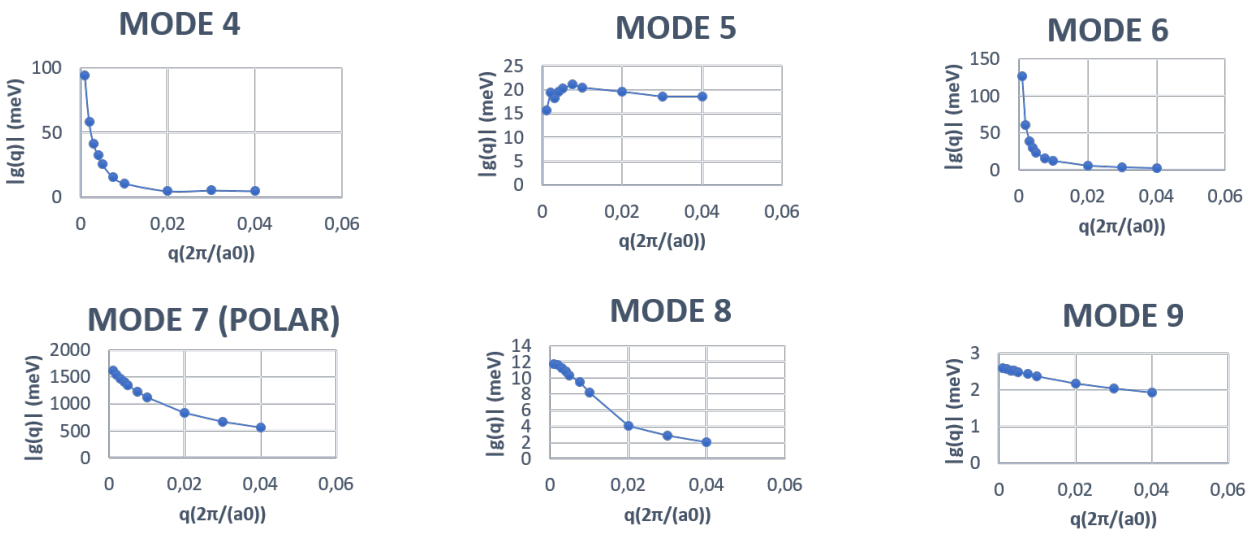

Figure 6.6: DFPT calculation of the el-ph matrix elements of SnS_2 ($\mathbf{q}$ along "Γ→K") at k=Γ for the lowest conduction band.

The deformation potential from the Fröhlich interaction of polar phonon was derived from the value of the scattering rate computed with a 2D model developed by Sohier et al [177]. In this model, the scattering rate for the polar optical phonon mode is estimated from the Fermi golden rule and writes:

$$\frac{1}{\tau_{\mathrm{pop},\mathbf{k}+\mathbf{q}}} = \frac{2\pi}{\hbar}\sum_{\mathbf{q}} |g(\mathbf{q})|^2 \delta(E(\mathbf{k}+\mathbf{q}) - E(\mathbf{k}) \pm \hbar\omega_{\mathrm{pop}}) \left[N_{\mathrm{pop}} + \frac{1}{2} \pm \frac{1}{2}\right] \tag{6.4}$$

where g is the electron-phonon matrix element. The + sign corresponds to the phonon emission process and the − to the phonon absorption process. The 2D model provides an expression for the polar electron-phonon matrix elements in the long-wavelength limit, where the dielectric environment is taken into account. It reads as:

$$g(\mathbf{q}) = \frac{C_Z}{\epsilon^0_{eff} + r_{eff}|\mathbf{q}|} \tag{6.5}$$

where we define C_Z, which is the bare magnitude of the polar-optical coupling, as:

$$C_Z = \frac{2\pi e^2}{A}\sum_a \frac{\mathbf{e_q} Z_a \mathbf{e}^a_{\mathbf{q},LO}}{\sqrt{2M_a\omega_{\mathbf{q},LO}}} \tag{6.6}$$

where A is the area of the unit-cell, a runs over all the atoms of the unit-cell, $\mathbf{e_q}$ is the unit vector along $\mathbf{q}$, $\mathbf{e}^a_{\mathbf{q},LO}$ is the phonon eigenvectors for the LO mode normalized over the unit cell, M_a is the atomic mass of atom a, Z_a is the Born effective charge tensor. Moreover, ϵ^0_{eff} and r_{eff} are functions of the dielectric permittivity of the monolayer and of the surrounding dielectric environment. Their definitions are provided in Ref.[177]. To compute the electron-phonon

matrix elements, we need the Born effective charge tensor and the dielectric permittivity of the materials. Both are provided by DFPT calculations. Hence, the deformation potential was obtained by considering that the imaginary part of the self-energy is the carriers lifetime [66], then the scattering rate is expressed as:

$$\frac{D_{\text{op}}^2}{\rho\omega_{\text{pop}}}\left(\frac{\sqrt{m_x m_y}}{2\pi\hbar^2}\right)N_{\text{pop}} = \frac{1}{\tau_{\text{pop}}} \tag{6.7}$$

Where m_x and m_y are the effective mass along x and y direction, respectively, used to estimate the density of states of the corresponding valley. The left term corresponds to the absorption scattering rate computed from the el-ph self-energy at the lowest energy value of the conduction band or the highest value for the valence band. Since we estimate the deformation potential for both materials, the right term is evaluated from Eq. (6.4) taken at **k** corresponding to the conduction band minimum for SnS_2 and the valence band maximum for 1ML $HfSe_2$.

Phonon and electron-phonon interaction parameters obtained by DFPT for monolayer $HfSe_2$ and SnS_2 are listed in the table 6.1.

Table 6.1: Phonon and electron-phonon properties of 1T-SnS_2 and 1T-$HfSe_2$.

	$v_s(\text{m.s}^{-1})$	$D_{\text{ac}}(\text{eV})$	$\hbar\omega_{\text{pop}}(\text{meV})$	$\frac{1}{\tau_{\text{pop}}}(\text{s}^{-1})$	$D_{\text{pop}}(\text{eV.cm}^{-1})$
1T-SnS_2	6120	5.47	30	6.7×10^{11}	3.7×10^{8}
1T-$HfSe_2$	4740	2.49	25	8.9×10^{10}	2.7×10^{8}

6.3 Results and discussions

Now that we have presented the theoretical model, we can discuss the device architectures and the obtained results.

6.3.A Device design

To achieve high ON-current with tunnel FET, it is needed to have a source region that is in a broken-gap configuration with the channel and drain regions. In this regard, for the n-type TFET, monolayer 1T-$HfSe_2$ was chosen for the source as its valence bands are close to the lowest conduction band of the monolayer 1T-SnS_2 that was used for the channel and drain region. In the case of the VdW tunnel FET, the VdW heterojunction extended with a monolayer 1T-SnS_2 was chosen for the channel to obtain a better OFF-state as the modulation with the gate voltage of the large bandgap of SnS_2 will help to suppress the current in the OFF state [26] and to provide a better subthreshold swing.

To enhance the electric control of the devices, a high-κ oxide with a dielectric constant of 25 (similar to that of HfO_2 [148, 107]) and a thickness of 3.2 nm have been chosen to obtain an EOT of 0.5 nm.

In all our simulations, the Fermi level of the source was taken as the energy reference. The device is meant to be a low-power device thus we have chosen a power supply of $V_{DD} = 0.35$ V and the OFF-state was fixed at $I_{OFF} = 0.1$ nA/μm.

6.3.B Lateral tunnel FET

The sketch of the lateral tunnel FET is illustrated in Fig. 6.7 and we remind that the bandstructrures of the two materials were displayed in Fig. 6.1 showing that the heterostructure is of type II. For the lateral tunnel FET, a single gate is used to control the channel. The length of the source and drain regions is 8 nm for each which is long enough to ensure that the electrostatic field from the channel is screened. We choose a channel length of around 15 nm for an efficient device performance in OFF-state as we are not studying the scaling of TFETs. The drain region is n-doped with a density of $N_D = 10^{13}$ cm^{-2} and the source region is p-doped with a density of $N_A = 10^{13}$ cm^{-2} in order to degenerate both regions to achieve a high ON-current by increasing the energy window inside which the current will flow. The channel region is n-doped with a density of $N_D = 0.5 \times 10^{13}$ cm^{-2}. This doping level is necessary to reduce the depletion region in the source. It produces a sharp band profile at the interface between the source and the channel which results in a reduced tunnelling path for a high ON-current. Controlled doping of these regions may be done by chemical doping as realized in some recent works [69, 52].

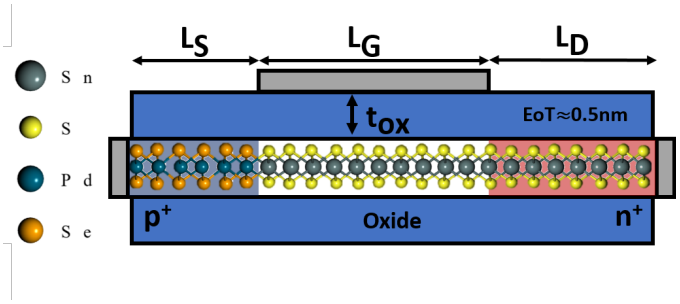

Figure 6.7: Sketch of the lateral heterostructure of $HfSe_2/SnS_2$ TFET device where the blue and red shaded region represent p- and n-doped region, respectively.

In order to have a complete picture of the impact of el-ph interaction in TFETs, we have computed the current for different scattering rates. Figure 6.8 shows the transfer characteristics obtained by using either the ballistic approximation or the el-ph interaction with different values of the POP scattering rate. The values of τ_{pop} for the two materials are reported in Table 6.1. As one might expect, the lateral TFET exhibits a subthreshold slope much steeper than the MOSFET limit with a value around 15 mV/dec for the ballistic curve, which is also a signature of the excellent electrostatic control of the device. It can be emphasized that the phonons have a small impact on the device performances with a rather increase of the OFF-current for POP scattering rates up to $\frac{1}{\tau_{\text{pop}}}$. For higher scattering rate equals to $\frac{10}{\tau_{\text{pop}}}$, the current is increased in the full range of bias with a degradation of the SS that increases by 19 mV/dec to reach 25 mV/dec.

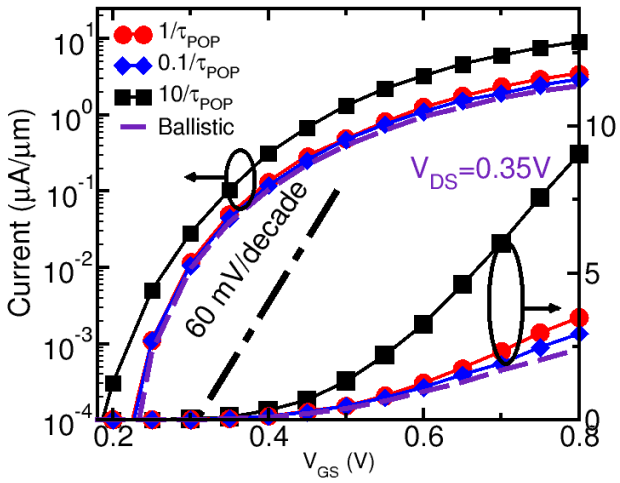

Figure 6.8: Transfer characteristics for the lateral TFET in the ballistic approximation and for different scattering rates at $V_{DS} = 0.35$ V.

Figure 6.9 represents the current spectrum along the device in the transport direction. It illustrates the role played by the phonons with a scattering rate of $6.7{\times}10^{12}$ s^{-1} in the OFF-state. We choose to plot it at $V_{\mathrm{GS}} = 0.45$ V and for the strongest scattering rate ($\frac{10}{\tau_{\mathrm{pop}}}$) that we simulated for the sake of clarity. The phonons are responsible for an increase of current, as they promote valence-band electrons to higher energies by allowing them to cross the tunneling barrier. This phenomenon induces an increase of the subthreshold swing from 15 mV/dec in the ballistic regime to 24 mV/dec with a strong scattering rate ($\frac{10}{\tau_{\mathrm{pop}}}$).

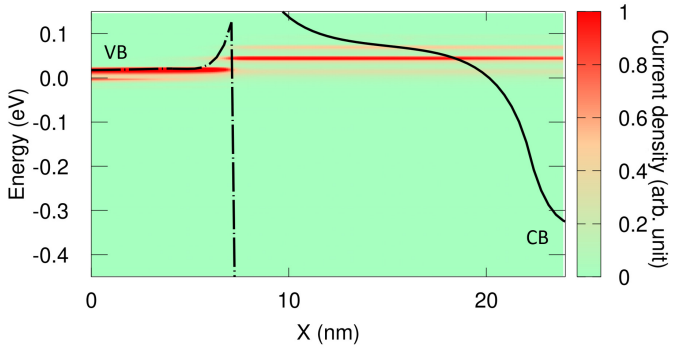

Figure 6.9: Contour color plot of the current spectrum. CB (solid line) and VB (dashed line) profiles along the transport direction for $V_{GS} = 0.45$ V with a scattering rate at $\frac{10}{\tau_{\mathrm{pop}}}$.

When looking at the ON-state, quite unusually, it is found that when the el-ph coupling is strong enough, the electron-phonon interaction also increases the ON-current significantly compared to the ballistic case. This phenomenon can be explained with the contribution of inter-valley transmission mediated by phonons in the valence band of monolayer $HfSe_2$ at the Γ point to the conduction band valley of monolayer SnS_2 in S point (see Fig. 6.3).

This is illustrated in Fig. 6.10 and Fig. 6.11, showing the current spectrum in the ON-state and the LDOS for two different lateral wave vectors k_y. In the $k_y = 0$ case, the charge flow is localized in the source region composed of $HfSe_2$, while in the $k_y = 0.5 \times \frac{2\pi}{a_y}$ case, in the channel region made of the SnS_2. This is due to the fact that monolayer 1T-SnS_2 has a second conduction valley at the edge of the first Brillouin zone (high-symmetry point S) close to the first valley (high-symmetry point Γ), which is not the case for the valence bands of monolayer 1T-$HfSe_2$. This can be seen in the band structure presented in Fig. 6.1. Similarly, the LDOS along the devices can illustrate it. It is represented in Fig. 6.10, where at $k_y = 0$, their is a no-null local density of states in the conduction bands of the channel which is in front of the LDOS of the valence bands in the source. This is the origin of the band-to-band current for this transverse wave vector. However, at $k_y = 0.5 \times \frac{2\pi}{a_y}$, it is not the case and no current results from this transverse wave vector. In Fig. 6.11, we can observe the spectral current density in the presence of phonons, the inter-valley transmission mediated by phonons makes it possible for carriers to populate the LDOS in the channel at $k_y = 0.5 \times \frac{2\pi}{a_y}$, which results in a current increase. We remark that this phenomenon cannot be observed with a mere ballistic simulation since the valence band of $HfSe_2$ for $k_y = 0.5 \times \frac{2\pi}{a_y}$ is far below the Fermi level in the source.

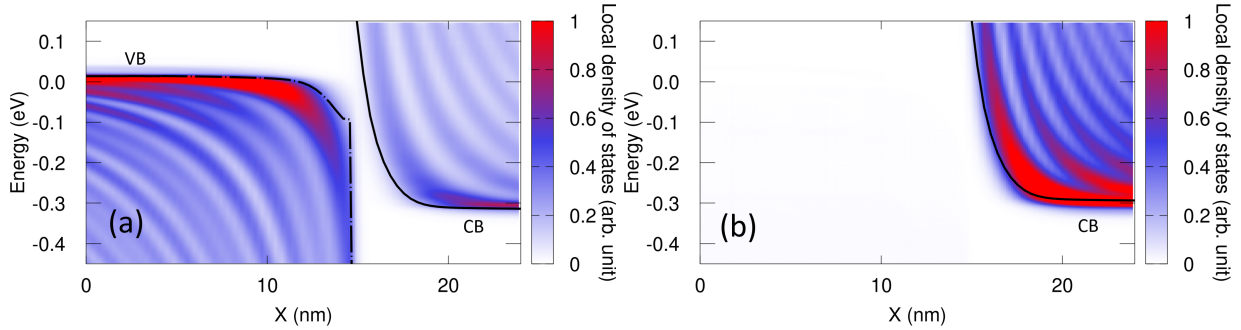

Figure 6.10: LDOS of the lateral $SnS_2/HfSe_2$ TFET at (a) k_y=0×$\frac{2\pi}{a_y}$ and (b) k_y=0.5×$\frac{2\pi}{a_y}$.

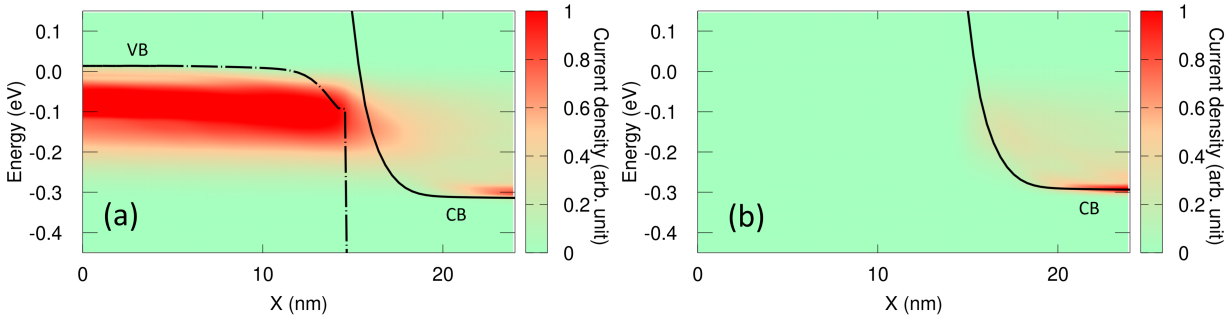

Figure 6.11: Contour plot of the current spectrum of the lateral $SnS_2/HfSe_2$ TFET at (a) k_y=0×$\frac{2\pi}{a_y}$ and (b) k_y=0.5×$\frac{2\pi}{a_y}$.

Even though the tunnelling path is short because of the small gap of 1ML $HfSe_2$, a relatively low ON-current (3 $\mu A/\mu m$) is observed for the lateral TFET. The reason behind this is due to the quite large effective mass of both materials $0.16m_0$ for the isotropic 1ML 1T-$HfSe_2$ highest valence band valley at Γ and $0.17m_0$ for 1T-SnS_2 lowest

conduction band at Γ along the transport direction and $0.78m_0$ for the transverse directions. These values along the transport direction are not small enough to obtain a high current due to the tunnelling path. The low density of states in the valence bands of 1ML 1T-$HfSe_2$ also makes it difficult to reach high current in this condition.

6.3.C Van der Waals tunnel FET

The architechture of the Van der Waals tunnel FET is presented in Fig. 6.12. For the same reason as the lateral TFET and to study them on the same footing, the length of the source and drain are 8 nm for each and we choose a channel length around 15 nm for the channel region. The overlap part of the channel is 10 nm long. The role of the gate extended region is to reduce the OFF-state current. It provides a better OFF state by using this large bandgap material as a filter [26]. For the VdW tunnel FET, two gates are needed with one grounded in order to generate a large enough vertical electric field in the channel region. The electric field is used to achieve the band inversion in the ON-state. To achieve the band inversion when we switch ON the device, a work function difference between the two gates of 0.7 eV is chosen. This work function difference corresponds to the potential difference necessary between the two gates in order to generate an electric field strong enough to achieve the band inversion in the channel region in the ON-state.

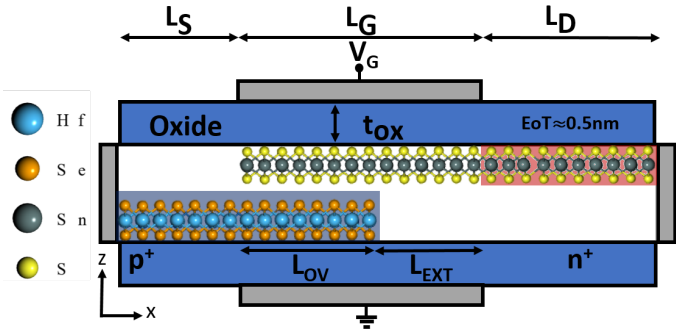

Figure 6.12: Sketch of the vdW heterostructure of monolayer $HfSe_2/SnS_2$ TFET device where the blue and red shaded region represent p- and n-doped region respectively.

Since we extended the gate over the SnS_2 monolayer in the channel, we aim to tune the conduction bands of vdW heterostructure to reach the band inversion between the valence and conduction bands according to what we have seen before from DFT calculations in Sec. 6.1. Thus, the gate is chosen to be the one close to the 1ML SnS_2 layer as it provides a better control on the conduction band of the vdW heterostructure. If we would have chosen the gate close to 1ML $HfSe_2$, part of the field would have been screened by the 1ML $HfSe_2$ resulting in a less efficient electrostatic control. This is more efficient for attaining the band inversion as the conduction band in the 1ML SnS_2 will be more modulated than the valence band in the 1ML $HfSe_2$. 1ML $HfSe_2$ is p-doped with the same doping density as the source $N_A = 10^{13}$ cm^{-2}. The drain is n-doped with a density of $N_D = 10^{13}$ cm^{-2}. This particular doping, in the vdW heterostructures, set the top of the valence band of the $HfSe_2$ in the channel at the same energy

level as in the source in the OFF-state.

We found that the electron-phonon coupling is responsible for numerous phenomena impacting the current control of the device, as shown in Fig. 6.13, where we plot the transfer characteristics with different scattering rates of the polar optical phonon scattering $\frac{1}{\tau_{\text{pop}}} = 6.7 \times 10^{11}$ s^{-1} to emphasize its impact. A first remark about the I-V curves is that the ballistic approximation always results in a lower current compared to the calculation with the phonons, even in the ON-state. The reason for that is the same as we discussed for the lateral TFET. Then, it can be observed that the current is higher than in the lateral device and can reach high values for low power devices with an ON-state current of 298 μA/μm for a $V_{DD} = 0.35$ V. In the ballistic approximation, the subtheshold swing is similar to that of the lateral TFET with an SS of 15 mV/dec. It is similar to the lateral TFET.

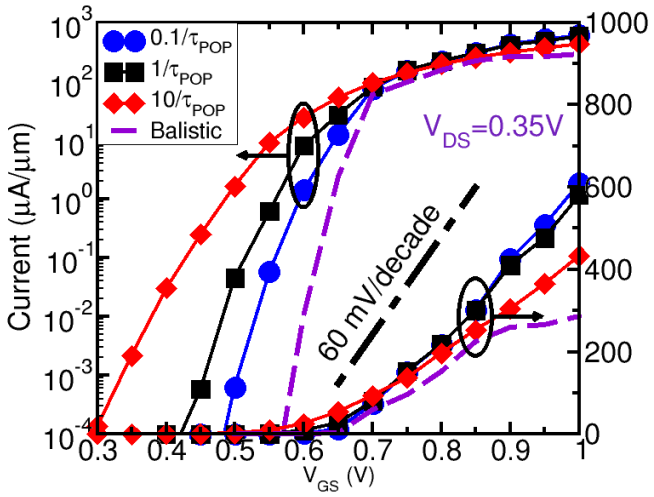

Figure 6.13: Transfer characteristics for different scattering rates and V$_{DS}$=0.35V.

However, the vertical device is much sensible to the electron-phonon coupling as the el-ph coupling strength greatly increases the subtheshold swing from 15 mV/dec in the ballistic case to 40 mV/dec when strong scattering rate ($\frac{10}{\tau_{\text{pop}}}$) is used.

Figure 6.14 shows the current spectrum in the OFF state. It illustrates the importance of phonon absorption mechanism in the vdW region that. Moreover, the small bandgap of the vdW heterostructure is small enough for carriers to go from the valence band to the conduction band of the vdW material with only a few phonon absorptions. It enables a larger current in the OFF-state that is responsible of a significant degradation of the sub-threshold slope.

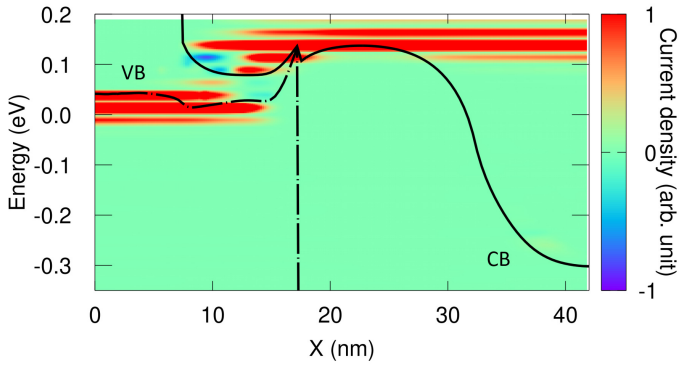

Figure 6.14: Current density of the vertical TFET at $V_{GS} = 0.5$ V

For the ON-state, Fig. 6.15 shows the corresponding current spectrum when band inversion between the valence and conduction bands is achieved in the vdW region for two scattering rates. This band inversion is one of the reason of the higher ON-current achieved compared to the lateral TFET. As it suppresses the tunnel barrier at the interface, a significant current for a TFET of 298 μA/μm is obtained for the ballistic approximation. With the phonon interactions, a further resistive effect due to the phonon emission in the drain is visible and explains the decrease of the ON-state current as a function of the polar optical phonon coupling visible in Fig. 6.13. Indeed, as the resistive effect increases, the energy window for the carriers to pass through the device tends to close, reducing the current.

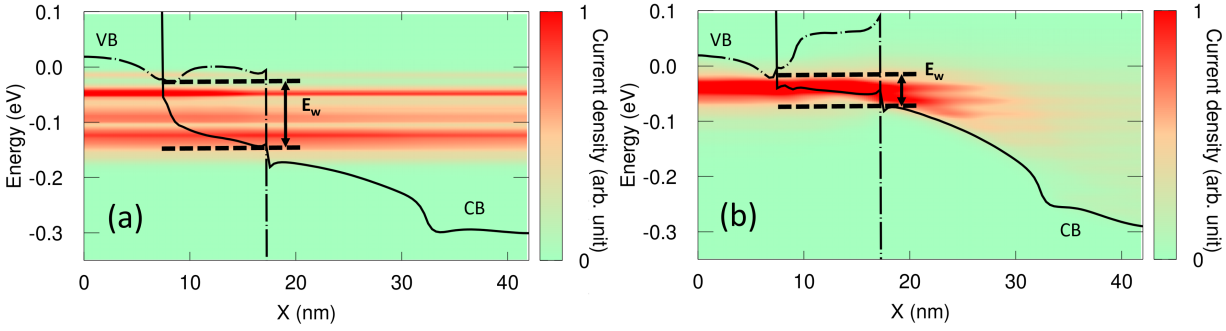

Figure 6.15: Current density at $V_{GS} = 1\,V$ for the VdW TFET with scattering rate at $\frac{1}{\tau_{\text{pop}}}$ (a) and $\frac{10}{\tau_{\text{pop}}}$ (b)

We also have studied the impact of the overlap length. Figure 6.16 presents the I-V curves for different overlap lengths. The ON-current increases weakly with this design parameter. It suggests that the carrier transport is a point-tunnelling transmission at the edge of the heterostructure like the lateral device. If it was not a point-tunnelling transmission, a larger current increase would be expected as the region for the carrier to tunnel from one layer to the other is extended. The higher current observed in the vdW TFET compared to the lateral heterostructure is certainly due to the inversion of the conduction and valence band in the heterojunction and the reduction of the

tunnelling path that comes with it.

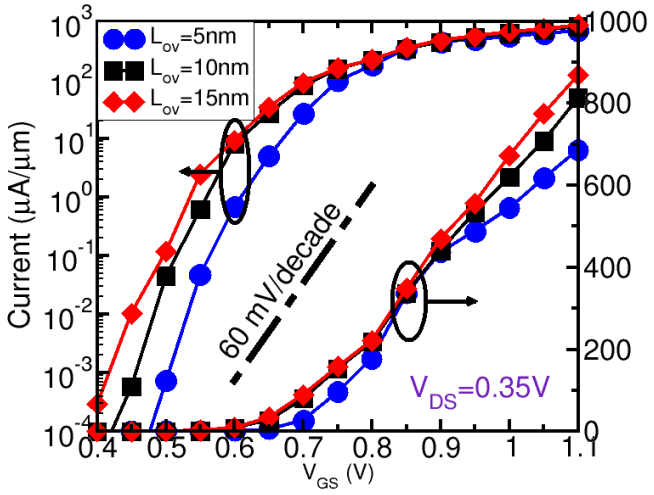

Figure 6.16: Transfer characteristics for different L_{OV} and at $V_{DS} = 0.35$ V.

Last but not least, Fig. 6.17 summarizes the performance evolution of the two architectures. It represents the SS and I_{ON} of the van der Waals and the lateral TFET. It can be noted that the lateral TFET is much more robust to phonons, but its low electrical performance for the I_{ON} makes it less attractive than its Van der Waals counterpart. The latter is still qualified to satisfy the recommendations of IRDS for the next technological nodes that predict a low power device with a ON-state current higher than 500 μA/μm and with a SS smaller than 65 mV/dec.

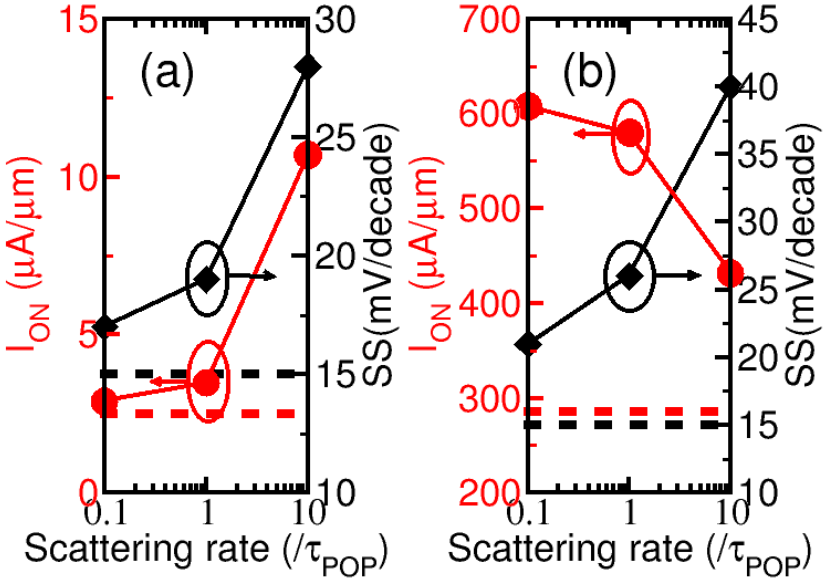

Figure 6.17: Average SS over the window 10^{-4} μA/μm $<$ I$_{DS}$ $<$ 1 μA/μm and I$_{ON}$ as a function of the scattering rate for the lateral TFET (a) and van der Waals TFET (b). Dotted horizontal line correspond to the ballistic case.

6.4 Summary

This chapter presented a transport study for two device architectures of 2D materials heterostructures. We extended the ballistic study of tunnel FETs, in chapter 5, with the implementation of electron-phonon interaction whose parameters were completely determined by a first principles approach. The results suggest that vdW TFETs are significantly affected by phonon scattering due to the large coupling of electrons with polar phonons where the lateral architecture remains more robust to the el-ph interaction. However, the vdW architecture still represents a promising option for the next generation electronics due to the high current around 580 μA/μm and a SS of 26 mV/dec compared to the lateral TFET that is not able to reach high current with these materials.

From these observations, two strategies can be chosen to design efficient TFETs based on 2D materials. The first one would be to stick to the van der Waals architecture and to optimize the dielectric environment to screen the polar phonon modes and thus reducing their impact. The second one would be to work on lateral heterostructures in order to increase the ON-current by finding a better couple of 2D materials with a smaller effective masses along the transport direction.

These results illustrate the importance to add electron-phonon scattering to realistically simulate nanoscale electron-devices. However, the approach that we proposed here mainly describes the el-ph coupling in the long-wavelength limit and the phonon wave vector dependence of the el-ph coupling is completely neglected, whereas it can play a key role for the important inter-valley scattering. Thus, it is important to develop a more accurate model which fully takes into account the complex phonon branch dispersions as well as the phonon wave vector dependency of the el-ph coupling. This is the topic of the next chapter.

Chapter 7

Ab-initio electron-phonon calculations using DFT, DFPT and NEGF methods: applications to 2-D materials mobility and nanoscale FETs

7.1 Motivation

The progress in *ab-initio* methods makes it nowadays possible to calculate not only the electronic structure but also the electron-phonon (el-ph) coupling by using respectively DFT and DFPT [16]. The non-equilibrium Green's function approach with a DFT Hamiltonian and comprising a DFPT based el-ph coupling is arguably the most fundamental and predictive approach for transport in electron devices.

At the material level, *ab-initio* calculations of the phonon-limited mobility have been recently reported for several TMDs [71, 179]. For now, the mobility calculations obtained with *ab-initio* methods have resulted in a quite large spreading of values that are still under scrutiny [62]. Most of the previous works tend to show that an accurate description of the el-ph matrix elements is essential for mobility calculation [144].

In Chapter 6, we made a first attempt to describe the electron-phonon interaction using the deformation potential approximation, showing that this has a non-negligible impact on the device performance. However, the model presented in Chapter 6 neglects many important properties of the el-ph interaction as it remains a long wavelength description of the el-ph coupling. More precisely, the phonon dispersion is approximated with a simple model of dispersionless optical phonons and a linear dispersion for the acoustic phonons. Moreover, the phonon momentum

dependence of the el-ph interaction is neglected which could lead to different results regarding the inter-valley scattering. Knowing that the latter is a major scattering mechanism at high carrier density, the question of the validity of such an approximation arises.

In order to accurately describe the el-ph interaction at room temperature, we here derive a method to include a DFPT based el-ph coupling in our *ab-initio* quantum transport approach [137, 2]. We show that our basis consisting of URBF is especially suitable for a DFPT based el-ph description. We first report calculations of el-ph limited mobility in the ML MoS_2, duly compared to the literature. Then, we apply our methodology to n- and p-type FETs and show comparisons with a simplified el-ph formulation based on an effective deformation potential.

7.2 Theory

7.2.A Electron-Phonon Interaction in the URBF basis

The formalism, presented in this chapter, assumes a 2-D semiconductor, where $\mathbf{r} = (x, y)$ is the in-plane coordinate and z is the out-of-plane direction. If not otherwise stated, we will also assume that the unit cell is orthorhombic, as illustrated in Fig. 7.1, and denote with a_x and a_y the lattice constant in the x and y direction, respectively. The Bloch functions of the system, $\Psi_{n,\mathbf{k}}(\mathbf{r}, z)$, can be obtained directly from the eigenvectors of the plane-wave DFT Hamiltonian provided by QE, with n and $\mathbf{k} = (k_x, k_y)$ being respectively the band index and the wave vector inside the 1st Brillouin zone. The reformulation of the derivations for a 3-D system is straightforward.

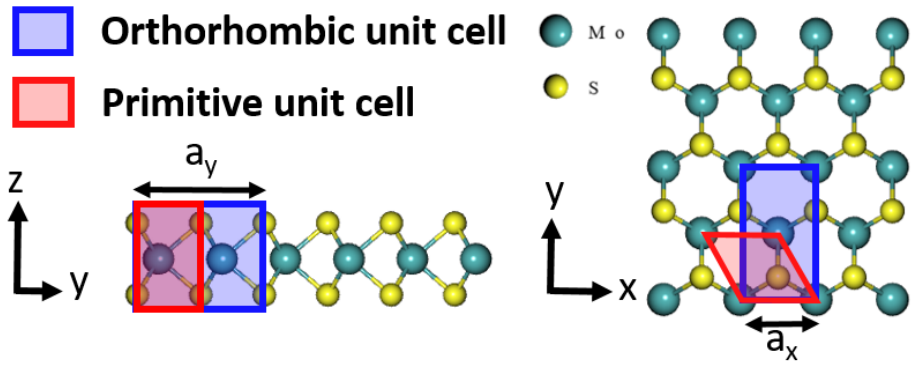

Figure 7.1: Atomic structure of monolayer MoS_2 with primitive cell (red) and orthorhombic unit cell.

We briefly recall the construction of the URBF basis from the Bloch functions. The Bloch functions $|\Psi_{n,\mathbf{k}}\rangle_{\text{u.c.}}$ restricted to a single unit cell along the transport direction x used in our basis are not orthogonal for different k_x (and same k_y), and in Eq. (7.2) we denote with $S(k_y)$ the corresponding matrix of the overlap integrals defined as

$$S_{n,k_x,m,k'_x}(k_y) = \langle \Psi_{n,k_x,k_y} | \Psi_{m,k'_x,k_y} \rangle \tag{7.1}$$

We then define new othogonal basis functions as

$$\Phi_{c,i,k_y}(\mathbf{r},z)=\begin{cases}\sum_{n,k_x}[S^{-\frac{1}{2}}(k_y)]_{n,k_x,i}\frac{u_{n,k_x,k_y}(\mathbf{r},z)}{\sqrt{N_{cy}}}e^{i\mathbf{k}.\mathbf{r}} \text{ for } c\ a_x<x<(c+1)a_x \\ 0 \text{ otherwise}\end{cases} \tag{7.2}$$

where c denotes a specific unit cell along x, the periodic parts $u_{n,k_x,k_y}(\mathbf{r},z)$ of the Bloch functions are normalized over a single unit-cell, and N_{cy} is the number of unit cells along the periodic direction y. The index i refers to a basis function corresponding to a given c and k_y , and each Φ_{c,i,k_y} is defined as an appropriate linear combination of N_B Bloch functions Ψ_{m,k_x,k_y} (restricted to the unit cell c along x). Since the overall number of $(n,\mathbf{k})$ states is N_B, $S(k_y)$ and $S^{-\frac{1}{2}}(k_y)$ are Hermitian matrices with rank N_B.

As for the el-ph interaction, our starting point is the definition of the matrix elements, $g_{m,n,\nu}(\mathbf{k},\mathbf{q})$, calculated by QE using the DFPT method:

$$g_{m,n,\nu}(\mathbf{k},\mathbf{q})=\sqrt{\frac{\hbar}{2\omega_{\nu,\mathbf{q}}}}\langle\Psi_{m,\mathbf{k}+\mathbf{q}}|\frac{\partial V_{KS}}{\partial\mathbf{U}_{\nu\mathbf{q}}}|\Psi_{n,\mathbf{k}}\rangle \tag{7.3}$$

where V_{KS} denotes the Kohn-Sham potential, $\mathbf{U}_{\nu\mathbf{q}}$ is the amplitude of the ν–th phonon mode and $\omega_{\nu,\mathbf{q}}$ is the phonon frequency. The matrix elements $g_{n\,m,\nu}(\mathbf{k},\mathbf{q})$ are calculated in the unit-cell by using the same $\mathbf{k}/\mathbf{q}$-grid used to compute the Bloch's functions .

The URBF basis set is particularly suitable for a DFPT description of the el-ph coupling. Indeed, the URBF basis is defined in Eq. (7.2) as a linear combination of Bloch functions in the unit cell, so that the corresponding matrix elements can be written in terms of the matrix elements $g_{m,n,\nu}(\mathbf{k},\mathbf{q})$. In Appendix B, we derive an expression for the electron-phonon interaction, $\hat{H}_{ep}$, that is local within a single unit cell, namely the $\hat{H}_{ep}$ couples only URBFs restricted to the same unit cell, and can be written as

$$\hat{H}_{ep}\approx\frac{1}{\sqrt{\Omega}}\sum_{j,i,k_y}\sum_{\mathbf{q},\nu}M^{\Phi}_{\mathbf{q},\nu}(j,i,k_y)(\hat{b}_{\mathbf{q},\nu}+\hat{b}^{\dagger}_{-\mathbf{q},\nu})\hat{a}^{\dagger}_{c',j,k'_y}\hat{a}_{c,i,k_y} \tag{7.4}$$

where $M^{\Phi}_{\mathbf{q},\nu}(j,i,k_y)$ is derived in Appendix B and is expressed in terms of the matrix elements $g_{m,n,\nu}(\mathbf{k},\mathbf{q})$ as:

$$M^{\Phi}_{\mathbf{q},\nu}(j,i,k_y)=\sum_{\substack{n,k_x\\m,k'_x}}\left[S^{-\frac{1}{2}}(k_y+q_y)\right]_{j,m,k'_x}g_{m,n,\nu}(\mathbf{k},\mathbf{q})\left[S^{-\frac{1}{2}}(k_y)\right]_{n,i,k_x}\delta_{k_x+q_x,k'_x} \tag{7.5}$$

In Eq. (7.4), $\hat{b}_{\mathbf{q},\nu}$ and $\hat{b}^{\dagger}_{-\mathbf{q},\nu}$ are the annihilation and creation operators, respectively, for a phonon in branch ν and with wave-vector $\mathbf{q}$, whereas $\hat{a}_{c,j,k_y}$ and $\hat{a}^{\dagger}_{c,j,k_y}$ are the corresponding operators for the URBFs restricted to the same unit cell c

7.2.B NEGF based transport model

For transport properties, the NEGF equations are solved with the Green-Tea solver [137]. The $\hat{H}_{ep}$ in Eq. (7.4) establishes the link between the lesser-than electron-phonon self-energy and the lesser-than Green's function for electrons. For computational reasons, we calculate only the diagonal elements of the lesser-than electron-phonon self-energy expressed in terms of the diagonal elements corresponding to the i-th Bloch function, $\Sigma^{<}_{i,i}(k_y, E)$, and of the lesser-than electrons Green's function, $\mathbf{G}^{<}_{i,i}(k_y, E)$. From the interaction potential $\hat{H}_{ep}$ in Eq. (7.4), the lesser-than self-energy reads:

$$\Sigma^{<}_{i,i}(k_y, E) = \frac{1}{A}\sum_{\nu,\mathbf{q}}\left[N_{\nu,\mathbf{q}} + \frac{1}{2} \pm \frac{1}{2}\right]\sum_{j}\left|M^{\Phi}_{\mathbf{q},\nu}(j, i, k_y)\right|^2 G^{<}_{j,j}(k_y + q_y, E \pm \hbar\omega_{\nu,\mathbf{q}}) \tag{7.6}$$

where A is just the normalization area in the x-y plane and the upper and lower sign describe respectively the phonon emission and phonon absorption process. $N_{\nu,\mathbf{q}}$ denotes the equilibrium phonon number, which is given by the phonon energy $\hbar\omega_{\nu,\mathbf{q}}$ through the Bose-Einstein occupation function.

While Eqs. (7.5-7.6) provide the desired DFPT-based el-ph coupling formulation. For comparison we will also use an el-ph treatment based on deformation potentials presented in the Chapter 6.

For the greater-than self-energies analogous equations apply.

$$\Sigma^{>}_{i,i}(k_y, E) = \frac{1}{A}\sum_{\nu,\mathbf{q}}\left[N_{\nu,\mathbf{q}} + \frac{1}{2} \mp \frac{1}{2}\right]\sum_{j}\left|M^{\Phi}_{\mathbf{q},\nu}(j, i, k_y)\right|^2 G^{>}_{j,j}(k_y + q_y, E \pm \hbar\omega_{\nu,\mathbf{q}}) \tag{7.7}$$

The electron-phonon retarded self-energy, Σ^r, was then obtained by using the conventional approximation consisting in considering only its anti-Hermitian part, namely we used

$$\Sigma^r \approx \Sigma^r_{AH} = \frac{\Sigma^{>} - \Sigma^{<}}{2} \tag{7.8}$$

with

$$\Sigma^{\lessgtr} = \Sigma^{\lessgtr}_{L} + \Sigma^{\lessgtr}_{R} + \Sigma^{\lessgtr}_{el-ph} \tag{7.9}$$

The real part of the retarded self-energy, which is related to the imaginary part by an Hilbert transform, is thus neglected and the impact from it is expected to be negligible for our device application as it would mainly result in an electron energy shift.

7.3 Results and analysis

A sketch of the atomic structure of monolayer 2H-MoS_2 is presented on Fig. 7.1, showing both the primitive and the orthorhombic unit cell used for transport calculations.

For the DFT and DFPT calculations in QE we employed the PBE exchange-correlation functional and the non-local van der Waals functional vdW-DF3 [31]. For DFPT, we also took into account the effects of dimensionality by using the recently developed DFPT framework for gated 2D materials [178]. The plane-wave energy cutoff was set to 60 Ry, and a 20×20×1 k-grid sampling was used for the electronic band structure. For the electron-phonon matrix elements and phonon dispersion, DFPT calculation was performed with an 8×10×1 q-grid sampling for the phonon dispersion. After a complete geometry optimization, the lattice parameter of ML MoS_2 is found to be a_{lat}=0.318nm, which is consistent with literature [128].

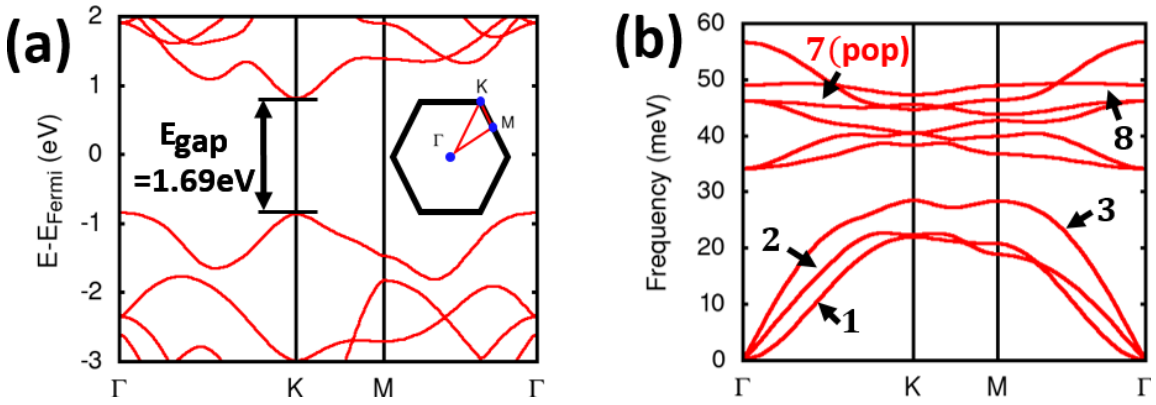

Figure 7.2: Electronic band structure and phonon dispersion computed for the primitive cell of 1ML MoS_2.

Figure 7.2 shows the MoS_2 band structure and the phonon dispersion computed for the primitive cell. In Fig. 7.2(a), an electronic direct bandgap of 1.69 eV is found. It is also important to note that the value of the energy separation between the K and Q valleys of 1L MoS_2 reported in the literature present a large spreading [91, 187, 180, 62] as it is sensible to the different flavour of the DFT calculation performed. In our work, an energy separation of $E_{K-Q} = 0.23$ eV was found, which is consistent with previous theoretical works [62, 180] and is close to ARPES and scanning tunneling spectroscopy (STS) measurements that provided a value of $E_{K-Q} = 0.24$ eV [103] . This quantity is important for diffusive transport since it will largely affect the inter-valley scattering in the material. In Fig. 7.2(b), with three atoms in the unit cell, we can remark the nine phonon branches of 1ML MoS_2 that are numbered from 1 to 9. For clarity, we indicated on the plot, the most dominant modes for the el-ph coupling. The three lower branches correspond to the three acoustic phonon modes. The first phonon branch present a quadratic dispersion relation close to the Γ point whereas the two other acoustic modes present a classic linear dispersion. The lowest acoustic mode corresponds to the flexural mode or ZA mode. For 2D materials with horizontal mirror symmetry, the electron-ZA phonon mode interaction is prohibited for symmetry reasons [58]. At higher energy, we can identify the six optical modes that are separated from the acoustic modes by an energy gap. Among them, the polar phonon branch, corresponding to the LO mode, can be recognized as the seventh branch around 48 meV.

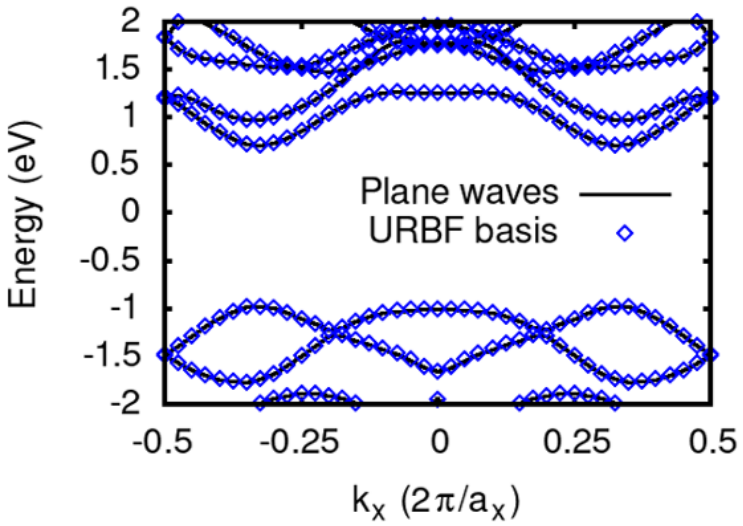

Figure 7.3: Band structure computed on a plane waves basis (black lines) and on the URBF basis (blue symbols) at $k_y = 0$.

Figure 7.3 shows band structure of the orthorhombic cell used for transport calculations for the case $k_y = 0$. The URBF basis used can achieve a very accurate band-structure reconstruction for MoS_2 with a basis size N_B=66.

Then, the el-ph matrix elements were computed from DFPT. Figure 7.4 shows the MEs of the el-ph interaction in MoS_2 along high-symmetry points paths for both acoustic and optical modes. It can be noticed that the polar mode (seventh branch in Fig. 7.2) close to the Γ point presents the largest value compared to the other modes. When the phonon momentum approaches the Γ point, the ME of the polar mode do not diverge in opposition to the 3D case. This effect is related to the behaviour of the long-range Coulomb interaction [177]. However, describing correctly the long-range nature of the polar mode at Γ point is out of range of our simulation as it would require to consider a large supercell. With DFPT calculations, we estimated the value of the el-ph ME associated with polar phonon mode by taking large phonon momentum (down to $|\mathbf{q}| = 10^{-3}\frac{2\pi}{a_0}$ in our calculation) until its does not change by increasing the momentum. The value for the monolayer MoS_2 was evaluated at 0.31 eV, which is consistent with a previous work [177]. We can remark that the polar optical mode is known to have a long-range coupling with electrons resulting in this high peak for the coupling at Γ point. However, all others modes also present a complex dependency with the phonon momentum and a much lower coupling value (<0.1eV). This highlights the importance of an accurate description of the el-ph interaction as a function of the exchanged momentum.

Our methodology offers a direct and accurate description of polar modes which have an influence on transport in TMDs and whose description can be more problematic with other approaches such as for MLWFs, which requires the use of an analytical formulation of the coupling with the polar mode. Indeed, the Wannier interpolation requires the assumption that the el-ph coupling is localized and slowly varying in real-space to describe it with MLWFs. As a result, the Wannier interpolation is not suited to correctly reproduce the DFPT el–ph ME of the polar optical phonon at small $\mathbf{q}$.

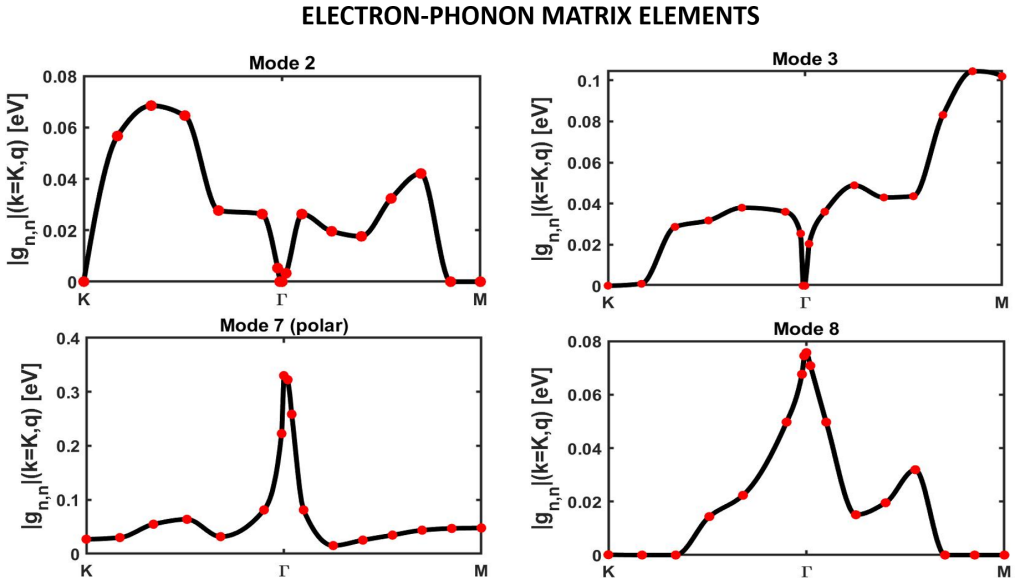

Figure 7.4: Absolute value of a few representative matrix elements of the electron-phonon interaction computed with the DFPT for the monolayer MoS_2. The MEs correspond to the lowest conduction band minima at $\mathbf{k} = K$ as a function of the phonon wave vector $\mathbf{q}$. Phonon modes 2 and 3 can be identified as transverse and longitudinal acoustic modes, respectively, mode 7 as the polar optical phonon mode, which is the longitudinal optical mode, and mode 8 as the transverse optical mode.

7.3.A Calculation of the phonon-limited mobility

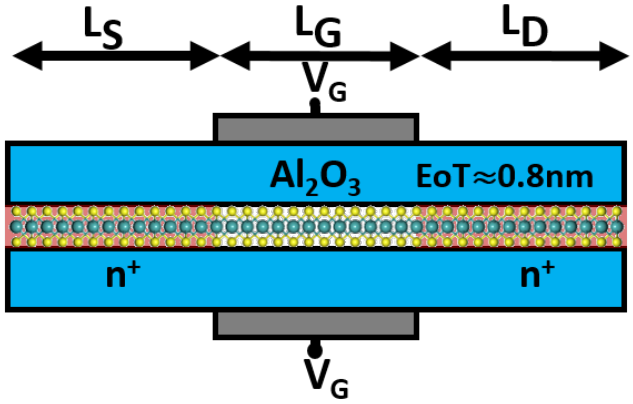

Figure 7.5: Sketch of MoS_2 double gate MOSFET where the red shaded region represents n-doped MoS_2.

For transport calculations, we considered the orthorhombic unit cell sketched in Fig. 7.1, and then simulated electron devices by combining many unit cells with different potentials along the transport direction x. Concerning the wave vectors of electrons used for the transport calculation, the transverse wave vectors, k_y for the electrons and q_y for the phonons were chosen to sample the first Brillouin zone with a constant step of $\Delta k_y = 0.1\frac{2\pi}{a_y}$.

The phonon-limited mobility of the ML MoS_2 was estimated at 300 K, according to the methodology detailed in [158], which has already been used to derive mobility with NEGF simulations [187]. In the linear transport regime, the total resistance of the device is estimated as the sum of the resistance of the intrinsic device channel and the

extrinsic resistance containing the resistance from the access regions. The resistance of the intrinsic device channel, R_{in}, can be expressed as a function of the mobility as:

$$R_{in} = \frac{L_{ch}}{e\mu_{\mathrm{ph}} n_{2D}} \tag{7.10}$$

where μ_{ph} is the phonon-limited mobility of the channel material, e is the absolute elementary charge of electron, n_{2D} is the 2D charge density in the channel region and L_{ch} is the length of the channel.

To access the value of the mobility, we can perform an NEGF simulation of a MOSFET at low V_{DS} and at a given charge density. This simulation provides us the total resistance of the device. By repeating the calculation for different channels lengths, we can estimate the first derivative of the total resistance with respect to the channel lengths. The extrinsic resistance should remain constant for different channel lengths as it includes only the resistive effects from the access regions, that are kept identical for all the calculations. Thus, it disappears in the derivative term. By differentiating Eq. (7.10), the mobility is then evaluated by considering only the intrinsic resistance of the channel with the following expression:

$$\mu_{\mathrm{ph}} = \frac{1}{e\ n_{2D}} \frac{1}{\frac{dR_{tot}}{dL_{ch}}} \tag{7.11}$$

The resistance is evaluated for the FET device sketched in Fig. 7.5 with an operating voltage of $V_{DS} = 10^{-2}$ V. In this device, we solve self-consistently the NEGF method equation with a non-linear Poisson equation to provide an accurate charge description in the device. The electron density of the channel is controlled by the gate voltage. In order to estimate the derivative of the resistance, we used three different lengths for the channel, i.e. 57 nm, 67 nm and 77 nm.

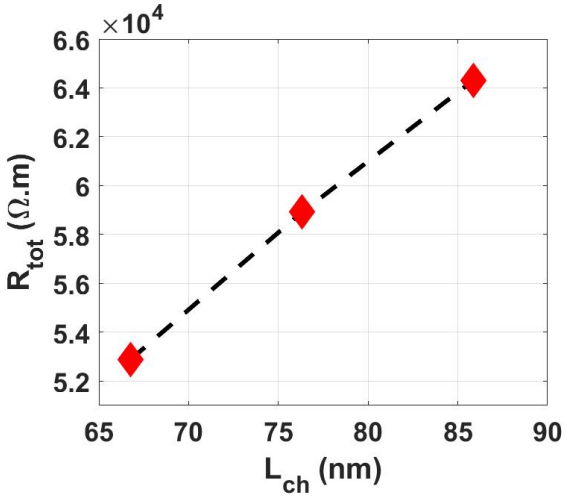

Figure 7.6: Channel resistance of a ML MoS_2 MOSFET sketched in Fig. 7.5 computed with a DFPT self-energy at $V_{DS} = 10^{-}2$ V and with a charge density of $n_{2D} = 10^{12}$ cm^{-2} in the channel region.

Fig. 7.6 shows the channel resistance computed for the three lengths with a charge density in the channel region of $n_{2D} = 10^{12}$ cm^{-2}. We can observe that the resistance increases linearly with the channel length. The linearity

confirms that we are dealing with a diffusive transport regime.

We used this kind of calculation to compute the mobility for three different cases. First, the el-ph self-energy was computed with the DFPT ME (see Eqs. (7.6-7.7) and we used it to compute the mobility in ML MoS_2. Second, we have checked the impact of the interaction with polar optical phonons by computing the mobility without the corresponding mode, for comparison. Lastly, the el-ph self-energy was computed from the deformation potential approximation as presented in Chapter 6. For the simulations employing the deformation potential approximation, we used data from literature with D_{op}=5.8 $\times$ 10^8 eV/cm with a corresponding phonon energy of $\hbar\omega_{op}$ = 48.9 meV for inelastic optical phonons, and D_{ac}=4.7 eV for elastic acoustic phonons where both results where estimated from DFPT calculation [108].

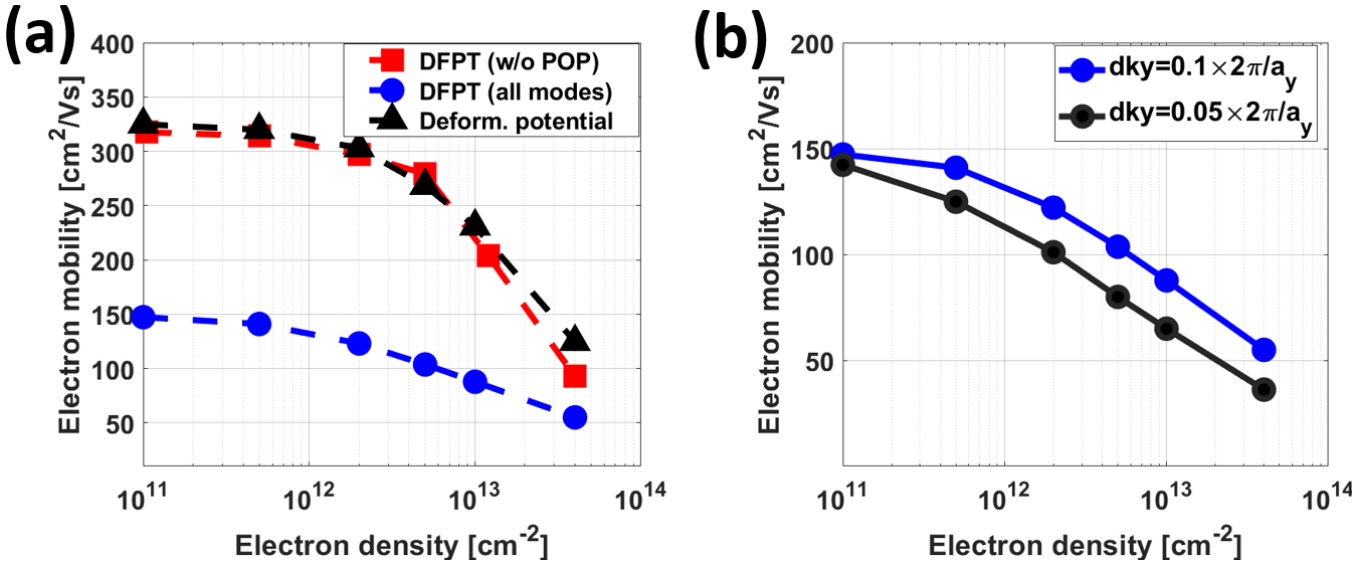

Figure 7.7: (a) Phonon-limited mobility of the monolayer MoS_2 computed by considering (squares) all the modes but the polar optical phonon (POP) one (mode 7 in Fig. 7.4), all the phonon modes (circles) and computed with deformation approximation (triangles). (b) Same but computed by considering all the phonon modes with a step for the transverse wave vector of $\Delta k_y = 0.1\frac{2\pi}{a_y}$ (blue) and $\Delta k_y = 0.05\frac{2\pi}{a_y}$ (black).

The results for the three cases are shown in Fig. 7.7(a), where the mobility is shown as a function of electron density. The black curve with triangles has been obtained with the deformation potential approach. It presents a plateau at low electron density close to 320 cm^2/Vs. When the electron density reaches an electron density of 10^{13} cm^{-2}, the mobility presents an important drop. This is related to the intervalley scattering as the Fermi level of the electrons is brought much closer to the Q valley, which enables the transmission from the K to the Q valley. The red curve with squares corresponds to the mobility calculation with the DFPT self-energy when we remove the polar mode. Its tendency is almost identical to the mobility computed with the deformation potentials. The blue curve with circles is the results of the mobility with DFPT calculation by considering all the modes and in particular the polar one. We can clearly remark that the interaction with polar optical phonons has a strong influence on the mobility, particularly at low electron densities where the mobility is halved compared to the case without the polar optical

mode. However, the latter is in fairly good agreements with the results obtained with the deformation potential approximation. This could have been expected as the deformation potentials were derived from DFPT calculation without the Fröhlich coupling. At high charge densities, the long range polar optical phonon is expected to be screened and the inter-valley scattering becomes dominant if another valley contributes to the electron transport. This may explain the mobility tendency resulting from our calculations where all three calculations tend to the same values at high electron density. Indeed, an inter-valley scattering is expected to occur at higher electron density when the Q valley represented in Fig. 7.2(a) is close to the Fermi level of the channel, making possible the inter-valley process between the K and the Q valley.

The convergence of the mobility with respect to the number of wave vectors is also investigated in Fig. 7.7(b). It was performed by using a denser grid with a 20×40×1 k-grid for electrons and an 8×20×1 q-grid for phonons. With these grids, corresponding to a sampling of the Brillouin zone with a constant step for the transverse wave vector of $\Delta k_y = 0.05\frac{2\pi}{a_y}$. As it can be seen with the plot of the mobility obtained for the two different **k**- and **q**- grids, the convergence of the mobility, with respect to the number of wave vectors, is not perfectly reached. However, the discrepancies between the two calculations remain small at low electron density with less than 5 $cm^2/(Vs)$ ($\approx 3\%$) of difference at $n_{2D} = 10^{11}$ cm^2, whereas we find a discrepancy of 24 $cm^2/(Vs)$ ($\approx 20\%$) at $n_{2D} = 5 \times 10^{12}$ cm^2.

Mobility [$cm^2/(Vs)$]		DFT	El-Ph	Ref.
n =$2{\cdot}10^{12}$	n=$1{\cdot}10^{13}$ [cm^{-2}]	calculations	Matr. Elem.	
400	343	LCAO, LDA	DP (FP)	[91]
368	340	LCAO, LDA	FP	[71]
$\sim$ 290	244	PW(QE) ONCV-PBE	DP (DFPT)	[143]
302	230	PW (QE) ONCV-PBE	DP (DFPT)	This work
297	203	PW (QE) ONCV-PBE	DFPT (w/o polar mode)	This work
168	—	PW(QE), LDA	DFPT, WI	[208]
Intrinsic, 147		PW (QE) ONCV-PBE	DFPT, WI	[62]
104	64	PW (QE) ONCV-PBE	DFPT	This work

Table 7.1: Electron mobility for MoS_2 calculated in this work and compared to recent findings. The methodology for the calculation of the bandstructure and el-ph MEs are also given. Acronyms: FP (Frozen Phonon), DP (Deformation Potential), LCAO (Linear Combination of Atomic Orbitals), LDA (Local density Approximation), Optimized Norm-Conserving Vanderbilt (ONCV), PBE (Perdew-Burke-Ernzerhof), PW (Plane-Wave), QE (Quantum Espresso), WI (Wannier Interpolation).

A comparison of the phonon-limited mobility computed in this work with some previous works is reported in Table 7.1. It is important to mention that the discrepancies for these calculations may find their origins in several factors. The method (EMA, TB, DFT...) to compute the band structure may results in different effective mass values, which play an important role for the mobility. The energy separation between the valleys is also crucial as it determines at

which charge density the inter-valley valley scattering will start. For the DFT calculations, the exchange-correlation functional and the inclusion of spin-orbit coupling have also a non-negligible effect on the mobility [62].

The chosen method used to compute the el-ph ME also introduces more variability. The Fröhlich interaction, being a long-range interaction, is difficult to compute by using a localized function basis. It includes all the frozen phonons methods and DFPT calculations that use a Wannier interpolation. To compensate the absence of the Fröhlich interaction in the calculation, it can be added by means of an analytical formula that fit the DFPT calculation at the long-range limit [176].

We can observe that our results neglecting polar optical modes and the ones obtained with the deformation potential approximation are in good agreement with the three first works reported in the Table 7.1. In Ref. [91], Kaasbjerg et al. used an effective mass approach to model the lowest conduction band in K with an effective mass of $0.48m_0$ along the longitudinal direction, which is slightly bigger than the $0.43m_0$ that we obtained. The el-ph MEs were computed from a DFT Hamiltonian described in a LCAO basis, which does not describe accurately the Fröhlich interaction for the same reason as with the MLWFs basis. The Fröhlich interaction was approximated by an analytical formula with Born effective charges. It was estimated at 98 meV for $\mathbf{q} = 0$ which is three times smaller than our value. From it, the authors derived two intra-valley deformation potential, one of 2.6×10^8 eV/cm, which is half of the value of 5.8×10^8 eV/cm that we used in our calculation and another for the homopolar optical mode with $D_{op} = 4.1 \times 10^8$ eV/cm. The inter-valley scattering between the K and Q valley has been neglected as they found an energy separation of 200 meV. The smaller value of intra-valley deformation potential compared to ours and the absence of inter-valley scattering at high electron density may be the reason of the higher values that they obtained compared to the other results.

Gunst et al. [71] and Pilotto et al. [143] performed to calculations where the Fröhlich coupling was neglected. In their work, Gunst et al. used a DFT band structure for MoS_2. The el-ph ME were computed with the frozen phonon approach which makes them neglect the Fröhlich coupling. For the mobility calculations, they considered only the intra-valley scattering. Because of the latter assumption, they did not observe a drop of the mobility at higher electron density. In their work, Pilotto et al. used the same deformation potential that we used from Ref. [108]. The two valleys at K and Q of ML MoS_2 were described by an effective mass approximation. They report an effective mass of $0.47m_0$ for the K valley along the longitudinal direction and an energy separation between the K and Q valley of 160 meV. This value is smaller than the energy separation that we obtained and it should result in a smaller mobility at high electron density than the one that we report. However, they used a multi-valley deformation potential approximation in their calculations, where the transition from K to Q valley corresponds to a deformation potential of 1.9×10^8 eV/cm and for the transition from K to Q', it corresponds to a deformation potential of 5.6×10^8 eV/cm. In our calculation with the deformation potential, only one value of 5.8×10^8 eV/cm is used which may result in an overestimation of the mobility at high electron density.

In all the previously discussed works, the inclusion of the el-ph interaction due to polar optical phonon mode was

neglected. When it is taken into account, smaller mobility values are obtained. It is linked to the fact that polar optical phonon modes are known to be the dominant scattering process at low charge densities. The works of Zhou et al. [208] presents results with el-ph ME computed from a Wannier interpolation of DFPT ME. As mentioned previously, this approach is not able to describe properly the Fröhlich interaction due to its long-range nature. For this reason, the ME of the polar phonon mode were evaluated analytically with Born effective charges and dielectric tensors for small values of $\mathbf{q}$. Unlike us, SOC coupling is considered for the DFT calculations in this work. However, Gaddemane et al. remarked that the SOC has negligible impact on electron mobility for Mo-based TMDs. No other information was provided about the DFT calculations or the band structure of MoS_2. However, the higher mobility value, compared to our calculation with DFPT ME, that they obtained may be explained by the use of effective screening terms in the long-range el-ph ME. For Gaddemane et al [62], the same method as Zhou et al. was used to compute the el-ph ME but with EPW [145] instead of their own software. However, it is unknown if they used the polar correction for their MEs. For MoS_2, the Hamiltonian was provided by DFT calculations. They present an exhaustive work on the impact of the different flavour of DFT calculation on phonon-limited mobility. As such, they provide calculation with the same configuration as our work. The properties of ML MoS_2 are relatively the same as us with an energy separation of 265 meV between the K and Q valley. They computed the mobility for the intrinsic material making the comparison difficult with our results. However, it is interesting to remark that their mobility value is close to ours at low electron densities. It suggests that mobility calculations, with all the ME computed from DFPT, generally result in lower values compared to the frozen phonon and the deformation potential approaches.

7.3.B Comparison with the deformation potential model

In order to highlight the importance of including the el-ph MEs formalism developed in this PhD thesis, a comparison with the ballistic transport in ML MoS_2 has been made. We simulated a n-type FETs using the architecture sketched in Fig. 7.5. We considered double gate FETs featuring a gate length L_G of 9 nm, access regions 13 nm long and with a doping density of 4×10^{13} cm^{-2}, and an Al_2O_3 gate oxide with a relative permittivity of $\epsilon_R = 9.3$ and a thickness of 1.9 nm resulting in an equivalent oxide thickness of 0.8 nm.

This comparison illustrates the importance to accurately describe the el-ph interaction in such a device, by showing the output characteristics, in Fig. 7.8, of the ML MoS_2 FET with and without the presence of phonon scattering. Ballistic simulation results in a large negative differential resistance effect. However, this effect has never been observed in experiments and therefore it is most likely an effect from the ballistic approximation, which is suppressed by the el-ph interaction.

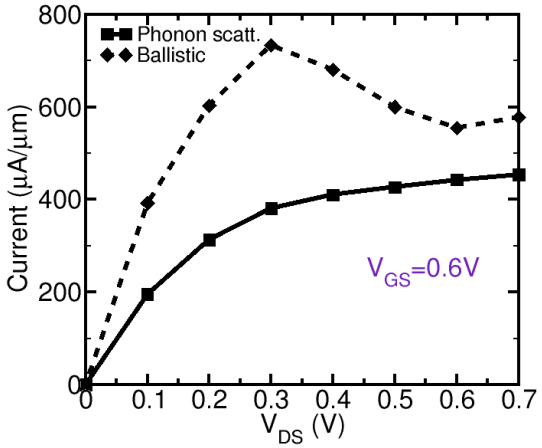

Figure 7.8: Output characteristics for V_{GS} = 0.6 V of the ML MoS_2 FET sketched in Fig. 7.5 obtained either for ballistic transport (diamonds) or by including the DFPT based el-ph scattering (squares).

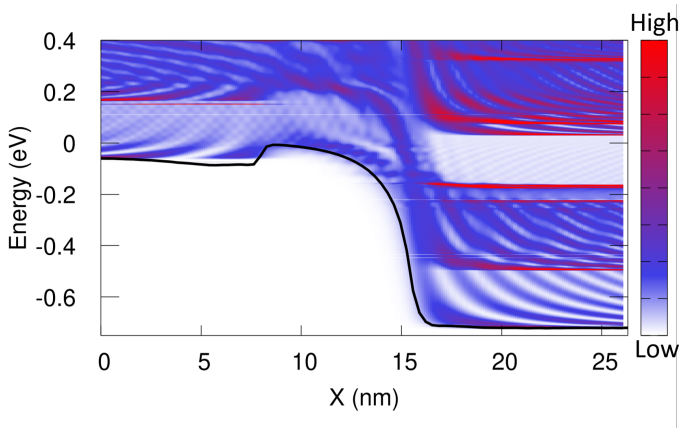

Figure 7.9: (Color) Local density of states of the conduction bands for V_{GS} = 0.6 V and V_{DS} = 0.7 V of the ML MoS_2 FET sketched in Fig. 7.5 obtained with ballistic approximation calculation.

To better understand the origin of the NDR observed, we plot the LDOS along the device in Fig. 7.9. This figure shows a minimum in the conduction band density of states of ML MoS_2 of the drain region (for x>15 nm) under coherent transport conditions which has already been reported [187] and linked to the ballistic approximation calculation. The low density of states in the drain region acts as a filter of the current in the ballistic regime when this zone is at the same energy as the electron injection energy in the source. By adding the el-ph interaction, the electrons are energetically relaxed in the drain region which results in the suppression of the NDR effect. For this reason, all simulations reported in the remainder of this section will thus include el-ph scattering with deformation potential for comparison with DFPT ME approach.

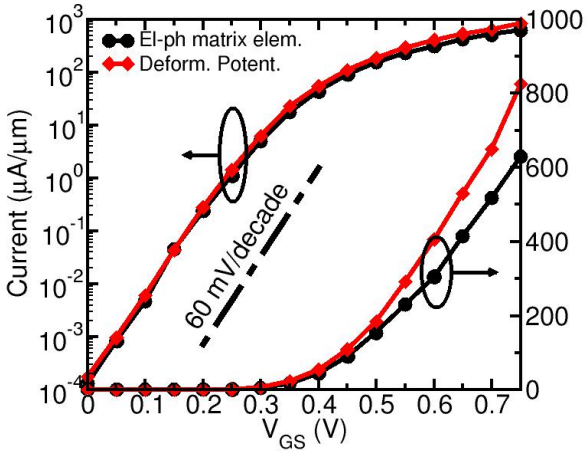

Figure 7.10: Transfer characteristics for $V_{DS} = 0.68$ V of the ML MoS_2 FET sketched in Fig. 7.5 computed with phonon self-energies using either the DFPT MEs (Eq. (5)) or the deformation potential theory (Eqs.(6-7)).

Figure 7.10 compares the transfer characteristics calculated by using the el-ph self-energies obtained either with the DFPT based MEs given by Eq. (7.6), or with the deformation potential approach. Even if the transfer characteristics obtained with the two methods in Fig. 7.10 are qualitatively similar, we have systematically observed a sizeable current reduction for the DFPT-based compared to the deformation potential-based description of the el-ph interaction, particularly at large overdrive regime, whenever a more accurate description of the el-ph interaction is included. Moreover, Figure 7.11 compares the current spectra at high gate voltages obtained with either the DFPT or the deformation potential formulations of the el-ph self-energies. As it can be seen, the simulations corresponding to the DFPT approach present a stronger electron relaxation in the drain region which may be explained by the strength of the polar phonon interaction. Moreover, we observe a wider and more complex rearrangement in energy and space of the scattered carriers in the drain region.

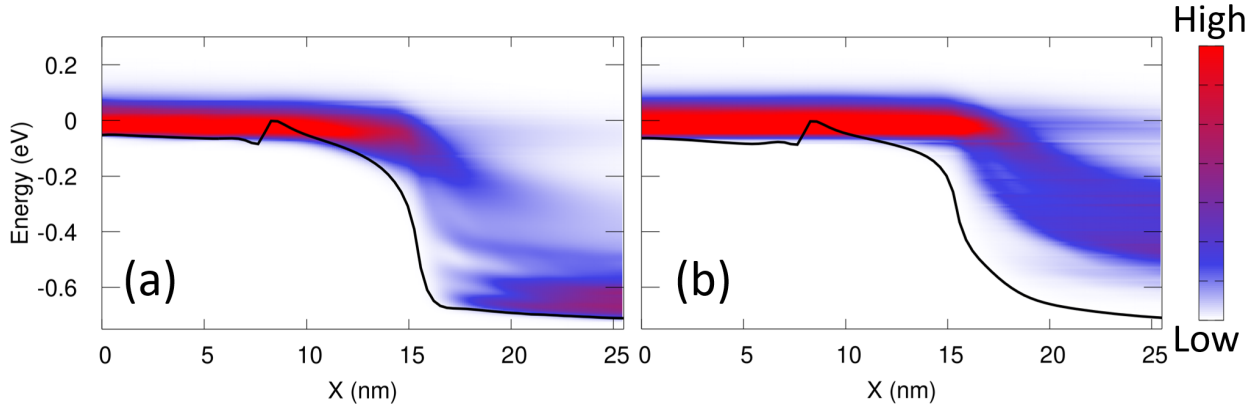

Figure 7.11: (Color) current spectra along the transport direction at $V_{GS} = 0.7$ V and $V_{DS} = 0.68$ V for a ML MoS_2 FET. a) simulations employing the DFPT based el-ph interaction formulation in Eq. (7.6); b) simulations employing the deformation potential approximation.

To investigate more closely the differences between the two approaches, we realized a simulation with a constant

electric field in the channel to observe the current density spectrum under a constant electric field. Knowing that we have a strong polar optical phonon, we choose an electric field of $\approx 3.5 \times 10^5$ V/cm as optical phonons are expected to be more important under this high electric field condition. Figure 7.12 clearly shows the different energy distribution of electrons achieved with both approaches. For the deformation potential approach, a step-like current density spectra is observed. It results directly from the dispersionless approximation for the optical phonon of this approach. Compared to the DFPT approach, where the phonon frequency is **q** dependant and all the modes are taken into account. The resulting electron energy relaxation is more continuous due to the contribution of a finite range of phonon frequencies, which is responsible for the differences in current spectra observed in Fig. 7.11. For the device, this difference in energy relaxation may lead to a different heat dissipation in the drain region, and may have more prominent effects on the terminal currents of tunnelling FETs that are very sensitive to phonon assisted tunnelling processes.

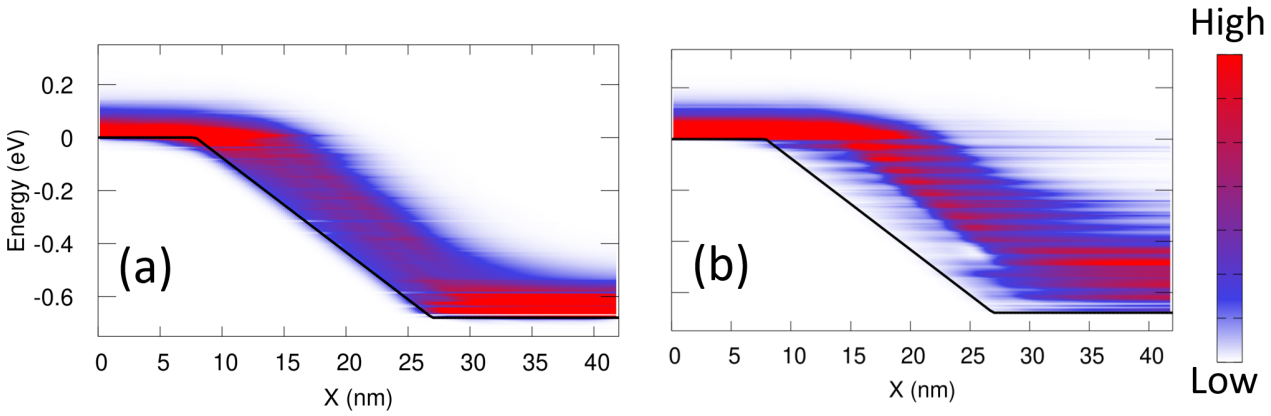

Figure 7.12: (Color) current spectra along the transport direction under a constant electric field of $\approx 3.5 \times 10^5$ V/cm for (a) simulations employing the DFPT based el-ph interaction formulation in Eq. (7.6); (b) simulations employing the deformation potential approximation.

7.3.C Simulation of CMOS FETs

CMOS technology based on n-type ML MoS_2 FET and p-type ML WSe_2 FET became popular among researchers since both materials exhibit relatively large mobility and high $\frac{I_{ON}}{I_{OFF}}$ ratio [146, 147, 93, 153]. As mentioned in Sec. 1.5.B, recent progress have been made for n-type ML MoS2 FET, with contact resistance of only 123 Ω μm between ML MoS_2 and semi-metal bismuth [171], and a p-type with a contact resistance of 3.3 kΩ μm [202] for ML WSe_2 and further developments are expected in this area. These results encourage to continue the work on the intrinsic properties of 2D FET where electron transport in the channel can be strongly impacted by scattering mechanisms. Our new approach is suitable to perform this predictive work with a rigorous description of the el-ph interaction inside the device.

In this regards, we simulated n- and p-type FETs using the architecture sketched in Fig. 7.5. We target low-power

transistors as recently discussed in [38]. Mono-layer MoS_2 and WSe_2 are used as channel material for n- and p-type FETs, respectively. We considered double gate FETs featuring a gate length L_G ranging from 9 nm to 5 nm in order to investigate the scalability of the device. Access regions are 13 nm-long and with a strong doping density of $N_D = N_A = 4 \times 10^{13}$ cm^{-2}, and an Al_2O_3 gate oxide with a relative permittivity of $\epsilon_R = 9.3$ and a thickness of 1.9 nm resulting in an equivalent oxide thickness of 0.8 nm . The design of the device presented is taken from Chen et al [38] where this configuration should result in performances fulfilling the criteria of IRDS for the next-generation, low-power, logic device. Thanks to the generality of our *ab-initio* simulation approach, it is easy to change the device material and to focus on the ML WSe_2 as a promising candidate for a p-type FET. Similarly, to ML MoS_2, the el-ph matrix elements were computed from DFPT. It is important to mention that in ML WSe_2, the spin-orbit coupling results in a large band splitting of $\approx$ 500meV for the top of valence band. Since we neglected the spin-orbit coupling in our DFT calculations, the transport properties for ML WSe_2 were computed by removing the spin degeneracy factor which should results in a better estimation of the device performances. Indeed, the contribution to the current of the second valence band of should be negligible due to the large splitting.

The $I - V$ curves are presented in Fig. 7.13 for ML MoS_2 and Fig. 7.14 for ML WSe_2, respectively. As it can be seen, both FETs employing 2-D materials exhibit an excellent electrostatic integrity even at $L_G \approx 7$ nm, with a sub-threshold swing (SS) of about 65 mV/dec for ML MoS_2 and about 63 mV/dec for ML WSe_2. The onset of a sizeable SS degradation is pushed down to $L_G \approx 5$ nm where a SS of 86 mV/dec is found for both devices. Such an ideal sub-threshold operation can be explained with the sub-nanometric thickness of the channel material, as well as with the relatively large transport masses, $0.43m_0$ for the lowest conduction band of ML MoS_2 and $0.35m_0$ for the highest valence band of ML WSe_2, compared to silicon, which can effectively contrast the source-to-drain tunnelling.

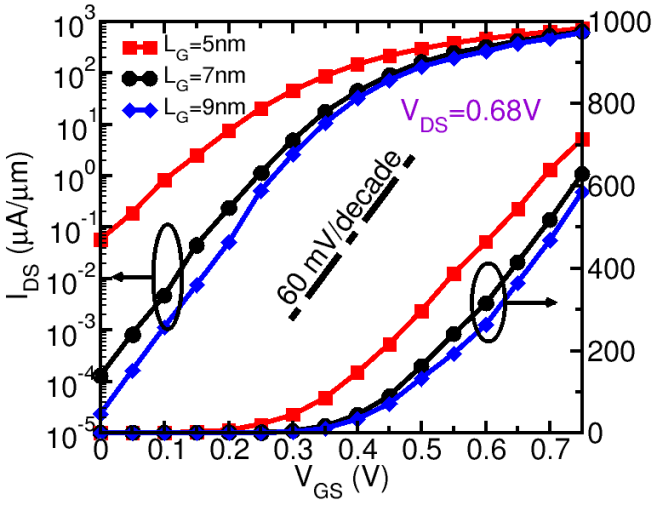

Figure 7.13: Transfer characteristics at $V_{DS} = 0.68$ V for the ML MoS_2 based n-type FETs featuring a gate length $L_G \approx 5$ nm (squares), $L_G \approx 7$ nm (circles) and $L_G \approx 9$ nm (diamonds).

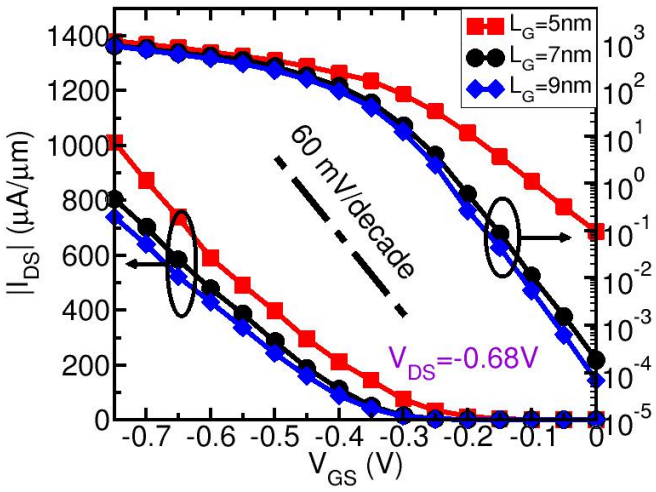

Figure 7.14: Transfer characteristics at V_{DS} =-0.68 V for the ML WSe_2 based p-type FETs featuring a gate length $L_G \approx 5$ nm (squares), $L_G \approx 7$ nm (circles) and $L_G \approx 9$ nm (diamonds).

For the ON-state, both devices exhibit large ON-current even in the presence of el-ph interaction. The ON-current of both devices are close to the targeted value of 656 μA/μm fixed by IRDS for year 2028 with 518 μA/μm for ML MoS_2 and even larger for ML WSe_2 with 717 μA/μm. The ratio between ON and OFF current is evaluated to be $\frac{I_{ON}}{I_{OFF}} = 5.10^6$ for ML MoS_2 and ML WSe_2 for the gate length of 7 nm. This is remarkable as the impact of the el-ph interaction for the ON-state results in a decreasing ON-current. With the relatively low mobility obtained for ML MoS_2 at high electron density in Fig 7.7, we could have expected a lower ON-current. For this CMOS, the p-type presents a higher ON-current. This can be explained by the contribution to the current of several valleys in 1ML WSe_2 compared to 1ML MoS_2, which increases the density of states and thus the current.

With these results, both devices show a high-ON current and SS that satisfy the IRDS requirement projections for 2028 and are satisfying in term of scalability as they provide more room for development with a more aggressive EOT. In Sec. 7.3.A, our work showed that the main scattering process in the simulation was coming from the coupling with the polar optical phonon mode. In this regards, we stress again that we omitted to include the electron screening on the el-ph ME which is expected to reduce the long-range potential coupled to electron. This may potentially result in higher current for our devices.

7.4 Summary

In this chapter, we have presented the theory and implementation of an original methodology to include a full *ab-initio* treatment of el-ph interaction based on the DFPT approach in self-consistent NEGF simulations of electronic systems. This approach is general and can be applied to 2D systems as well as 3D one. It provides a rigorous description of the phonon dispersion as well as the phonon momentum dependence for electron-phonon matrix elements which is neglected with a simple method like deformation potential. This information is necessary for predicting the physics of

materials which present a large anisotropy or several valleys at low energy since the phonon momentum dependence plays an important role in the inter-valley scattering.

Our findings show that the URBF basis set is especially suitable for a DFPT based el-ph description. The generality of our approach is illustrated by showing both mobility calculations and complete simulations of n- and p-type FETs consisting of mono-layer TMDs. Mobility calculation highlights the strong coupling to the polar phonon in ML MoS_2 at low electron density. Compared to literature, the mobility is slightly smaller where the mobility without the polar optical phonon mode gives similar results.

A comparison between the el-ph description with the deformation potential approximation has been performed. Results show that a not negligible difference is obtained for the current. The current spectrum along the device presents a larger discrepancy between the two methods, with a more complex and stronger energy relaxation for the el-ph description and it could result in larger differences in physical systems other than the MOSFET, when the phonon assisted tunnelling plays a predominant role. It may also be important for the self-heating effects, which are expected to produce an increase of the phonon population resulting in a degradation of the current. This phenomena should be influenced by the difference of rearrangement in energy and space of the scattered carriers in the drain region presented in Sec. 7.3.B, which is sensibly different for the two methods used.

Ultimately, the simulation of the CMOS FETs based on a n-type ML MoS_2 FET and a p-type ML WSe_2 FET reveal that even under strong scattering mechanism, they remain a promising candidates for the next technological nodes due to their scalability and excellent electrostatic control illustrated from our results.

As a perspectives, it would be interesting to study the self-heating emerging from these calculations by coupling the kinetics equations of electron with the heat equation in order to evaluate the self-heating of such device as it is known that the heat dissipation plays a major role in the device performance [203].

A better estimation of the el-ph matrix elements could also be addressed by trying to include the non-locality of polar phonon in the self-energy and adding the screening from the carriers, which is known to reduce the polar phonon coupling at high electron density. However, the advantages of our approach is that it is relatively simple to add more physics into the model, since we are working directly with the matrix elements of the el-ph interaction.

Chapter 8

Summary and perspectives

In this PhD thesis, the development of advanced physical models for quantum transport in nanoscale devices was carried out. All models were used to simulate and explore technologically relevant devices from III-V TFETs to 2D based MOSFETs and TFETs. The developed models were based on pseudopotential Hamiltonians computed either with an empirical approach (EPM) or an *ab-initio* approach with the use of the DFT. A special focus was given on the modelling of heterostructures and the inclusion of the electron-phonon interaction in our transport calculations.

In the first Chapter, an outline was given on the history of electronic devices for logic applications and the state of the art of device simulation. First, we summarized the history of MOSFET, its working principle and its limitation related to the aggressive shrinking that the industry tries to maintain. As it reaches the nanoscale, quantum phenomena and electrostratics degradation start to arise and deteriorate the performance of the device under the generic term of short-channel effects (SCE). The possible perspectives for electronic devices are presented. Among them, the use of III-V materials or 2D materials are promising options. They possess electrical properties that may help to design efficient electron devices, in particular for the TFETs, since the low effective mass of some III-V and 2D materials can help to achieve high ON-currents. To evaluate the potential of these materials, an accurate material description is required, and to this aim, a summary of the different methods is presented. The model requires to be computationally efficient and accurate enough to describe ultra-thin structure. We choose the EPM and DFT method to explore the different materials investigated in this thesis. They are both based on the pseudopotential approach. It means that the development done with one method can be extended, with some corrections, to the other one. They both provide a full band description of the materials. EPM is computationally more efficient due to the use of empirical parameters, which requires a basis set with a smaller number of elements. However, it is also the reason of a limitation for the method since the empirical parameters are not available for all the materials (2D materials for example). DFT, as an *ab-initio* approach, does not suffer from this limitation, but it demands more computational resources. From the material modeling, we obtain the Hamiltonian that is required to simulate the device, while, to compute the transfer characteristics, we rely on the NEGF method. This formalism provides an

accurate evaluation of the out-of-equilibrium properties of our device. It takes into account quantum phenomena (tunnelling, coherence,..), which are relevant in nanoscale devices.

In Chapters 2 and 3, we present the theoretical basis used in the thesis. Chapter 2 focuses on the material description aspect. The basic ideas behind pseudopotential Hamiltonians are presented. The plane-wave basis is chosen in our work as it provides several advantages, like an accurate atomic relaxation and a simple way to improve the completeness of the basis. It allows us to develop the model of EPM and DFT for electrons and DFPT for phonons. In chapter 3, the second quantization is used to derive the Green's functions and the NEGF formalism. The Green's functions formalism allows us to compute the dynamical (transmission, scattering rates,..) and statistical (current, carrier density,..) properties. The development of the formalism is the first step before the device simulation. In order to solve the kinetic equations of NEGF formalism with the previously derived Hamiltonian in the plane-wave basis, the Hamiltonian has to undergo two transformations. The first one consists in transferring the plane-wave basis into a hybrid basis with real space in the transport direction and reciprocal space in the transverse directions. In this basis, the Hamiltonian is a block tridiagonal matrix which allows us to use computationally efficient recursive algorithm to calculate the Green's functions. This transformation is also required to be able to connect the device region simulated to two semi-infinite leads where we inject and collect the current. The second one focuses on reducing the rank of the matrix. We presented two methods to achieve that: the mode-space and the unit-cell restricted Bloch functions basis. The first one consists in a basis of the selected eigenvectors computed from the Schrödinger equations in a section of the material and at $k_x = 0$. Because of that, the mode-space basis works correctly to reduce the size of the matrix if the Hamiltonian is local, hence, it fits perfectly with the local EPM Hamiltonian. The URBF basis solves the problem of the non-local DFT Hamiltonian by considering unit-cell restricted Bloch functions. This basis was recently proposed to use DFT Hamiltonian in quantum transport simulation. Lastly, the interactions with external perturbations is presented by means of the self-energies. We present the self-energies describing the coupling with the two semi-infinite leads and with phonons in a deformation potential approach.

In the following chapters, the obtained results were presented. Generally, we focused on two aspects of the modelling for nanoscale devices. The first one, concerning chapters 4 and 5, deals with the description of heterostructures with pseudopotential Hamiltonians, heterostructures being a promising technological option to design efficient device. After the presentation of the methodology developed to model heterostructures, we focus on the inclusion of electron-phonon interaction which generally has a non-negligible impact on the device performance.

In Chapter 4, we studied a specific III-V heterostructure (GaSb and InAs) for tunnel devices applications since it presents a suitable type III band alignment which improves the ON-current by enhancing the transmission probability. An Esaki diode as well as an n-type and a p-type TFETs designed with this heterostructure were studied by means of EPM Hamiltonians. To describe more accurately the interface, we developed a model to mimic the transition of materials at the interface within the virtual crystal approximation. In the TFET cases, we also included a local confinement operator to simulate ultra-thin channels. For both devices, Esaki diode and TFETs, different lengths of

the gradual transition were investigated. Generally, it was found that the transmission properties were degraded due to an increase of the tunnelling path induced by the transition region. In the Esaki diode cases, a comparison with recent experimental measurements was realised. A quite satisfying agreement was found between our simulations (current density slope in the range of 14-20 MA/(V.cm^2)) and the experimental measurements (current density slope of 17 MA/(V.cm^2)) in the reverse bias region. However, the peak-to-valley current ratio is highly overestimated in our results due to the absence of trap assisted tunnelling, which is expected to provide more realistic values of valley-current. Both n- and p-type TFETs present a decrease of the ON-current when increasing the length of the transition region. The n-type TFET provides a steep sub-threshold swing of 30 mV/dec, whereas the p-type performance is slightly below 60 mV/dec. For the p-type, the current is relatively low at the ON-state with 148 μA/μm. For the n-type TFET, it is the opposite with an ON-state current that is almost twice the value of the p-type. This is mainly due to the large degeneracy of the source region of the p-type TFET, which is also responsible of the higher SS compared to the n-type. The approach developed in this chapter was useful to develop the next steps.

In Chapter 5, the theoretical development focused on the study of heterostructures with *ab-initio* Hamiltonian. For newly found or understudied materials (2D materials in our case), no empirical data for EPM are accessible in the literature. An alternative is provided by DFT simulations. As also in this case we use a pseudopotential Hamiltonian computed on a plane-wave basis, the same approach was used to compute the heterostructure. The main difference resides in the coupling matrix at the interface. To estimate it, we proposed two methods, the first one consists in using the coupling matrix of one of the two isolated materials that constitute the heterostructure; the second transformation is based on the idea that we can extract the coupling matrix between the two materials in a Hamiltonian of a supercell which includes the interface. This approach should provide a better approximation compared to the first method as the coupling matrix is computed from DFT. To test out the two methods, we investigated two different systems. First, we simulated a MOSFET made of 1ML and 2ML $PtSe_2$ heterostructure in order to exploit its property to switch from semiconductor to metal by adding more layers. Although simple, the first method turns out to be close to the second method in the case of the multi-layer $PtSe_2$. For such a system, the two methods provide similar results as a consequence of the the fact that the two materials are related and that the current is essentially lateral and involves only atoms directly connected at the interface.

The second application consists in a lateral heterostructure composed of 1ML $PdSe_2$ and 1ML SnS_2. In this case, the results of the two methods are different. The calculation with the coupling matrix extracted from the supercell presents the lowest ON-current. In this case, the origin is directly related to the coupling matrix as the bands of the different cases are strictly the same. The results obtained in this part present interesting performance for the multi-layer $PtSe_2$ FET. However, for the TFET, the electrical performance is much worse than that of the III-V TFET simulated in chapter 4. Moreover, it is important to note that these results are obtained in a ballistic approximation, which provides a best case scenario for the device performance. The origin for the low current of the TFET is in the large effective mass and the tunneling path that the carriers have to cross. To improve the situation, a solution with

vdW heterostructures is particularly interesting as it consists to close the gap in the channel of a TFET to provide large ON-current. This is realised with the vdW heterostructure that we studied in the next chapter.

In Chapter 6, as we decided to continue to study technologically relevant devices such as the TFETs made of lateral and van der Waal heterostructures, we developed a model to study 2D materials heterostructure with DFT Hamiltonians. The materials used in this study are the 1ML 1T-$HfSe_2$ and the 1ML 1T-SnS_2 as they form an heterostructure with a type II band alignment close to a type III. DFT calculations indicates that a band inversion between the top of the valence band and the bottom of the conduction band is possible for the van der Waals heterostructure under the influence of an external electric field. Moreover, since the electron-phonon interaction is known to be relevant in these devices, we developed a model to compute deformation potentials for electron-acoustic phonon and electron-optical phonon coupling. In the two materials, the phonons and electron-phonon coupling properties computed with DFPT indicate that the polar phonon mode is the dominant mode among the optical phonon modes. The el-ph interaction modeling is simple as it considers that the deformation potential is independent of the phonon momentum, dispersionless optical phonon branches and linear dispersion for the acoustic modes branches. Our simulations suggested that the van der Waals TFET provides better transport properties than the lateral TFET even if it is more sensible to the el-ph coupling. The el-ph interaction gives rise to an interesting effect in these devices, resulting in higher current compared to the ballistic case. An investigation points out the origin of this effect as an intervalley transmission enabled by the el-ph interaction. This allows the ON-current to reach relatively high values for the vdW TFET with a current of 580 μA/μm at V_{DS}=0.35 V. For the lateral heterostructure, the ON-current was also increased by the phonon coupling, however, it remains lower with a value of 3 μA/μm. This is related to the relatively large tunnelling path, which disappears in the vdW heterostructure when it attains the band inversion. In ballistic approximation, both device architectures present a very steep SS of 15 mV/dec. However, when the el-ph couplins is turned on, the *SS* of the lateral TFET only increases up to 24 mV/dec, whereas in the van der Waals TFET it reaches 40 mV/dec. Even if its *SS* is more degraded by the presence of phonons, the vdW TFET remains the most promising option since it satisfies the recommendations of IRDS that give the target of an ON-state current higher that 500 μA/μm and a *SS* lower that 65 mV/dec.

Finally, in the Chapter 7, we worked to improve the model used to describe the el-ph interaction. Indeed, the lack of independence on the phonon momentum of the matrix elements and the simplified phonon bandstructure may result in an inaccurate representation of the phonon scattering. To improve the model used in Chapter 6, we exploited the fact that our basis composed of unit cell restricted Bloch functions is perfectly suitable to incorporate the el-ph ME computed from DFPT. Using el-ph matrix elements from DFPT calculations to compute the el-ph self-energy is arguably one of the most accurate methods to model this interaction. In this model, we obtained a self-energy that is dependent over the electron momentum and also the phonon momentum and mode. Even the phonon frequency keeps its dependence over the phonon momentum and the phonon mode. We apply the model to study the use of 1ML MoS_2 in *n*-type FETs, as it is one of the most studied material in literature. The first comparison is

performed by computing the phonon-limited mobility of 1ML MoS_2. A general comparison with literature is carried out. We observe that the polar phonon mode in our simulation induces a large mobility reduction at low carrier density compared to the case without it. Our results without the polar phonon mode are close to the results in literature, where the polar phonon mode is lacking, with a mobility close to 300 $cm^2/(V.s)$ at a carrier concentration of 2×10^{12} cm^{-2}. When we turn the polar mode on, our results decrease to 111 $cm^2/(V.s)$ at a carrier concentration of 2×10^{12} cm^{-2}. However, this is still consistent with mobility calculations where el-ph ME computed from DFPT are used. Then a comparison with the deformation potential approach is done for a MOSFET device. The results show that the DFPT model provides a stronger effect to the el-ph coupling and it results in a more complex energy relaxation. Finally, we close our study by using the generality of our model to simulate technologically relevant CMOS FETs where we used 1ML MoS2 for the n-type FET and WSe_2 for the p-type FET. The DFPT el-ph self-energy is included in the simulation and the scaling of the gate length is analysed. It results that even with the el-ph interaction, both devices with a gate length of 7 nm present promising performances that fulfil the IRDS prediction for 2028 with ON-current of 717 $\mu A/\mu m$ for WSe_2 and 518 $\mu A/\mu m$ for MoS_2. An excellent electrostatic control is found in the two devices with a *SS* of 63 mV/dec for WSe_2 and 65 mV/dec for MoS_2. These results highlight the potential of these materials for next generation electronics.

At the end of this PhD, some questions remain to be answered in the future. These questions open perspectives for future works in this field.

First, on the device aspect, it would be interesting to apply the methodology developed in the chapter 7 to study tunnel devices like the one of chapter 6. It is unknown if it will result in the same current increase by including the phonon coupling. As we have seen in the chapter 6, the impact of the phonons is non-negligible and as such it is necessary to employ an accurate description of the el-ph interaction. It would also be interesting to use it to study other paradigms not mentioned in this thesis. For instance, the cold-source FET presents excellent performance in the ballistic approximation [150, 111]. However, depending on the material used, the impact of the el-ph interaction may help [115] or degrade [51] the performance of this device. For TFETs, it is important to develop a model to take into account the presence of point-like defects, which are responsible of important *SS* degradation [142].

A second perspective on our work concerns the improvement of the model for the el-ph interaction. Several approximations have been used in order to implement the el-ph ME self-energy. However, we had to assume the el-ph interaction to be local, at least within the unit cell used for transport calculations. This is not true in the case of the polar phonons that produce a non-local interaction. Taking the non-local nature of the interaction into account would results in the loss of the block tridiagonal matrix preventing us from using the recursive Green's function algorithm. However, it is possible to keep a local approach and to correct the impact of the approximation by using a scaling factor [129]. Another approximation is also made by considering bare el-ph ME in our calculations, but it is known that the potential generated by the phonons is sensible to the screening from the electrons, which results in lower value for the el-ph ME. The alleviation of such approximations would lead us to address a more realistic

description of inelastic scattering in nanoscale devices based on 2D materials.

Appendix A

First Appendix: Recursive methods for Green's functions

From the definition of the retarded Green's functions in Eq.(3.39), computing $\mathbf{G}^r$ would require to inverse the Hamiltonian which is computationally intractable in most cases. Moreover, by looking at the formula of the observables, it can be remarked that all the elements of the Green's function are not necessary. A solution is to employ a recursive method by using the Dyson's equation to compute the lesser-than and retarded Green's function diagonal elements.

For the retarded Green's function, we start by decomposing the device in N_C cells along the transport direction x. The first step begin with the first layer of the device which is just connected to the semi-infinite leads through the self energy Σ_L. It begins the calculation of the "left connected" Green's function which only take the interaction from the left cells into account. For the first cell the Green's function writes as:

$$\mathbf{g}_{1,1}^{left} = [E\mathbf{I} - \mathbf{H}_{1,1} - \mathbf{\Sigma}_L] \tag{A.1}$$

Where $\mathbf{H}_{1,1}$ is the corresponding Hamiltonian of the cell. Then by using the Dyson equation from Eq.(3.37) to add the interaction of the first cells to the next ones with the equation:

$$\mathbf{g}_{i,i}^{left} = [E\mathbf{I} - \mathbf{H}_{i,i} - \mathbf{H}_{i,i-1}\mathbf{g}_{i-1,i-1}^{left}\mathbf{H}_{i-1,i}] \tag{A.2}$$

Which an iterative step that use the previously cell Green's function "left connected" and that started with Eq.(A.1). When the most right cell is reached, it is possible to compute the correct Green's function since the interaction from the right side comes from the semi-infinite lead that is known. Thus the Green's function writes as:

$$\mathbf{G}_{N_C,N_C} = [E\mathbf{I} - \mathbf{H}_{N_C,N_C} - \mathbf{H}_{N_C,N_C-1}\mathbf{g}_{N_C-1,N_C-1}^{left}\mathbf{H}_{N_C-1,N_C} - \Sigma_R] \tag{A.3}$$

Since we know the Green's function at the right of the device, we can now link the interactions from the right cell by starting the the last cell. It gives us the correct Green's function for every cells and it expresses as:

$$\mathbf{G}_{i,i} = [\mathbf{g}_{i,i}^{left} - \mathbf{g}_{i,i}^{left}\mathbf{H}_{i,i+1}\mathbf{G}_{i+1,I+1}\mathbf{H}_{i+1,i}\mathbf{g}_{i,i}^{left}] \tag{A.4}$$

Where it is computed one after the other by starting form $\mathbf{G}_{N_C,N_C}$.

For the lesser-than Green's function, similarly, we start with the "left connected" Green's function and by using Eq.(3.41), we define the lesser-than Green's function "left connected as:

$$\mathbf{g}_{1,1}^{<,left} = \mathbf{g}_{1,1}^{left}\Sigma_L^{<}\mathbf{g}_{1,1}^{\dagger,left} \tag{A.5}$$

Then we do the same for each cells where we take into account only the interaction coming from the left side:

$$\mathbf{g}_{i,i}^{<,left} = \mathbf{g}_{i,i}^{left}[\mathbf{H}_{i,i-1}\mathbf{g}_{i-1,i-1}^{left}\mathbf{H}_{i-1,i}]\mathbf{g}_{i,i}^{\dagger,left} \tag{A.6}$$

Once we reach the last cell, we also connect it to the semi-infinite lead:

$$\mathbf{G}_{N_C,N_C}^{<} = \mathbf{G}_{N_C,N_C}[\mathbf{H}_{N_C,N_C-1}\mathbf{g}_{N_C-1,N_C-1}^{left}\mathbf{H}_{N_C-1,N_C} + \mathbf{Sigma}_R^{<}]\mathbf{G}_{N_C,N_C}^{\dagger} \tag{A.7}$$

Similarly to the retarded Green's function, it can be used to compute the lesser-than Green's function recursively as:

$$\mathbf{G}_{i,i}^{<} = \mathbf{g}_{i,i}^{<,left} + \mathbf{g}_{i,i}^{left}[\mathbf{H}_{i,i+1}\mathbf{G}_{i+1,i+1}^{<,left}\mathbf{H}_{i+1,i}^{\dagger}]\mathbf{g}_{i,i}^{\dagger,left} + \mathbf{g}_{i,i}^{left}\mathbf{H}_{i,i+1}\mathbf{G}_{i+1,i}^{\dagger,left} + \mathbf{G}_{i,i+1}^{left}\mathbf{H}_{i+1,i}^{\dagger}\mathbf{g}_{i,i}^{\dagger,left} \tag{A.8}$$

Lastly, for the current density, the off-diagonal term of the lesser-than Green's function are needed and they can be computed from the elements that we already derived as:

$$\mathbf{G}_{i,i+1}^{<} = \mathbf{G}_{i+1,i+1}\mathbf{H}_{i+1,i}\mathbf{g}_{i,i}^{left} + \mathbf{G}_{i+1,i+1}^{<}\mathbf{H}_{i+1,i}^{\dagger}\mathbf{g}_{i,i}^{\dagger,left} \tag{A.9}$$

Which conclude our recursive calculations for the Green's function.

Appendix B

Second Appendix: Derivation of the electron-phonon interaction operator

By recalling Eq.(7.3), we here introduce the lattice-periodic part, $f_{\nu,\mathbf{q}}(\mathbf{r})$, of the perturbation potential energy due to phonons as

$$\sqrt{\frac{\hbar}{2\omega_{\nu,\mathbf{q}}}}\frac{\partial V_{KS}}{\partial U_{\nu,\mathbf{q}}} = f_{\nu,\mathbf{q}}(\mathbf{r},z)e^{i\mathbf{q}\cdot\mathbf{r}} \tag{B.1}$$

and then define the operator of the phonons perturbation potential energy as:

$$\hat{V}_{ep}(\mathbf{r},z) = \frac{1}{\sqrt{\Omega}}\sum_{\mathbf{q},\nu} f_{\nu,\mathbf{q}}(\mathbf{r},z)e^{i\mathbf{q}\cdot\mathbf{r}}(\hat{b}_{\mathbf{q},\nu} + \hat{b}^{\dagger}_{-\mathbf{q},\nu}) \tag{B.2}$$

where $\hat{b}_{\mathbf{q},\nu}$ and $\hat{b}^{\dagger}_{-\mathbf{q},\nu}$ are respectively the annihilation and creation operator for a phonon in branch ν and with wave-vector $\mathbf{q}$, and Ω is a normalization volume. Moreover, we recall also the definition for the field operators of electrons as

$$\begin{aligned}\hat{\Phi}(\mathbf{r},z) &= \sum_{i,c,k_y}\Phi_{i,c,k_y}\hat{a}_{i,c,k_y}\\ \hat{\Phi}^{\dagger}(\mathbf{r},z) &= \sum_{i,c,k_y}\Phi^{*}_{i,c,k_y}\hat{a}^{\dagger}_{i,c,k_y}\end{aligned} \tag{B.3}$$

where the basis functions Φ_{i,c,k_y} are defined in Eq.(7.2), $\hat{a}_{i,c,k_y}$ and $\hat{a}^\dagger_{i,c,k_y}$ are the corresponding electron annihilation and creation operators. The electron-phonon interaction in the URBFs basis set can thus be written as

$$\begin{aligned}\hat{H}_{ep} &= \int_\Omega \hat{\Phi}^\dagger(\mathbf{r},z)\,\hat{V}_{ep}(\mathbf{r},z)\hat{\Phi}(\mathbf{r},z)\,d\mathbf{r}\,dz \\ &= \sum_{\substack{i,c,k_y \\ j,c',k_y'}} \frac{1}{\sqrt{\Omega}} \sum_{\mathbf{q},\nu} M^\Phi_{\mathbf{q},\nu}(j,c',k_y',i,c,k_y)(\hat{b}_{\mathbf{q},\nu}+\hat{b}^\dagger_{-\mathbf{q},\nu})\hat{a}^\dagger_{j,c',k_y'}\hat{a}_{i,c,k_y}\end{aligned} \tag{B.4}$$

where we have introduced the matrix element

$$M^\Phi_{\mathbf{q},\nu}(j,c',k_y',i,c,k_y) = \int_\Omega \Phi^*_{j,c',k_y'} f_{\nu,\mathbf{q}}(\mathbf{r},z)\,e^{i\mathbf{q}\cdot\mathbf{r}}\Phi^*_{i,c,k_y}\,d\mathbf{r}\,dz \tag{B.5}$$

Because the basis functions are restricted to a single unit-cell along x, then $M^\Phi_{\mathbf{q},\nu}(j,c',k_y',i,c,k_y)$ is by definition zero for $c' \neq c$. We now assume $c' = c$, drop c in the notation, and substitute Eq.(7.2) into Eq.(B.5), which leads to

$$M^\Phi_{\mathbf{q},\nu}(j,c',k_y',i,c,k_y) = \sum_{\substack{n,k_x \\ m,k_x'}} \left[S^{-\frac{1}{2}}(k_y')\right]_{j,m,k_x'} M_{\mathbf{q},\nu}(m,\mathbf{k}_{x'y'},k_y',n,\mathbf{k}_{xy}) \left[S^{-\frac{1}{2}}(k_y)\right]_{n,i,k_x} \tag{B.6}$$

where we have used the fact that $S^{-\frac{1}{2}}(k_y)$ is Hermitian and, moreover, we have introduced the new matrix element

$$M_{\mathbf{q},\nu}(m,\mathbf{k}',n,\mathbf{k}) = \frac{1}{N_{cy}} \int_z \int_{L_y} \int_{ca_x}^{(c+1)a_x} u^*_{m\mathbf{k}'}(\mathbf{r},z) f_{\nu,\mathbf{q}}(\mathbf{r},z)\,e^{i(\mathbf{k}-\mathbf{k}'+\mathbf{q})\cdot\mathbf{r}} u_{n\mathbf{k}}(\mathbf{r},z)\,d\mathbf{r}\,dz \tag{B.7}$$

with L_y being the normalization length in the periodic direction y. The integral over L_y can be split into integrals over single unit cells by using $y = (R_l + \rho_y)$, where R_l is the y component of a direct lattice vector (with $l = 1, 2, \cdots N_{cy}$), and ρ_y varies in a single orthorhombic unit cell. By doing so, the matrix element $M_{\mathbf{q},\nu}(m,\mathbf{k}',n,\mathbf{k})$ can be rewritten as

$$\begin{aligned}M_{\mathbf{q},\nu}(m,\mathbf{k}',n,\mathbf{k}) =& \frac{1}{N_{cy}} \sum_{l=1}^{N_{cy}} e^{i(k_y-k_y'+q_y)R_y} \times \\ & \int_{\Omega,c} u^*_{m\mathbf{k}'}(x,\rho_y,z) f_{\nu,\mathbf{q}}(x,\rho_y,z)\,e^{i(k_x-k_x'+q_x)x}\,e^{i(k_y-k_y'+q_y)\rho_y}\,u_{n\mathbf{k}}(x,\rho_y,z)\,dx d\rho_y\,dz\end{aligned} \tag{B.8}$$

where the integral is taken over the volume, Ω, of any unit cell at the position c along x. Because k_y, k_y' and q_y are multiple of $2\pi/L_y$, the sum over l can be evaluated analytically as

$$\frac{1}{N_{cy}} \sum_{l=1}^{N_{cy}} e^{i(k_y-k_y'+q_y)R_y} = \delta_{k_y-k_y'+q_y} \tag{B.9}$$

In Eq.(B.9) and in the remainder of these derivations it is understood that the wave-vectors k and q are internal to the reduced zone, which may involve a reciprocal lattice vector $\mathbf{g}$ such that $\mathbf{k'} = \mathbf{k} + \mathbf{q} - \mathbf{g}$ is also internal to the reduced zone; the terms having $\mathbf{g} \neq 0$ are known as umklapp processes, and we hereafter omit g to simplify the notation. By using Eq.(B.9), $M_{\mathbf{q},\nu}(m, \mathbf{k'}, n, \mathbf{k})$ can be rewritten as

$$M_{\mathbf{q},\nu}(m, \mathbf{k'}, n, \mathbf{k}) = \delta_{k_y - k'_y + q_y} \int_{\Omega,c} u^*_{m\mathbf{k'}}(x,y,z) f_{\nu,\mathbf{q}}(x,y,z) e^{i(k_x - k'_x + q_x)x} u_{n\mathbf{k}}(x,y,z) dxdydz \tag{B.10}$$

where ρ_y has been renamed y. Eq.(B.10) deserves a few comments. Because the basis functions are restricted to a single unit cell along c, in the calculations of the $M_{\mathbf{q},\nu}(m, \mathbf{k'}, n, \mathbf{k})$ we do not find a $\delta_{\mathbf{k'},\mathbf{k}+\mathbf{q}}$ term, as we have for the periodic direction y. This implies that, for a given k_x and phonon wave-vector $\mathbf{q} = (q_x, q_y)$, Eq.(B.10) provides a non null $M_{\mathbf{q},\nu}(m, \mathbf{k'}, n, \mathbf{k})$ even for $k'_x = k_x + q_x$. In this regard, however, we now introduce an approximation and assume that, in the calculation of $M^{\Phi}_{\mathbf{q},\nu}(j, i, k_y)$ by means of Eq.(B.6), the dominant terms are those corresponding to $k'_x = k_x + q_x$; incidentally, the condition $k'_x = k_x + q_x$ also makes the integral in Eq.(B.10) independent of the position c along x. This approximation can be equivalently rephrased by stating that, to the purpose of evaluating $M^{\Phi}_{\mathbf{q},\nu}(j, i, k_y)$ in Eq.(B.6), the $M_{\mathbf{q},\nu}(m, \mathbf{k'}, n, \mathbf{k})$ can be approximated as

$$\begin{aligned} M_{\mathbf{q},\nu}(m, \mathbf{k'}, n, \mathbf{k}) &\approx g_{m,n,\nu}(\mathbf{k}, \mathbf{q}) \delta_{\mathbf{k'},\mathbf{k}+\mathbf{q}} \\ g_{m,n,\nu}(\mathbf{k}, \mathbf{q}) &= \int_{\Omega} u^*_{m\mathbf{k}+\mathbf{q}}(\mathbf{r}) f_{\nu,\mathbf{q}}(\mathbf{r}) u_{n\mathbf{k}}(\mathbf{r}) d\mathbf{r} \end{aligned} \tag{B.11}$$

where $g_{m,n,\nu}(\mathbf{k}, \mathbf{q})$ is the matrix element already defined in Eq.(7.3) and actually calculated by QE. By substituting Eq.(B.11) into Eq.(B.6), $M^{\Phi}_{\mathbf{q},\nu}(j, i, k_y)$ can be simplified as

$$M^{\Phi}_{\mathbf{q},\nu}(j, i, k_y) = \sum_{\substack{n,k_x \\ m,k'_x}} \left[S^{-\frac{1}{2}}(k_y + q_y) \right]_{j,m,k'_x} g_{m,n,\nu}(\mathbf{k}, \mathbf{q}) \left[S^{-\frac{1}{2}}(k_y) \right]_{n,ik_x} \delta_{k_x+q_x,k'_x} \tag{B.12}$$

The derivations from Eq.(B.5) to Eq.(B.12) allow us to state that our treatment of the phonon scattering is local to a unit cell, namely the electron-phonon scattering couples only URBFs restricted to the same unit cell, and that Eq.(B.4) can be recast in the form

$$\hat{H}_{ep} \approx \frac{1}{\sqrt{\Omega}} \sum_{j,ik_y} \sum_{\mathbf{q},\nu} M^{\Phi}_{\mathbf{q},\nu}(j, i, k_y) (\hat{b}_{\mathbf{q},\nu} + \hat{b}^{\dagger}_{-\mathbf{q},\nu}) \hat{a}^{\dagger}_{j,c',k'_y} \hat{a}_{i,c,k_y} \tag{B.13}$$

where $M^{\Phi}_{\mathbf{q},\nu}(j, ik_y)$ is given in Eq.(B.12). Hence Eq.(B.13) finally provides the derivation of Eq.3 in the main text.

Bibliography

[1] Techinsights. https://www.techinsights.com/products/sar-1504-801). Accessed: 2010-09-30.

[2] https://github.com/marcopala/Green-Tea-.\.

[3] *2021 IEEE International Roadmap for Devices and Systems Outbriefs*. IEEE Computer Society. ISBN 1-66548-639-2.

[4] MS Windows NT kernel description. http://web.archive.org/web/20080207010024/http://www.808multimedia.com/winnt/kernel.htm. Accessed: 2010-09-30.

[5] D. S. 1954-... *Quantum transport : atom to transistor*. Cambridge, UK New York : Cambridge University Press , cop. 2005, Cambridge, UK New York, 2005. ISBN 978-1-107-63213-4.

[6] A. Afzalian, E. Akhoundi, G. Gaddemane, R. Duflou, and M. Houssa. Advanced dft–negf transport techniques for novel 2-d material and device exploration including hfs_2/wse_2 van der waals heterojunction tfet and wte_2/ws_2 metal/semiconductor contact. *IEEE Transactions on Electron Devices*, 68(11):5372–5379, 2021. doi: 10.1109/TED.2021.3078412.

[7] M. P. Anantram, M. S. Lundstrom, and D. E. Nikonov. Modeling of nanoscale devices. *Proceedings of the IEEE*, 96(9):1511–1550, 2008. doi: 10.1109/JPROC.2008.927355.

[8] L. Ansari, S. Monaghan, N. McEvoy, C. O. Coileáin, C. P. Cullen, J. Lin, R. Siris, T. Stimpel-Lindner, K. F. Burke, G. Mirabelli, R. Duffy, E. Caruso, R. E. Nagle, G. S. Duesberg, P. K. Hurley, and F. Gity. Quantum confinement-induced semimetal-to-semiconductor evolution in large-area ultra-thin $ptse_2$ films grown at 400c. *NPJ 2D materials and applications*, 3(1), 2019. ISSN 2397-7132.

[9] J. Appenzeller, Y.-M. Lin, J. Knoch, and P. Avouris. Band-to-band tunneling in carbon nanotube field-effect transistors. *Physical review letters*, 93(19):196805.1–196805.4, 2004. ISSN 0031-9007.

[10] I. Asselberghs, Q. Smets, T. Schram, B. Groven, D. Verreck, A. Afzalian, G. Arutchelvan, A. Gaur, D. Cott, T. Maurice, S. Brems, K. Kennes, A. Phommahaxay, E. Dupuy, D. Radisic, J.-F. de Marneffe, A. Thiam,

W. Li, K. Devriendt, C. Huyghebaert, D. Lin, M. Caymax, P. Morin, and I. Radu. Wafer-scale integration of double gated ws2-transistors in 300mm si cmos fab. In *2020 IEEE International Electron Devices Meeting (IEDM)*, pages 40.2.1–40.2.4. IEEE, 2020. ISBN 9781728188881.

[11] M. M. Atalla. Semiconductor devices having dielectric coatings. (US3206670A), Mar. 8, 1960.

[12] U. E. Avci, D. H. Morris, and I. A. Young. Tunnel field-effect transistors: Prospects and challenges. *IEEE Journal of the Electron Devices Society*, 3(3):88–95, 2015. doi: 10.1109/JEDS.2015.2390591.

[13] G. Bae, D.-I. Bae, M. Kang, S. Hwang, S. Kim, B. Seo, T. Kwon, T. Lee, C. Moon, Y. Choi, K. Oikawa, S. Masuoka, K. Chun, S. Park, H. Shin, J. Kim, K. Bhuwalka, D. Kim, W. Kim, J. Yoo, H. Jeon, M. Yang, S.-J. Chung, D. Kim, B. Ham, K. Park, W. Kim, S. Park, G. Song, Y. Kim, M. Kang, K. Hwang, C.-H. Park, J.-H. Lee, D.-W. Kim, S.-M. Jung, and H. Kang. 3nm gaa technology featuring multi-bridge-channel fet for low power and high performance applications. In *2018 IEEE International Electron Devices Meeting (IEDM)*, pages 28.7.1–28.7.4. IEEE, 2018. ISBN 1728119871.

[14] E. Baravelli, E. Gnani, R. Grassi, A. Gnudi, S. Reggiani, and G. Baccarani. Optimization of n- and p-type tfets integrated on the same / technology platform. *IEEE transactions on electron devices*, 61(1):178–185, 2014. ISSN 0018-9383.

[15] S. Baroni, P. Giannozzi, and A. Testa. Green's-function approach to linear response in solids. *Phys. Rev. Lett.*, 58:1861–1864, May 1987. doi: 10.1103/PhysRevLett.58.1861. URL `https://link.aps.org/doi/10.1103/PhysRevLett.58.1861`.

[16] S. Baroni, S. de Gironcoli, A. Dal Corso, and P. Giannozzi. Phonons and related crystal properties from density-functional perturbation theory. *Rev. Mod. Phys.*, 73:515–562, Jul 2001. doi: 10.1103/RevModPhys.73.515. URL `https://link.aps.org/doi/10.1103/RevModPhys.73.515`.

[17] G. Bastard, J. Brum, and R. Ferreira. Electronic states in semiconductor heterostructures. In H. Ehrenreich and D. Turnbull, editors, *Semiconductor Heterostructures and Nanostructures*, volume 44 of *Solid State Physics*, pages 229–415. Academic Press, 1991. doi: https://doi.org/10.1016/S0081-1947(08)60092-2. URL `https://www.sciencedirect.com/science/article/pii/S0081194708600922`.

[18] P. Batude, M. Vinet, A. Pouydebasque, C. Le Royer, B. Previtali, C. Tabone, J.-M. Hartmann, L. Sanchez, L. Baud, V. Carron, A. Toffoli, F. Allain, V. Mazzocchi, D. Lafond, O. Thomas, O. Cueto, N. Bouzaida, D. Fleury, A. Amara, S. Deleonibus, and O. Faynot. Advances in 3d cmos sequential integration. In *2009 IEEE International Electron Devices Meeting (IEDM)*, pages 1–4. IEEE, 2009. ISBN 9781424456390.

[19] L. Bellaiche and D. Vanderbilt. Virtual crystal approximation revisited: Application to dielectric and piezoelectric properties of perovskites. *Phys. Rev. B*, 61:7877–7882, Mar 2000. doi: 10.1103/PhysRevB.61.7877. URL `https://link.aps.org/doi/10.1103/PhysRevB.61.7877`.

[20] P. B. Bennett, Z. Pedramrazi, A. Madani, Y.-C. Chen, D. G. de Oteyza, C. Chen, F. R. Fischer, M. F. Crommie, and J. Bokor. Bottom-up graphene nanoribbon field-effect transistors. *Applied physics letters*, 103(25):253114, 2013. ISSN 0003-6951.

[21] A. Betti, G. Fiori, and G. Iannaccone. Strong mobility degradation in ideal graphene nanoribbons due to phonon scattering. *Applied physics letters*, 98(21):212111, 2011. ISSN 0003-6951.

[22] M. Bohr. The evolution of scaling from the homogeneous era to the heterogeneous era. In *2011 International Electron Devices Meeting*, pages 1.1.1–1.1.6. IEEE, 2011. ISBN 1457705060.

[23] F. Bonaccorso, A. Lombardo, T. Hasan, Z. Sun, L. Colombo, and A. C. Ferrari. Production and processing of graphene and 2d crystals. *Materials today (Kidlington, England)*, 15(12):564–589, 2012. ISSN 1369-7021.

[24] B. M. Borg, K. A. Dick, B. Ganjipour, M.-E. Pistol, L. Wernersson, and C. Thelander. Inas/gasb heterostructure nanowires for tunnel field-effect transistors. *Nano letters*, 10 10:4080–5, 2010.

[25] S. Brocard, M. G. Pala, and D. Esseni. Large on-current enhancement in hetero-junction tunnel-fets via molar fraction grading. *IEEE Electron Device Letters*, 35(2):184–186, 2014. doi: 10.1109/LED.2013.2295884.

[26] J. Cao, D. Logoteta, S. ozkaya, B. Biel, A. Cresti, M. G. Pala, and D. Esseni. Operation and design of van der waals tunnel transistors: A 3-d quantum transport study. *IEEE Transactions on Electron Devices*, 63(11): 4388–4394, 2016. doi: 10.1109/TED.2016.2605144.

[27] J. Cao, J. Park, F. Triozon, M. Pala, and A. Cresti. Simulation of 2d material-based tunnel field-effect transistors: planar vs. vertical architectures. *Composants nanoélectroniques*, 18, 02 2018. doi: 10.21494/ISTE.OP.2018.0222.

[28] M. Cardona and F. H. Pollak. Energy-band structure of germanium and silicon: The k·p method. *Phys. Rev.*, 142:530–543, Feb 1966. doi: 10.1103/PhysRev.142.530. URL `https://link.aps.org/doi/10.1103/PhysRev.142.530`.

[29] A. M. Carlin. Growth and strain relaxation in anion-graded gaxin1-xasyp1-y. 2010.

[30] D. M. Ceperley and B. J. Alder. Ground state of the electron gas by a stochastic method. *Phys. Rev. Lett.*, 45:566–569, Aug 1980. doi: 10.1103/PhysRevLett.45.566.

[31] D. Chakraborty, K. Berland, and T. Thonhauser. Next-generation nonlocal van der waals density functional. *Journal of Chemical Theory and Computation*, 16(9):5893–5911, 2020. doi: 10.1021/acs.jctc.0c00471. URL https://doi.org/10.1021/acs.jctc.0c00471. PMID: 32786912.

[32] J. R. Chelikowsky and M. L. Cohen. Electronic structure of silicon. *Phys. Rev. B*, 10:5095–5107, Dec 1974. doi: 10.1103/PhysRevB.10.5095. URL https://link.aps.org/doi/10.1103/PhysRevB.10.5095.

[33] J. R. Chelikowsky and M. L. Cohen. Electronic structure of silicon. *Physical review. B, Solid state*, 10(12): 5095–5107, 1974. ISSN 0556-2805.

[34] J. R. Chelikowsky and M. L. Cohen. Nonlocal pseudopotential calculations for the electronic structure of eleven diamond and zinc-blende semiconductors. *Phys. Rev. B*, 14:556–582, Jul 1976. doi: 10.1103/PhysRevB.14.556.

[35] J. R. Chelikowsky, N. Troullier, and Y. Saad. Finite-difference-pseudopotential method: Electronic structure calculations without a basis. *Phys. Rev. Lett.*, 72:1240–1243, Feb 1994. doi: 10.1103/PhysRevLett.72.1240. URL https://link.aps.org/doi/10.1103/PhysRevLett.72.1240.

[36] K. Chen, X. Wan, J. Wen, W. Xie, Z. Kang, X. Zeng, H. Chen, and J.-B. Xu. Electronic properties of mos2-ws2 heterostructures synthesized with two-step lateral epitaxial strategy. *ACS nano*, 9(10):9868–9876, 2015. ISSN 1936-086X.

[37] P. Chen, N. Li, X. Chen, W.-J. Ong, and X. Zhao. The rising star of 2d black phosphorus beyond graphene: synthesis, properties and electronic applications. *2d materials*, 5(1):14002, 2017. ISSN 2053-1583.

[38] P.-F. Chen, E. Chen, and Y.-R. Wu. Design of monolayer mos2 nanosheet transistors for low-power applications. *IEEE Transactions on Electron Devices*, 69(1):358–363, 2022. doi: 10.1109/TED.2021.3130840.

[39] J. Choukroun, M. Pala, S. Fang, E. Kaxiras, and P. Dollfus. High performance tunnel field effect transistors based on in-plane transition metal dichalcogenide heterojunctions. *Nanotechnology*, 30(2):025201–025201, 2018. ISSN 0957-4484.

[40] A. Ciarrocchi, A. Avsar, D. Ovchinnikov, and A. Kis. Thickness-modulated metal-to-semiconductor transformation in a transition metal dichalcogenide. *Nature communications*, 9(1):919–6, 2018. ISSN 2041-1723.

[41] M. L. Cohen. *Electronic Structure and Optical Properties of Semiconductors*. Springer Series in Solid-State Sciences, 75. 2nd ed. 1989.. edition, 1989. ISBN 3-540-51391-4.

[42] M. L. Cohen and T. K. Bergstresser. Band structures and pseudopotential form factors for fourteen semiconductors of the diamond and zinc-blende structures. *Phys. Rev.*, 141:789–796, Jan 1966. doi: 10.1103/PhysRev.141.789.

[43] M. L. Cohen and V. Heine. The fitting of pseudopotentials to experimental data and their subsequent application. volume 24 of *Solid State Physics*, pages 37–248. Academic Press, 1970. doi: https://doi.org/10.1016/S0081-1947(08)60070-3. URL https://www.sciencedirect.com/science/article/pii/S0081194708600703.

[44] J.-P. Colinge. Multiple-gate soi mosfets. *Solid-state electronics*, 48(6):897–905, 2004. ISSN 0038-1101.

[45] F. Conzatti, M. G. Pala, and D. Esseni. Surface-roughness-induced variability in nanowire inas tunnel fets. *IEEE Electron Device Letters*, 33(6):806–808, 2012. doi: 10.1109/LED.2012.2192091.

[46] F. Conzatti, M. G. Pala, D. Esseni, E. Bano, and L. Selmi. Strain-induced performance improvements in inas nanowire tunnel fets. *IEEE Transactions on Electron Devices*, 59:2085–2092, 2012.

[47] F. Conzatti, M. G. Pala, D. Esseni, E. Bano, and L. Selmi. Strain-induced performance improvements in inas nanowire tunnel fets. *IEEE transactions on electron devices*, 59(8):2085–2092, 2012. ISSN 0018-9383.

[48] A. C. Cornescu, R. Morariu, A. Ofiare, A. Al-Khalidi, J. Wang, J. M. L. Figueiredo, and E. Wasige. High-efficiency bias stabilization for resonant tunneling diode oscillators. *IEEE transactions on microwave theory and techniques*, 67(8):3449–3454, 2019. ISSN 0018-9480.

[49] S. B. Desai, S. R. Madhvapathy, A. B. Sachid, J. P. Llinas, Q. Wang, G. H. Ahn, G. Pitner, M. J. Kim, J. Bokor, C. Hu, H.-S. P. Wong, and A. Javey. Mos_2 transistors with 1-nanometer gate lengths. *Science*, 354(6308):99–102, 2016. doi: 10.1126/science.aah4698. URL https://www.science.org/doi/abs/10.1126/science.aah4698.

[50] D. Domaretskiy, M. Philippi, M. Gibertini, N. Ubrig, I. Gutiérrez-Lezama, and A. F. Morpurgo. Quenching the bandgap of two-dimensional semiconductors with a perpendicular electric field. *Nature Nanotechnology*, aug 2022. doi: 10.1038/s41565-022-01183-4.

[51] R. Duflou, M. Houssa, and A. Afzalian. Electron-phonon scattering in cold-metal contacted two-dimensional semiconductor devices. In *2021 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*, pages 94–97. IEEE, 2021. ISBN 9781665406857.

[52] T. Eknapakul, I. Fongkaew, S. Siriroj, W. Jindata, S. Chaiyachad, S. Mo, S. Thakur, L. Petaccia, H. Takagi, S. Limpijumnong, and W. Meevasana. Direct observation of strain-induced orbital valence band splitting in hfse2 by sodium intercalation. 2018.

[53] J. K. Ellis, M. J. Lucero, and G. E. Scuseria. The indirect to direct band gap transition in multilayered mos 2 as predicted by screened hybrid density functional theory. *Applied physics letters*, 99(26):261908–261908–3, 2011. ISSN 0003-6951.

[54] M. Ernzerhof and G. E. Scuseria. Assessment of the perdew–burke–ernzerhof exchange-correlation functional. *The Journal of Chemical Physics*, 110(11):5029–5036, 1999. doi: 10.1063/1.478401. URL https://doi.org/10.1063/1.478401.

[55] D. Esseni, P. Palestri, and L. Selmi. Nanoscale mos transistors: Semi-classical transport and applications. 2011.

[56] D. Esseni, M. Pala, P. Palestri, C. Alper, and T. Rollo. A review of selected topics in physics based modeling for tunnel field-effect transistors. *Semiconductor Science and Technology*, 32(8):083005, jul 2017. doi: 10.1088/1361-6641/aa6fca. URL https://dx.doi.org/10.1088/1361-6641/aa6fca.

[57] R. P. Feynman. Forces in molecules. *Physical review*, 56(4):340, 1939.

[58] M. V. Fischetti and W. G. Vandenberghe. Mermin-wagner theorem, flexural modes, and degraded carrier mobility in two-dimensional crystals with broken horizontal mirror symmetry. *Phys. Rev. B*, 93:155413, Apr 2016. doi: 10.1103/PhysRevB.93.155413. URL https://link.aps.org/doi/10.1103/PhysRevB.93.155413.

[59] P. Friedel, M. S. Hybertsen, and M. Schlüter. Local empirical pseudopotential approach to the optical properties of si/ge superlattices. *Phys. Rev. B*, 39:7974–7977, Apr 1989. doi: 10.1103/PhysRevB.39.7974. URL https://link.aps.org/doi/10.1103/PhysRevB.39.7974.

[60] R. Frisenda, E. Navarro-Moratalla, P. Gant, D. Pérez De Lara, P. Jarillo-Herrero, R. V. Gorbachev, and A. Castellanos-Gomez. Recent progress in the assembly of nanodevices and van der waals heterostructures by deterministic placement of 2d materials. *Chemical Society reviews*, 47(1):53–68, 2018. ISSN 0306-0012.

[61] S. Fujita. Partial self-energy parts of kadanoff-baym. *Physica*, 30(4):848–856, 1964. ISSN 0031-8914. doi: https://doi.org/10.1016/0031-8914(64)90127-2.

[62] G. Gaddemane, S. Gopalan, M. L. Van de Put, and M. V. Fischetti. Limitations of ab initio methods to predict the electronic-transport properties of two-dimensional semiconductors: the computational example of 2h-phase transition metal dichalcogenides. *Journal of computational electronics*, 20(1):49–59, 2021. ISSN 1569-8025.

[63] B. Ganjipour, J. Wallentin, M. T. Borgstrom, L. Samuelson, and C. Thelander. Tunnel field-effect transistors based on inp-gaas heterostructure nanowires. *ACS nano*, 6(4):3109–3113, 2012. ISSN 1936-0851.

[64] M. Gell-Mann and F. Low. Bound states in quantum field theory. *Phys. Rev.*, 84:350–354, Oct 1951. doi: 10.1103/PhysRev.84.350.

[65] P. Giannozzi, O. Andreussi, T. Brumme, O. Bunau, M. B. Nardelli, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, M. Cococcioni, N. Colonna, I. Carnimeo, A. D. Corso, S. de Gironcoli, P. Delugas, R. A. DiStasio, A. Ferretti, A. Floris, G. Fratesi, G. Fugallo, R. Gebauer, U. Gerstmann, F. Giustino, T. Gorni, J. Jia,

M. Kawamura, H.-Y. Ko, A. Kokalj, E. Kuçukbenli, M. Lazzeri, M. Marsili, N. Marzari, F. Mauri, N. L. Nguyen, H.-V. Nguyen, A. O. de-la Roza, L. Paulatto, S. Poncé, D. Rocca, R. Sabatini, B. Santra, M. Schlipf, A. P. Seitsonen, A. Smogunov, I. Timrov, T. Thonhauser, P. Umari, N. Vast, X. Wu, and S. Baroni. Advanced capabilities for materials modelling with quantum ESPRESSO. *Journal of Physics: Condensed Matter*, 29(46): 465901, oct 2017. doi: 10.1088/1361-648x/aa8f79. URL `https://doi.org/10.1088/1361-648x/aa8f79`.

[66] F. Giustino. Electron-phonon interactions from first principles. *Rev. Mod. Phys.*, 89:015003, Feb 2017. doi: 10.1103/RevModPhys.89.015003. URL `https://link.aps.org/doi/10.1103/RevModPhys.89.015003`.

[67] Y. Gong, J. Lin, X. Wang, G. Shi, S. Lei, Z. Lin, X. Zou, G. Ye, R. Vajtai, B. I. Yakobson, H. Terrones, M. Terrones, B. K. Tay, J. Lou, S. T. Pantelides, Z. Liu, W. Zhou, and P. M. Ajayan. Vertical and in-plane heterostructures from ws2/mos2 monolayers. *Nature materials*, 13(12):1135–1142, 2014. ISSN 1476-1122.

[68] Y. Gong, H. Yuan, C.-L. Wu, P. Tang, S.-Z. Yang, A. Yang, G. Li, B. Liu, J. van de Groep, M. L. Brongersma, M. F. Chisholm, S.-C. Zhang, W. Zhou, and Y. Cui. Spatially controlled doping of two-dimensional sns2 through intercalation for electronics. *Nature nanotechnology*, 13(4):294–299, 2018. ISSN 1748-3387.

[69] Y. Gong, H. Yuan, C.-L. Wu, P. Tang, S.-Z. Yang, A. Yang, G. Li, B. Liu, J. van de Groep, M. L. Brongersma, M. F. Chisholm, S.-C. Zhang, W. Zhou, and Y. Cui. Spatially controlled doping of two-dimensional sns 2 through intercalation for electronics. *Nature nanotechnology*, 13(4):294, 2018. ISSN 1748-3395.

[70] S. Grimme. Semiempirical gga-type density functional constructed with a long-range dispersion correction. *Journal of computational chemistry*, 27(15):1787–1799, 2006. ISSN 0192-8651.

[71] T. Gunst, T. Markussen, K. Stokbro, and M. Brandbyge. First-principles method for electron-phonon coupling and electron mobility: Applications to two-dimensional materials. *Phys. Rev. B*, 93:035414, Jan 2016. doi: 10.1103/PhysRevB.93.035414. URL `https://link.aps.org/doi/10.1103/PhysRevB.93.035414`.

[72] P. Guo, Y. Yang, Y. Cheng, G. Han, J. Pan, Ivana, Z. Zhang, H. Hu, Z. Shen, C. K. Chia, and Y.-C. Yeo. Tunneling field-effect transistor with ge/in0.53ga0.47as heterostructure as tunneling junction. *Journal of Applied Physics*, 113:094502, 2013.

[73] Y. Guo and J. Robertson. Band engineering in transition metal dichalcogenides: Stacked versus lateral heterostructures. *Applied physics letters*, 108(23):233104, 2016. ISSN 0003-6951.

[74] I. Hamada. van der waals density functional made accurate. *Physical review. B, Condensed matter and materials physics*, 89(12), 2014. ISSN 1098-0121.

[75] J. Hamel, Y. Tang, and K. Osman. Technological requirements for a lateral sige hbt technology including theoretical performance predictions relative to vertical sige hbts. *IEEE Transactions on Electron Devices*, 49 (3):449–456, 2002. doi: 10.1109/16.987116.

[76] H. Haug and A.-P. Jauho. *Quantum Kinetics in Transport and Optics of Semiconductors*, volume 123 of *Springer Series in Solid-State Sciences*. Springer-Verlag, Berlin, Heidelberg, 2. aufl. edition, 2008. ISBN 9783540735618.

[77] H. Haug and A.-P. Jauho. *Quantum kinetics in transport and optics of semiconductors*, volume 2. Springer, 2008.

[78] H. Hellmann et al. Einfuhrung in die quantenchemie. 1937.

[79] L. A. Hemstreet, C. Y. Fong, and M. L. Cohen. Calculation of the band structure and optical constants of diamond using the nonlocal-pseudopotential method. *Phys. Rev. B*, 2:2054–2063, Sep 1970. doi: 10.1103/PhysRevB.2.2054.

[80] L. A. Hemstreet, C. Y. Fong, and M. L. Cohen. Calculation of the band structure and optical constants of diamond using the nonlocal-pseudopotential method. *Physical review. B, Solid state*, 2(6):2054–2063, 1970. ISSN 0556-2805.

[81] J. Heyd, G. E. Scuseria, and M. Ernzerhof. Hybrid functionals based on a screened coulomb potential. *The Journal of chemical physics*, 118(18):8207–8215, 2003. ISSN 0021-9606.

[82] S. Hofstein and G. Warfield. The insulated gate tunnel junction triode. *IEEE transactions on electron devices*, 12(2):66–76, 1965. ISSN 0018-9383.

[83] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Phys. Rev.*, 136:B864–B871, Nov 1964. doi: 10.1103/PhysRev.136.B864.

[84] C. Y. Huang, G. Dewey, E. Mannebach, A. Phan, P. Morrow, W. Rachmady, I. C. Tung, N. Thomas, U. Alaan, R. Paul, N. Kabir, B. Krist, A. Oni, M. Mehta, M. Harper, P. Nguyen, R. Keech, S. Vishwanath, K. L. Cheong, J. S. Kang, A. Lilak, M. Metz, S. Clendenning, B. Turkot, R. Schenker, H. J. Yoo, M. Radosavljevic, and J. Kavalieros. 3-d self-aligned stacked nmos-on-pmos nanoribbon transistors for continued moore's law scaling. In *2020 IEEE International Electron Devices Meeting (IEDM)*, pages 20.6.1–20.6.4. IEEE, 2020. ISBN 9781728188881.

[85] V. Hung Nguyen, J. Saint-Martin, D. Querlioz, F. Mazzamuto, A. Bournel, Y.-M. Niquet, and P. Dollfus. Bandgap nanoengineering of graphene tunnel diodes and tunnel transistors to control the negative differential resistance. *Journal of computational electronics*, 12(2):85–93, 2013. ISSN 1569-8025.

[86] G. Iannaccone, F. Bonaccorso, L. Colombo, and G. Fiori. Quantum engineering of transistors based on 2d materials heterostructures. *Nature Nanotechnology*, 13(3):183–191, Mar 2018. ISSN 1748-3395. doi: 10.1038/s41565-018-0082-6. URL `http://dx.doi.org/10.1038/s41565-018-0082-6`.

[87] J. O. Island, G. A. Steele, H. S. J. v. d. Zant, and A. Castellanos-Gomez. Environmental instability of few-layer black phosphorus. *2d materials*, 2(1):11002–6, 2015. ISSN 2053-1583.

[88] C. Jacoboni and P. Lugli. *Monte Carlo Method for Semiconductor Device Simulation*. Computational Microelectronics. Springer, Vienna, 1989. ISBN 9783709174531.

[89] J.-M. Jancu, R. Scholz, F. Beltram, and F. Bassani. Empirical spds* tight-binding calculation for cubic semiconductors: General method and material parameters. *Phys. Rev. B*, 57:6493–6507, Mar 1998. doi: 10.1103/PhysRevB.57.6493.

[90] C. Jungemann. *Hierarchical Device Simulation The Monte-Carlo Perspective*. Computational Microelectronics. 1st ed. 2003. edition, 2003. ISBN 3-211-01361-X.

[91] 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*, 85:115317, Mar 2012. doi: 10.1103/PhysRevB.85.115317.

[92] L. V. Keldysh. Diagram technique for nonequilibrium processes. *Sov. Phys. JETP*, 47:1515–1527, 1964.

[93] J. Y. Kim, H. J. Park, S.-h. Lee, C. Seo, J. Kim, and J. Joo. Distinctive field-effect transistors and ternary inverters using cross-type wse2/mos2 heterojunctions treated with polymer acid. *ACS Applied Materials & Interfaces*, 12(32):36530–36539, 2020. doi: 10.1021/acsami.0c09706. PMID: 32672032.

[94] G. Klimeck, S. Ahmed, H. Bae, N. Kharche, S. Clark, B. Haley, S. Lee, M. Naumov, H. Ryu, F. Saied, M. Prada, M. Korkusinski, T. Boykin, and R. Rahman. Atomistic simulation of realistically sized nanodevices using nemo 3-d-part i: Models and benchmarks. *IEEE transactions on electron devices*, 54(9):2079–2089, 2007. ISSN 0018-9383.

[95] C. Klinkert, A. Szabo, C. Stieger, D. Campi, N. Marzari, and M. Luisier. 2-d materials for ultrascaled field-effect transistors: One hundred candidates under the ab initio microscope. *ACS nano*, 14(7):8605–8615, 2020. ISSN 1936-0851.

[96] C. Klinkert, S. Fiore, J. Backman, Y. Lee, and M. Luisier. Impact of orientation misalignments on black phosphorus ultrascaled field-effect transistors. *IEEE Electron Device Letters*, 42:434–437, 2021.

[97] W. Kohn and L. J. Sham. Self-consistent equations including exchange and correlation effects. *Phys. Rev.*, 140: A1133–A1138, Nov 1965. doi: 10.1103/PhysRev.140.A1133.

[98] M. Kurpas and J. Fabian. Intrinsic and extrinsic spin-orbit coupling and spin relaxation in monolayer $ptse_2$. *Phys. Rev. B*, 103:125409, Mar 2021. doi: 10.1103/PhysRevB.103.125409. URL `https://link.aps.org/doi/10.1103/PhysRevB.103.125409`.

[99] K.-T. Lam, Z. Dong, and J. Guo. Performance limits projection of black phosphorous field-effect transistors. *IEEE Electron Device Letters*, 35(9):963–965, 2014. doi: 10.1109/LED.2014.2333368.

[100] D. C. Langreth. Linear and nonlinear response theory with applications. In *Linear and Nonlinear Electron Transport in Solids*, NATO Advanced Study Institutes Series, pages 3–32. Springer US, Boston, MA, 1976. ISBN 1475708777.

[101] A. Laturia, M. L. Van de Put, and W. G. Vandenberghe. Dielectric properties of hexagonal boron nitride and transition metal dichalcogenides: from monolayer to bulk. *NPJ 2D materials and applications*, 2(1), 2018. ISSN 2397-7132.

[102] C. Lee, I. Meric, S. Sorgenfrei, K. L. Shepard, C. R. Dean, A. F. Young, L. Wang, J. Hone, P. Kim, T. Taniguchi, and K. Watanabe. Boron nitride substrates for high-quality graphene electronics. *Nature nanotechnology*, 5 (10):722–726, 2010. ISSN 1748-3387.

[103] W. Lee, Y. Lin, L.-S. Lu, W.-C. Chueh, M. Liu, X. Li, W.-H. Chang, R. A. Kaindl, and C.-K. Shih. Time-resolved arpes determination of a quasi-particle band gap and hot electron dynamics in monolayer mos2. *Nano Letters*, 21(17):7363–7370, 2021. doi: 10.1021/acs.nanolett.1c02674. PMID: 34424691.

[104] S. Lei, X. Wang, B. Li, J. Kang, Y. He, A. George, L. Ge, Y. Gong, P. Dong, Z. Jin, G. Brunetto, W. Chen, Z.-T. Lin, R. Baines, D. S. Galvão, J. Lou, E. Barrera, K. Banerjee, R. Vajtai, and P. Ajayan. Surface functionalization of two-dimensional metal chalcogenides by lewis acid-base chemistry. *Nature nanotechnology*, 11(5):465–471, 2016. ISSN 1748-3387.

[105] H. Li, P. Li, J.-K. Huang, M.-Y. Li, C.-W. Yang, Y. Shi, X.-X. Zhang, and L.-J. Li. Laterally stitched heterostructures of transition metal dichalcogenide: Chemical vapor deposition growth on lithographically patterned area. *ACS Nano*, 10(11):10516–10523, 2016. doi: 10.1021/acsnano.6b06496. PMID: 27797472.

[106] M.-Y. Li, Y. Shi, C.-C. Cheng, L.-S. Lu, Y.-C. Lin, H.-L. Tang, M.-L. Tsai, C.-W. Chu, K.-H. Wei, J.-H. He, W.-H. Chang, K. Suenaga, and L.-J. Li. Epitaxial growth of a monolayer wse₂-mos₂ lateral p-n junction with an atomically sharp interface. *Science*, 349(6247):524–528, 2015. doi: 10.1126/science.aab4097.

[107] W. Li, J. Zhou, S. Cai, Z. Yu, J. Zhang, N. Fang, T. Li, Y. Wu, T. Chen, X. Xie, H. Ma, K. Yan, N. Dai, X. Wu, H. Zhao, Z. Wang, D. He, L. J. Pan, Y. Shi, and X. Wang. Uniform and ultrathin high-κ gate dielectrics for two-dimensional electronic devices. *Nature Electronics*, 2:1–9, 12 2019. doi: 10.1038/s41928-019-0334-y.

[108] X. Li, J. T. Mullen, Z. Jin, K. M. Borysenko, M. Buongiorno Nardelli, and K. W. Kim. Intrinsic electrical transport properties of monolayer silicene and mos_2 from first principles. *Phys. Rev. B*, 87:115418, Mar 2013. doi: 10.1103/PhysRevB.87.115418. URL `https://link.aps.org/doi/10.1103/PhysRevB.87.115418`.

[109] Y.-C. Lin, R. K. Ghosh, R. Addou, N. Lu, S. M. Eichfeld, H. Zhu, M.-Y. Li, X. Peng, M. J. Kim, L.-J. Li, R. M. Wallace, S. Datta, and J. A. Robinson. Atomically thin resonant tunnel diodes built from synthetic van der waals heterostructures. *Nature communications*, 6(1):7311–7311, 2015. ISSN 2041-1723.

[110] X. Ling, Y. Lin, Q. Ma, Z. Wang, Y. Song, L. Yu, S. Huang, W. Fang, X. Zhang, A. L. Hsu, Y. Bie, Y.-H. Lee, Y. Zhu, L. Wu, J. Li, P. Jarillo-Herrero, M. Dresselhaus, T. Palacios, and J. Kong. Parallel stitching of 2d materials. *Advanced Materials*, 28(12):2322–2329, 2016. doi: https://doi.org/10.1002/adma.201505070. URL `https://onlinelibrary.wiley.com/doi/abs/10.1002/adma.201505070`.

[111] F. Liu. Switching at less than 60 mv/decade with a "cold" metal as the injection source. *Physical review applied*, 13(6), 2020. ISSN 2331-7019.

[112] F. Liu, Y. Wang, X. Liu, J. Wang, and H. Guo. Ballistic transport in monolayer black phosphorus transistors. *IEEE Transactions on Electron Devices*, 61(11):3871–3876, 2014. doi: 10.1109/TED.2014.2353213.

[113] H. Liu, Y. Du, Y. Deng, and P. D. Ye. Semiconducting black phosphorus: synthesis, transport properties and electronic applications. *Chem. Soc. Rev.*, 44:2732–2743, 2015. doi: 10.1039/C4CS00257A. URL `http://dx.doi.org/10.1039/C4CS00257A`.

[114] D. Lizzit, D. Esseni, P. Palestri, P. Osgnach, and L. Selmi. Performance benchmarking and effective channel length for nanoscale inas, $in_{0.53}ga_{0.47}as$, and ssi n-mosfets. *IEEE Transactions on Electron Devices*, 61(6): 2027–2034, 2014. doi: 10.1109/TED.2014.2315919.

[115] D. Logoteta, J. Cao, M. Pala, P. Dollfus, Y. Lee, and G. Iannaccone. Cold-source paradigm for steep-slope transistors based on van der waals heterojunctions. *Physical review research*, 2(4), 2020. ISSN 2643-1564.

[116] P.-O. Löwdin and H. Shull. Natural orbitals in the quantum theory of two-electron systems. *Phys. Rev.*, 101: 1730–1739, Mar 1956. doi: 10.1103/PhysRev.101.1730. URL `https://link.aps.org/doi/10.1103/PhysRev.101.1730`.

[117] M. Luisier and G. Klimeck. Omen an atomistic and full-band quantum transport simulator for post-cmos nanodevices. In *2008 8th IEEE Conference on Nanotechnology*, pages 354–357, 2008. doi: 10.1109/NANO.2008.110.

[118] M. Luisier, A. Schenk, W. Fichtner, and G. Klimeck. Atomistic simulation of nanowires in the s p 3 d 5 s tight-binding formalism: From boundary conditions to strain calculations. *Physical review. B, Condensed matter and materials physics*, 74(20), 2006. ISSN 1098-0121.

[119] G. D. Mahan. *Many-Particle Physics*. Physics of Solids and Liquids. Springer, Boston, MA, third edition. edition, 2000. ISBN 1441933395.

[120] K. F. Mak, C. Lee, J. Hone, J. Shan, and T. F. Heinz. Atomically thin mos_2: a new direct-gap semiconductor. *Physical review letters*, 105(13):136805–136805, 2010. ISSN 1079-7114.

[121] E. G. Marin, D. Marian, M. Perucchini, G. Fiori, and G. Iannaccone. Lateral heterostructure field-effect transistors based on two-dimensional material stacks with varying thickness and energy filtering source. *ACS nano*, 14(2):1982–1989, 2020. ISSN 1936-0851.

[122] A. S. Mayorov, R. V. Gorbachev, S. V. Morozov, L. Britnell, R. Jalil, L. A. Ponomarenko, P. Blake, K. S. Novoselov, K. Watanabe, T. Taniguchi, and A. K. Geim. Micrometer-scale ballistic transport in encapsulated graphene at room temperature. *Nano letters*, 11(6):2396–2399, 2011. ISSN 1530-6984.

[123] S. R. Mehrotra, S. Kim, T. Kubis, M. Povolotskyi, M. S. Lundstrom, and G. Klimeck. Engineering nanowire n-mosfets at $l_g < 8$ nm. *IEEE Transactions on Electron Devices*, 60(7):2171–2177, 2013. doi: 10.1109/TED.2013.2263806.

[124] E. Memisevic, J. Svensson, M. Hellenbrand, E. Lind, and L.-E. Wernersson. Scaling of vertical inas-gasb nanowire tunneling field-effect transistors on si. *IEEE electron device letters*, 37(5):549–552, 2016. ISSN 0741-3106.

[125] E. Memisevic, J. Svensson, M. Hellenbrand, E. Lind, and L.-E. Wernersson. Vertical inas/gaassb/gasb tunneling field-effect transistor on si with s = 48 mv/decade and ion = 10 μa/μm for ioff = 1 na/μm at vds = 0.3 v. In *2016 IEEE International Electron Devices Meeting (IEDM)*, pages 19.1.1–19.1.4. IEEE, 2016. ISBN 9781509039029.

[126] F. Michelini, N. Cavassilas, R. Hayn, and M. Szczap. Multiband $k \cdot p$ models for strained zincblende crystals: Application to the fine structure of zno. *Phys. Rev. B*, 80:245210, Dec 2009. doi: 10.1103/PhysRevB.80.245210. URL `https://link.aps.org/doi/10.1103/PhysRevB.80.245210`.

[127] G. Mil'nikov, N. Mori, and Y. Kamakura. Equivalent transport models in atomistic quantum wires. *Phys. Rev. B*, 85:035317, Jan 2012. doi: 10.1103/PhysRevB.85.035317.

[128] 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*, 13(3):246–252, 2018. ISSN 1748-3387.

[129] M. Moussavou, M. Lannoo, N. Cavassilas, D. Logoteta, and M. Bescond. Physically based diagonal treatment of the self-energy of polar optical phonons: Performance assessment of iii-v double-gate transistors. *Phys. Rev. Appl.*, 10:064023, Dec 2018. doi: 10.1103/PhysRevApplied.10.064023. URL `https://link.aps.org/doi/10.1103/PhysRevApplied.10.064023`.

[130] M. Nakano, Y. Wang, Y. Kashiwabara, H. Matsuoka, and Y. Iwasa. Layer-by-layer epitaxial growth of scalable wse2 on sapphire by molecular beam epitaxy. *Nano Letters*, 17(9):5595–5599, 2017. doi: 10.1021/acs.nanolett.7b02420. URL https://doi.org/10.1021/acs.nanolett.7b02420. PMID: 28849935.

[131] 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*, 306(5696):666–669, 2004. doi: 10.1126/science.1102896. URL https://www.science.org/doi/abs/10.1126/science.1102896.

[132] K. S. Novoselov, D. Jiang, F. Schedin, T. J. Booth, V. V. Khotkevich, S. V. Morozov, and A. K. Geim. Two-dimensional atomic crystals. *Proceedings of the National Academy of Sciences - PNAS*, 102(30):10451–10453, 2005. ISSN 0027-8424.

[133] K. L. P. *Quantum statistical mechanics : Green's function methods in equilibrium and nonequilibrium problems.* Advanced book classics. Redwood City Calif. : Addison-Wesley Pub. Co., Advanced Book Program , cop. 1989, Redwood City (Calif.), 1989. ISBN 0-201-09422-3.

[134] M. G. Pala and D. Esseni. Interface traps in inas nanowire tunnel-fets and mosfets-part i: Model description and single trap analysis in tunnel-fets. *IEEE transactions on electron devices*, 60(9):2795–2801, 2013. ISSN 0018-9383.

[135] M. G. Pala and D. Esseni. Quantum transport models based on negf and empirical pseudopotentials for accurate modeling of nanoscale electron devices. *Journal of applied physics*, 126(5):55703, 2019. ISSN 0021-8979.

[136] M. G. Pala and D. Esseni. Quantum transport models based on negf and empirical pseudopotentials for accurate modeling of nanoscale electron devices. *Journal of applied physics*, 126(5):55703, 2019. ISSN 0021-8979.

[137] M. G. Pala, P. Giannozzi, and D. Esseni. Unit cell restricted bloch functions basis for first-principle transport models: Theory and application. *Phys. Rev. B*, 102:045410, Jul 2020. doi: 10.1103/PhysRevB.102.045410.

[138] K. C. Pandey and J. C. Phillips. Nonlocal pseudopotentials for ge and gaas. *Phys. Rev. B*, 9:1552–1559, Feb 1974. doi: 10.1103/PhysRevB.9.1552.

[139] D. J. Paul. 8-band **k**.**p** modeling of the quantum confined stark effect in ge quantum wells on si substrates. *Phys. Rev. B*, 77:155323, Apr 2008. doi: 10.1103/PhysRevB.77.155323. URL https://link.aps.org/doi/10.1103/PhysRevB.77.155323.

[140] J. P. Perdew, K. Burke, and M. Ernzerhof. Generalized gradient approximation made simple. *Phys. Rev. Lett.*, 77:3865–3868, Oct 1996. doi: 10.1103/PhysRevLett.77.3865.

[141] J. C. Phillips and L. Kleinman. New method for calculating wave functions in crystals and molecules. *Phys. Rev.*, 116:287–294, Oct 1959. doi: 10.1103/PhysRev.116.287.

[142] A. Pilotto, P. Dollfus, J. Saint-Martin, and M. Pala. Full quantum simulation of shockley–read–hall recombination in p-i-n and tunnel diodes. *Solid-State Electronics*, 198:108469, 2022. ISSN 0038-1101. doi: https://doi.org/10.1016/j.sse.2022.108469. URL https://www.sciencedirect.com/science/article/pii/S0038110122002404.

[143] A. Pilotto, P. Khakbaz, P. Palestri, and D. Esseni. Semi-classical transport in mos2 and mos2 transistors by a monte carlo approach. *Solid-State Electronics*, 192:108295, 2022. ISSN 0038-1101. doi: https://doi.org/10.1016/j.sse.2022.108295. URL https://www.sciencedirect.com/science/article/pii/S0038110122000673.

[144] S. Poncé, W. Li, S. Reichardt, and F. Giustino. First-principles calculations of charge carrier mobility and conductivity in bulk semiconductors and two-dimensional materials. *Reports on Progress in Physics*, 83, 2020.

[145] S. Poncé, E. Margine, C. Verdi, and F. Giustino. Epw: Electron–phonon coupling, transport and superconducting properties using maximally localized wannier functions. *Computer Physics Communications*, 209:116–133, 2016. ISSN 0010-4655. doi: https://doi.org/10.1016/j.cpc.2016.07.028. URL https://www.sciencedirect.com/science/article/pii/S0010465516302260.

[146] N. R. Pradhan, D. Rhodes, S. Memaran, J. M. Poumirol, D. Smirnov, S. Talapatra, S. Feng, N. Perea-Lopez, A. L. Elias, M. Terrones, P. M. Ajayan, and L. Balicas. Hall and field-effect mobilities in few layered p-wse_2 field-effect transistors. *Scientific reports*, 5(1):8979–8979, 2015. ISSN 2045-2322.

[147] N. R. Pradhan, D. Rhodes, S. Memaran, J. M. Poumirol, D. Smirnov, S. Talapatra, S. Feng, N. Perea-Lopez, A. L. Elias, M. Terrones, P. M. Ajayan, and L. Balicas. Hall and field-effect mobilities in few layered p-wse_2 field-effect transistors. *Scientific reports*, 5(1):8979–8979, 2015. ISSN 2045-2322.

[148] K. M. Price, S. Najmaei, C. E. Ekuma, R. A. Burke, M. Dubey, and A. D. Franklin. Plasma-enhanced atomic layer deposition of hfo2 on monolayer, bilayer, and trilayer mos2 for the integration of high-κ dielectrics in two-dimensional devices. *ACS Applied Nano Materials*, 2(7):4085–4094, 2019. doi: 10.1021/acsanm.9b00505. URL https://doi.org/10.1021/acsanm.9b00505.

[149] P. Pulay. Ab initio calculation of force constants and equilibrium geometries in polyatomic molecules. *Molecular Physics*, 17(2):197–204, 1969. doi: 10.1080/00268976900100941. URL https://doi.org/10.1080/00268976900100941.

[150] C. Qiu, F. Liu, L. Xu, B. Deng, M. Xiao, J. Si, L. Lin, Z. Zhang, J. Wang, H. Guo, H. Peng, and L.-M. Peng. Dirac-source field-effect transistors as energy-efficient, high-performance electronic switches. *Science*, 361(6400):387–392, 2018. doi: 10.1126/science.aap9195. URL https://www.science.org/doi/abs/10.1126/science.aap9195.

[151] D. Querlioz and P. Dollfus. *The Wigner Monte Carlo Method for Nanoelectronic Devices: A Particle Description of Quantum Transport and Decoherence*. Hoboken, 2010. ISBN 9781118618479.

[152] B. Radisavljevic, A. Radenovic, J. Brivio, V. Giacometti, and A. Kis. Single-layer mos2 transistors. *Nature nanotechnology*, 6(3):147–150, 2011. ISSN 1748-3387.

[153] A. Rawat, A. K. Gupta, and B. Rawat. Performance projection of 2-d material-based cmos inverters for sub-10-nm channel length. *IEEE Transactions on Electron Devices*, 68(7):3622–3629, 2021. doi: 10.1109/TED.2021.3072880.

[154] R. Resta. Deformation-potential theorem in metals and in dielectrics. *Phys. Rev. B*, 44:11035–11041, Nov 1991. doi: 10.1103/PhysRevB.44.11035. URL `https://link.aps.org/doi/10.1103/PhysRevB.44.11035`.

[155] S. Richard, F. Aniel, and G. Fishman. Energy-band structure of ge, si, and gaas: A thirty-band $\mathbf{k}\cdot\mathbf{p}$ method. *Phys. Rev. B*, 70:235204, Dec 2004. doi: 10.1103/PhysRevB.70.235204. URL `https://link.aps.org/doi/10.1103/PhysRevB.70.235204`.

[156] D. Rideau, M. Feraille, L. Ciampolini, M. Minondo, C. Tavernier, H. Jaouen, and A. Ghetti. Strained si, ge, and $si_{1-x}ge_x$ alloys modeled with a first-principles-optimized full-zone $k\cdot p$ method. *Phys. Rev. B*, 74:195208, Nov 2006. doi: 10.1103/PhysRevB.74.195208. URL `https://link.aps.org/doi/10.1103/PhysRevB.74.195208`.

[157] D. Rideau, M. Feraille, M. Michaillat, Y. Niquet, C. Tavernier, and H. Jaouen. On the validity of the effective mass approximation and the luttinger k.p model in fully depleted soi mosfets. *Solid-State Electronics*, 53(4):452–461, 2009. ISSN 0038-1101. doi: https://doi.org/10.1016/j.sse.2008.08.006. URL `https://www.sciencedirect.com/science/article/pii/S0038110108004024`. Special Issue with papers selected from the Ultimate Integration on Silicon Conference, ULIS 2008.

[158] K. Rim, S. Narasimha, M. Longstreet, A. Mocuta, and J. Cai. Low field mobility characteristics of sub-100 nm unstrained and strained si mosfets. In *Digest. International Electron Devices Meeting,*, pages 43–46, 2002. doi: 10.1109/IEDM.2002.1175775.

[159] J. Robertson and R. M. Wallace. High-k materials and metal gates for cmos applications. *Materials science & engineering. R, Reports : a review journal*, 88:1–41, 2015. ISSN 0927-796X.

[160] M. Roser and H. Ritchie. Technological change. *Our World in Data*, 2013. https://ourworldindata.org/technological-change.

[161] P. Ruffieux, J. Cai, N. C. Plumb, L. Patthey, D. Prezzi, A. Ferretti, E. Molinari, X. Feng, K. Mullen, C. A. Pignedoli, and R. Fasel. Electronic structure of atomically precise graphene nanoribbons. *ACS nano*, 6(8): 6930–6935, 2012. ISSN 1936-0851.

[162] A. Ruiz-Serrano, N. D. M. Hine, and C.-K. Skylaris. Pulay forces from localized orbitals optimized in situ using a psinc basis set. *The Journal of chemical physics*, 136(23):234101–234101, 2012. ISSN 0021-9606.

[163] M. P. L. Sancho, J. M. L. Sancho, and J. Rubio. Quick iterative scheme for the calculation of transfer matrices: application to mo (100). *Journal of physics. F, Metal physics*, 14(5):1205–1215, 1984. ISSN 0305-4608.

[164] P. Sarangapani, J. Charles, and T. Kubis. Tuning band tails in mono- and multilayered transition-metal dichalcogenides: A detailed assessment and a quick-reference guide. *Phys. Rev. Appl.*, 17:024005, Feb 2022. doi: 10.1103/PhysRevApplied.17.024005. URL `https://link.aps.org/doi/10.1103/PhysRevApplied.17.024005`.

[165] A. Schenk and S. Sant. Tunneling between density-of-state tails: Theory and effect on esaki diodes. *Journal of applied physics*, 128(1):14502, 2020. ISSN 0021-8979.

[166] M. Schröter, T. Rosenbaum, P. Chevalier, B. Heinemann, S. P. Voinigescu, E. Preisler, J. Böck, and A. Mukherjee. Sige hbt technology: Future trends and tcad-based roadmap. *Proceedings of the IEEE*, 105(6):1068–1086, 2017. doi: 10.1109/JPROC.2015.2500024.

[167] J. Schwinger. On the green's functions of quantized fields. i. *Proceedings of the National Academy of Sciences - PNAS*, 37(7):452–455, 1951. ISSN 0027-8424.

[168] Y. Shao and J. A. del Alamo. Sub-10-nm diameter vertical nanowire p-type gasb/inassb tunnel fets. *IEEE electron device letters*, 43(6):846–849, 2022. ISSN 0741-3106.

[169] Y. Shao, M. Pala, D. Esseni, and J. A. del Alamo. Scaling of gasb/inas vertical nanowire esaki diodes down to sub-10-nm diameter. *IEEE transactions on electron devices*, 69(4):2188–2195, 2022. ISSN 0018-9383.

[170] P.-C. Shen, Y. Lin, H. Wang, J.-H. Park, W. S. Leong, A.-Y. Lu, T. Palacios, and J. Kong. Cvd technology for 2-d materials. *IEEE Transactions on Electron Devices*, 65(10):4040–4052, 2018. doi: 10.1109/TED.2018.2866390.

[171] P.-C. Shen, C. Su, Y. Lin, A.-S. Chou, C.-C. Cheng, J.-H. Park, M.-H. Chiu, A.-Y. Lu, H.-L. Tang, M. M. Tavakoli, G. Pitner, X. Ji, Z. Cai, N. Mao, J. Wang, V. Tung, J. Li, J. Bokor, A. Zettl, C.-I. Wu, T. Palacios, L.-J. Li, and J. Kong. Ultralow contact resistance between semimetal and monolayer semiconductors. *Nature (London)*, 593(7858):211–217, 2021. ISSN 0028-0836.

[172] M. Shin. Full-quantum simulation of hole transport and band-to-band tunneling in nanowires using the k.p method. *Journal of applied physics*, 106(5):054505, 2009. ISSN 0021-8979.

[173] M. Shin, W. J. Jeong, and J. Lee. Density functional theory based simulations of silicon nanowire field effect transistors. *Journal of applied physics*, 119(15):154505, 2016. ISSN 0021-8979.

[174] M. Shin, W. J. Jeong, and J. Lee. Density functional theory based simulations of silicon nanowire field effect transistors. *Journal of applied physics*, 119(15):154505, 2016. ISSN 0021-8979.

[175] N. Singh, A. Agarwal, L. Bera, T. Liow, R. Yang, S. Rustagi, C. Tung, R. Kumar, G. Lo, N. Balasubramanian, and D.-L. Kwong. High-performance fully depleted silicon nanowire (diameter /spl les/ 5 nm) gate-all-around cmos devices. *IEEE electron device letters*, 27(5):383–386, 2006. ISSN 0741-3106.

[176] W. H. Sio and F. Giustino. Unified ab initio description of fröhlich electron-phonon interactions in two-dimensional and three-dimensional materials. *Phys. Rev. B*, 105:115414, Mar 2022. doi: 10.1103/PhysRevB.105.115414. URL `https://link.aps.org/doi/10.1103/PhysRevB.105.115414`.

[177] T. Sohier, M. Calandra, and F. Mauri. Two-dimensional fröhlich interaction in transition-metal dichalcogenide monolayers: Theoretical modeling and first-principles calculations. *Phys. Rev. B*, 94:085415, Aug 2016. doi: 10.1103/PhysRevB.94.085415. URL `https://link.aps.org/doi/10.1103/PhysRevB.94.085415`.

[178] T. Sohier, M. Calandra, and F. Mauri. Density functional perturbation theory for gated two-dimensional heterostructures: Theoretical developments and application to flexural phonons in graphene. *Phys. Rev. B*, 96: 075448, Aug 2017. doi: 10.1103/PhysRevB.96.075448. URL `https://link.aps.org/doi/10.1103/PhysRevB.96.075448`.

[179] T. Sohier, D. Campi, N. Marzari, and M. Gibertini. Mobility of two-dimensional materials from first principles in an accurate and automated framework. *Phys. Rev. Materials*, 2:114010, Nov 2018. doi: 10.1103/PhysRevMaterials.2.114010. URL `https://link.aps.org/doi/10.1103/PhysRevMaterials.2.114010`.

[180] T. Sohier, E. Ponomarev, M. Gibertini, H. Berger, N. Marzari, N. Ubrig, and A. F. Morpurgo. Enhanced electron-phonon interaction in multivalley materials. *Phys. Rev. X*, 9:031019, Aug 2019. doi: 10.1103/PhysRevX.9.031019.

[181] K. Sotthewes, R. van Bremen, E. Dollekamp, T. Boulogne, K. Nowakowski, D. Kas, H. J. W. Zandvliet, and P. Bampoulis. Universal fermi-level pinning in transition-metal dichalcogenides. *Journal of physical chemistry. C*, 123(9):5411–5420, 2019. ISSN 1932-7447.

[182] I. Souza, N. Marzari, and D. Vanderbilt. Maximally localized wannier functions for entangled energy bands. *Phys. Rev. B*, 65:035109, Dec 2001. doi: 10.1103/PhysRevB.65.035109. URL `https://link.aps.org/doi/10.1103/PhysRevB.65.035109`.

[183] S. Steiger, M. Povolotskyi, H.-H. Park, T. Kubis, and G. Klimeck. Nemo5: A parallel multiscale nanoelectronics modeling tool. *IEEE Transactions on Nanotechnology*, 10(6):1464–1474, 2011. doi: 10.1109/TNANO.2011.2166164.

[184] S. Su, P. Das, S. Ge, and R. K. Lake. Graphene contacts to a hfse2/sns2 heterostructure. *The Journal of Chemical Physics*, 146(6):064701, Feb 2017. ISSN 1089-7690. doi: 10.1063/1.4975178. URL `http://dx.doi.org/10.1063/1.4975178`.

[185] J. Sun, A. Ruzsinszky, and J. P. Perdew. Strongly constrained and appropriately normed semilocal density functional. *Phys. Rev. Lett.*, 115:036402, Jul 2015. doi: 10.1103/PhysRevLett.115.036402.

[186] J. H. Sung, H. Heo, S. Si, Y. H. Kim, H. R. Noh, K. Song, J. Kim, C.-S. Lee, S.-Y. Seo, D.-H. Kim, H. K. Kim, H. W. Yeom, T.-H. Kim, S.-Y. Choi, J. S. Kim, and M.-H. Jo. Coplanar semiconductor–metal circuitry defined on few-layer mote2 via polymorphic heteroepitaxy. *Nature nanotechnology*, 12(11):1064–1070, 2017. ISSN 1748-3387.

[187] A. Szabó, R. Rhyner, and M. Luisier. Ab initio simulation of single- and few-layer mos_2 transistors: Effect of electron-phonon scattering. *Phys. Rev. B*, 92:035435, Jul 2015. doi: 10.1103/PhysRevB.92.035435.

[188] A. Szabo, C. Klinkert, D. Campi, C. Stieger, N. Marzari, and M. Luisier. <italic>ab initio</italic> simulation of band-to-band tunneling fets with single- and few-layer 2-d materials as channels. *IEEE Transactions on Electron Devices*, 65(10):4180–4187, Oct 2018. ISSN 1557-9646. doi: 10.1109/ted.2018.2840436. URL `http://dx.doi.org/10.1109/TED.2018.2840436`.

[189] H. Taghinejad, M. Taghinejad, A. A. Eftekhar, Z. Li, M. P. West, M. H. Javani, S. Abdollahramezani, X. Zhang, M. Tian, T. Johnson-Averette, P. M. Ajayan, E. M. Vogel, S.-F. Shi, W. Cai, and A. Adibi. Synthetic engineering of morphology and electronic band gap in lateral heterostructures of monolayer transition metal dichalcogenides. *ACS Nano*, 14(5):6323–6330, 2020. doi: 10.1021/acsnano.0c02885.

[190] J. Tao, J. P. Perdew, V. N. Staroverov, and G. E. Scuseria. Climbing the density functional ladder: Nonempirical meta–generalized gradient approximation designed for molecules and solids. *Phys. Rev. Lett.*, 91:146401, Sep 2003. doi: 10.1103/PhysRevLett.91.146401.

[191] A. Tsormpatzoglou, C. Dimitriadis, R. Clerc, Q. Rafhay, G. Pananakakis, and G. Ghibaudo. Semi-analytical modeling of short-channel effects in si and ge symmetrical double-gate mosfets. *IEEE transactions on electron devices*, 54(8):1943–1952, 2007. ISSN 0018-9383.

[192] A. Tsormpatzoglou, C. Dimitriadis, R. Clerc, G. Pananakakis, and G. Ghibaudo. Semianalytical modeling of short-channel effects in lightly doped silicon trigate mosfets. *IEEE transactions on electron devices*, 55(10): 2623–2631, 2008. ISSN 0018-9383.

[193] E. Vandamme, P. Jansen, and L. Deferm. Modeling the subthreshold swing in mosfet's. *IEEE electron device letters*, 18(8):369–371, 1997. ISSN 0741-3106.

[194] W. G. Vandenberghe, A. S. Verhulst, K.-H. Kao, K. M. D. Meyer, B. Sorée, W. Magnus, and G. Groeseneken. A model determining optimal doping concentration and material's band gap of tunnel field-effect transistors. *Applied Physics Letters*, 100:193509, 2012.

[195] R. Venugopal, Z. Ren, S. Datta, M. S. Lundstrom, and D. Jovanovic. Simulating quantum transport in nanoscale mosfets: Real vs. mode space approaches. *Applied Physics Letters*, 2002.

[196] I. Vurgaftman, J. R. Meyer, and L. R. Ram-Mohan. Band parameters for iii–v compound semiconductors and their alloys. *Journal of Applied Physics*, 89:5815–5875, 2001.

[197] R. Wan, X. Cao, and J. Guo. Simulation of phosphorene schottky-barrier transistors. *Applied Physics Letters*, 105:163511, 2014.

[198] J. Wang, A. Al-Khalidi, L. Wang, R. Morariu, A. Ofiare, and E. Wasige. 15-gb/s 50-cm wireless link using a high-power compact iii-v 84-ghz transmitter. *IEEE transactions on microwave theory and techniques*, 66(11): 4698–4705, 2018. ISSN 0018-9480.

[199] S. Wang, D. Zhang, B. Li, C. Zhang, Z. Du, H. Yin, X. Bi, and S. Yang. Ultrastable in-plane 1t–2h mos2 heterostructures for enhanced hydrogen evolution reaction. *Advanced energy materials*, 8(25):n/a, 2018. ISSN 1614-6832.

[200] G. C. Wick. The evaluation of the collision matrix. *Phys. Rev.*, 80:268–272, Oct 1950. doi: 10.1103/PhysRev.80.268. URL `https://link.aps.org/doi/10.1103/PhysRev.80.268`.

[201] G. D. Wilk, R. M. Wallace, and J. M. Anthony. High-κ gate dielectrics: Current status and materials properties considerations. *Journal of applied physics*, 89(10):5243–5275, 2001. ISSN 0021-8979.

[202] W. Y., K. JC., L. Y., M. KY., H. S., K. M., S. HS., J. HY., and C. M. P-type electrical contacts for 2d transition-metal dichalcogenides. *Nature*, 6, 2022.

[203] E. Yalon, Ö. B. Aslan, K. K. H. Smithe, C. J. McClellan, S. V. Suryavanshi, F. Xiong, A. Sood, C. M. Neumann, X. Xu, K. E. Goodson, T. F. Heinz, and E. Pop. Temperature-dependent thermal boundary conductance of monolayer mos2 by raman thermometry. *ACS applied materials & interfaces*, 9 49:43013–43020, 2017.

[204] L. Yang, C.-H. Park, Y.-W. Son, M. L. Cohen, and S. G. Louie. Quasiparticle energies and band gaps in graphene nanoribbons. *Physical review letters*, 99(18):186801–186801, 2007. ISSN 0031-9007.

[205] T. Yang, B. Zheng, Z. Wang, T. Xu, C. Pan, J. Zou, X. Zhang, Z. Qi, H. Liu, Y. Feng, W. Hu, F. Miao, L. Sun, X. Duan, and A. Pan. Van der waals epitaxial growth and optoelectronics of large-scale wse2/sns2 vertical bilayer p–n junctions. *Nature communications*, 8(1):1–9, 2017. ISSN 2041-1723.

[206] K. Yılmaz, A. Farokhnejad, F. Criado, B. Iñíguez, F. Lime, and A. Kloes. Direct source-to-drain tunneling current in ultra-short channel dg mosfets by wavelet transform. In *2020 IEEE Latin America Electron Devices Conference (LAEDC)*, pages 1–4, 2020. doi: 10.1109/LAEDC49063.2020.9072953.

[207] M. Zhao, Y. Ye, Y. Han, Y. Xia, H. Zhu, S. Wang, Y. Wang, D. A. Muller, and X. Zhang. Large-scale chemical assembly of atomically thin transistors and circuits. *Nature nanotechnology*, 11(11):954–959, 2016. ISSN 1748-3387.

[208] J.-J. Zhou, J. Park, I.-T. Lu, I. Maliyov, X. Tong, and M. Bernardi. Perturbo: A software package for ab initio electron–phonon interactions, charge transport and ultrafast dynamics. *Computer Physics Communications*, 264:107970, 2021. ISSN 0010-4655. doi: https://doi.org/10.1016/j.cpc.2021.107970.

[209] Y. Zhu, M. K. Hudaita, D. Mohata, B. Rajamohanan, S. Datta, D. Lubyshev, J. M. Fastenau, and A. W. K. Liu. Structural , morphological , and defect properties of metamorphic in 0 . 7 ga 0 . 3 as / gaas 0 . 35 sb 0 . 65 p-type tunnel field effect transistor structure grown by molecular beam epitaxy. 2013.

Publications and conferences

In this thesis

Some of the research leading to this thesis has appeared previously in the following publications.

Articles in peer-review journals

- A. M'foukh, M. G. Pala and D. Esseni: **Full-Band Quantum Transport of Heterojunction Electron Devices With Empirical Pseudopotentials**. – *IEEE Transactions on Electron Devices*, vol. 67, no. 12, pp. 5662-5668, Dec. 2020
- A. M'foukh, J. Saint-Martin, P. Dollfus, M. Pala: **Phonon-assisted transport in van der Waals heterostructure tunnel devices**. – *Solid State Electronics*, vol. 194, p. 108344, Apr. 2022
- A. M'foukh, J. Saint-Martin, P. Dollfus, M. Pala: **Phonon-assisted transport in lateral and van der Waals heterostructure tunnel devices**. – In preparation
- A. M'foukh, P. Giannozzi, D. Esseni, M. Pala: ***Ab-initio* electron-phonon calculations using DFT, DFPT and NEGF methods: applications to 2-D materials mobility and nanoscale FETs**. – In preparation

Conferences

- A. M'foukh, M. Pala: ***Ab-initio* quantum transport simulation of lateral heterostructures based on 2D materials: assessment of the coupling Hamiltonians**.– *IWCN 2021: International Workshop on Computational Nanotechnology*, June 2021
- A. M'foukh, J. Saint-Martin, P. Dollfus, M. Pala: **Phonon-assisted transport in van der Waals heterostructure tunnel devices**. – *HOWDI annual meeting 2022*, May 2022
- A. M'foukh, J. Saint-Martin, P. Dollfus, M. Pala: ***Ab-initio* simulation of phonon-assisted electron transport in van der Waals heterostructures**. – *EUROSOI-ULIS Conference 2022*, May 2022

Résumé en français

Contexte

L'évolution du transistor, depuis son invention en 1947, a connue une tendance continue en suivant la loi de Moore qui consiste à augmenter la densité de transistors présente dans les circuits intégrés en diminuant leur taille. La miniaturisation du transistor par l'industrie de la microélectronique a nécessitée d'importants travaux de recherche pour améliorer les performances des transistors à chaque nouvelle génération. Aujourd'hui, cette strategie de miniaturisation commencent à atteindre ses limites. En atteignant l'échelle nanométrique, le contrôle électrostatique du transistor devient plus délicat et des phénomènes quantiques émergent, ayant pour effet de dégrader les performances du composant. Pour pouvoir continuer à suivre les prédictions de l'IRDS, présentées dans le tableau B.1, les prochaines générations de transistors devront évoluer avec une approche différente.

Table B.1: Caractéristiques électriques prévisionnelles du transistor pour les applications logiques [3]

Année de production	2022	2025	2028	2031	2034	2037
Tension d'alimentation V_{DD}	0.7	0.65	0.65	0.6	0.6	0.6
Pente sous le seuil (mV/dec) - HP	82	72	70	70	70	70
Pente sous le seuil (mV/dec) - HD	75	67	67	65	65	65
I_{ON} (μA/μm) at I_{OFF} = 10 nA/μm - HP	874	787	851	753	737	753
I_{ON} (μA/μm) at I_{OFF} = 100 pA/μm - HD	644	602	656	551	532	547

Une alternative prometteuse au MOSFET, qui est la référence actuelle, est le tunnel FET. En effet, le MOSFET possède une limite intrinsèque pour la pente sous le seuil qui ne peut être plus faible que 60 mV/dec à 300 K. Ce n'est pas le cas du tunnel FET qui peut atteindre des pentes sous le seuil bien plus faibles que cette limite. Cela permet de concevoir des tunnel FETs avec une tension d'alimentation réduite par rapport aux MOSFETs pour avoir un dispositif à faible coût énergétique.

Une autre voie pour atteindre ces prédictions de l'IRDS consiste à exploiter des matériaux différents du Si. Les principaux candidats pour remplacer le Si sont les matériaux III-V et les matériaux 2D. Ils possèdent des propriétés électroniques qui peuvent aider la conception de composants électroniques à faible consommation. Par exemple il est particulièrement intéressant d'avoir une faible masse effective dans la direction du transport, comme nous pouvons

le retrouver dans certains matériaux III-V afin d'atteindre de forts courants à l'état ON.

Pour les matériaux 2D, il est possible d'y trouver de nombreux avantages grâce à leur nature 2D. Leur finesse extrême permet d'obtenir un excellent contrôle électrostatique sur le canal, mais cela permet aussi d'envisager des empilements de différents matériaux 2D pour associer les avantages de différent matériaux et moduler leurs propriétés. Il existe des matériaux 2D métalliques, semi-conducteurs mais également isolants. De ce fait, il est possible de concevoir des dispositifs électroniques ultra-performants seulement à base de matériaux 2D.

Théorie

Pour évaluer le potentiel de ces matériaux, une description précise et efficace sur le plan calculatoire est nécessaire. Pour cela, nous avons choisi de les modéliser à l'aide des Hamiltoniens de type pseudopotentiel développés sur les ondes planes. Nous avons opté pour la base des ondes planes pour les différents avantages qu'elle confère. Elle permet d'obtenir une relaxation précise des atomes pour connaitre leur position, mais il est également aisé d'obtenir une base complète pour décrire les propriétés des matériaux. Dans cette base, les Hamiltoniens peuvent être obtenus soit de manière semi-empirique avec la méthode des pseudopotentiels empirique (EPM), soit de manière *ab-initio* avec la théorie de la fonctionnelle de la densité (DFT). Les deux approches possèdent leurs avantages et inconvénients, mais comme les Hamiltoniens sont décrits dans le même type de base d'ondes planes, les développements réalisés pour l'une peuvent être étendus à l'autre au prix de quelques corrections mineures. Les deux méthodes apportent une description microscopique de la physique des materiaux et nous permettent d'obtenir la structure de bande sur toute la zone de Brillouin. D'un point de vue calculatoire, la méthode EPM est plus efficace grâce à l'utilisation de paramètres empiriques, qui permettent de décrire les Hamiltoniens sur une base comportant un plus faible nombre d'éléments. Mais ceci constitue également l'une des limites de la méthode. En effet, elle nécessite d'avoir accès à des paramètres empiriques qui sont disponibles seulement pour certains matériaux. A l'opposé, la méthode DFT, étant une approche dite *ab initio*, n'est pas concernée par cette limitation et peut fournir un modèle pour tous les matériaux. Cependant, cela se paye par une demande en ressource calculatoire bien plus élevée et un Hamiltonien plus complexe.

Les Hamiltoniens obtenus par les méthodes précédentes nous permettent de calculer les caractéristiques de courant à l'aide du formalisme des fonctions de Green hors équilibre. Ce formalisme nous fournit une évaluation précise des propriétés hors équilibre du dispositif étudié. Les fonctions de Green hors équilibre prennent en compte des phénomènes quantiques (effet tunnel, cohérence,...) qui sont importants à l'échelle nanométrique. A partir des fonctions de Green, il est possible de calculer les propriétés dynamiques (transmission, taux de diffusion,..) et statistiques (courant, densité de charge,..) du système étudié. Pour résoudre efficacement les équations cinétiques de ce formalisme et obtenir les fonctions de Green, les Hamiltoniens obtenus avec les méthodes EPM et DFT dans une base d'ondes planes, ont parfois des dimensions trop importantes. On a alors procédé à deux changements de

base. Le premier consiste à passer de la base des ondes planes à une base hybride utilisant l'espace réel selon la direction du transport (x) et les ondes planes dans les deux directions orthogonales au transport (y-z). Dans cette base, la matrice du Hamiltonien a une forme tridiagonale à bloc ce qui nous permet d'utiliser un algorithme récursif efficace pour calculer les fonctions de Green. Ce changement de base est également nécessaire pour connecter la région du composant simulé à deux contacts semi-infinis par lesquels le courant est injecté et collecté. La seconde transformation sert à réduire le rang de la matrice pour rendre les simulations plus rapides. Pour cela, deux méthodes peuvent être exploitées: la transformation dans l'espace des modes (mode-space) et la transformation dans la base des fonctions de Bloch restreintes à la cellule unitaire (URBF). La première transformation consiste a construire une base avec les vecteurs propres obtenus en résolvant l'équation de Schrödinger dans une section du matériau pour le vecteur d'onde $k_x = 0$. Avec cette dernière condition, la base fonctionne correctement pour réduire le rang de la matrice seulement si la matrice du Hamiltonien est locale. Elle est ainsi particulièrement adaptée aux matrices Hamiltoniennes locaux obtenus avec la méthode des EPM. La base URBF permet, quant à elle, également de réduire le rang de la matrice des Hamiltoniens non locaux et peut donc être utilisée pour les Hamiltoniens DFT. Enfin, les interactions dues aux perturbations extérieures sont également prises en compte par la définition du terme de "self-energy". Les interactions que nous prenons en compte sont celles liées au couplage avec les deux contacts semi-infinis et le couplage avec les phonons.

Résultats

Au court de cette thèse, nous nous sommes principalement concentrés sur deux aspects de la modélisation des composants. Les premiers résultats concernent la modélisation des hétérostructures à partir des Hamiltoniens de type pseudopotentiels. En effet, les hétérostructures de matériaux semi-conducteurs représentent une option technologique prometteuse pour concevoir des composants à haute performance et à faible consommation. Une fois que la méthodologie pour les hétérostructures a été développée, nous nous sommes concentrés sur l'inclusion de l'interaction électron-phonon dans nos simulations sachant qu'elle a généralement un impact non-négligeable sur les performances des transistors.

Hétérostructures avec les Hamiltoniens de type pseudopotentiel empirique

Nous avons commencé par étudier le modèle des hétérostructures avec les Hamiltoniens obtenus par la méthode EPM. En effet, les matériaux III-V choisis (InAs et GaSb) sont particulièrement intéressants pour concevoir des composants basés sur l'effet tunnel et sont accessibles avec cette approche. Pour concevoir un tunnel FET, nous avons choisi une hétérostructure qui a la particularité de former un alignement des bandes de type III, ce qui signifie que le minimum de la bande de conduction de l'un des matériaux est plus bas que le maximum de la bande de valence de l'autre, c'est-à-dire qu'il n'y a pas de recouvrement entre les bandes interdites. Cela permet d'atteindre des courants élevés à l'état

ON puisque la barrière tunnel est supprimée (ou presque) dans cette configuration. Pour décrire plus précisément l'interface de l'hétérostructure, nous avons développés un modèle qui reproduit la transition entre les matériaux à l'interface de manière non abrupte avec l'approximation du crystal virtuels . Cette description se rapproche de ce que nous pouvons retrouver expérimentalement. Une diode Esaki ainsi que des tunnel FETs de type n et p ont été conçus avec l'hétérostructure InAs/GaSb et simulés avec le modèle développé pour les Hamiltoniens EPM. Dans le cas du tunnel FET, nous avons également inclus un opérateur local pour le confinement présent dans ce composant ultrafin. A la fois pour la diode Esaki et pour le tunnel FET, différentes longueurs de région de transition à l'interface ont été investiguées. Les caractéristiques de courant pour la diode Esaki sont montrées dans la figure B.1.

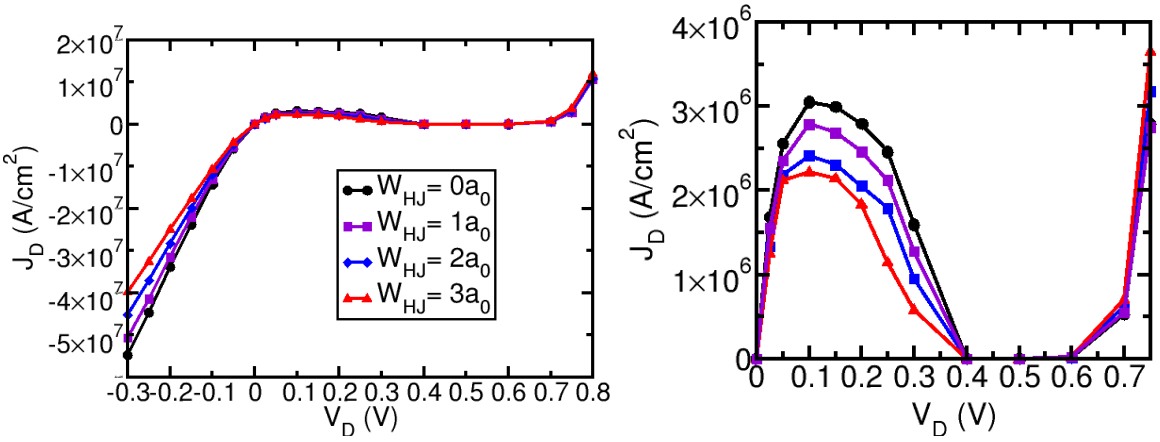

Figure B.1: (à gauche) Densité de courant, J_D, par rapport à la tension appliquée V_D, pour différentes longueurs de région de transition à l'interface de l'hétérostructure, W_{HJ}. (à droite) Zoom sur la région de la résistance différentielle négative.

Dans le le régime de polarisation directe de la diode, nous pouvons observer un phénomène de résistance différentielle négative (NDR), ce qui est dû à la modulation de la fenêtre d'énergie dans laquelle se produit l'effet tunnel. Au niveau du pic de courant, nous pouvons remarquer qu'avec une longueur de la région de transition longue, le courant devient faible. Cette observation peut également être faite pour le régime de polarisation inverse. Cela s'explique par une augmentation de la longueur de la barrière que les porteurs de charge doivent franchir par effet tunnel.

Les caractéristiques I-V pour les tunnel FET sont montrées dans la figure B.2. Les types n et p présentent tous les deux une diminution du courant à l'état ON lorsque la longueur de la région de transition à l'interface augmente.

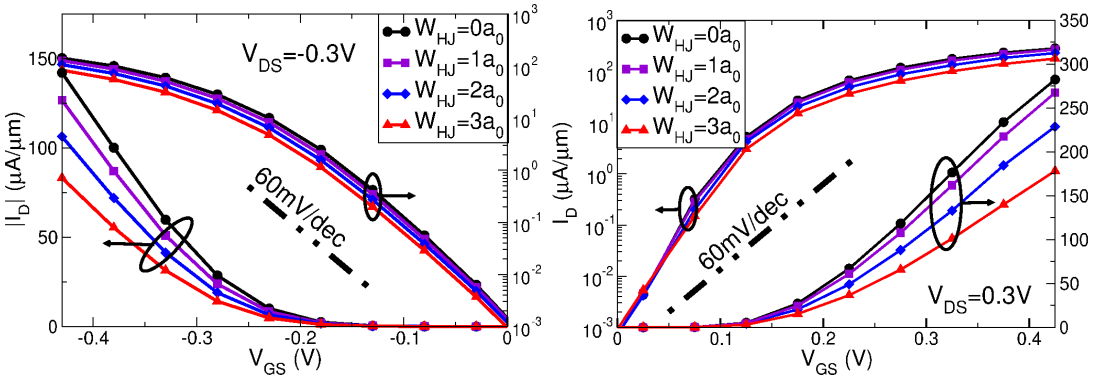

Figure B.2: Caracteristiques I_D-V_D à V_{DS}=0.3 V pour le tunnel FET de type p (à gauche) et de type n (à droite) pour différentes longueurs de région de transition à l'interface de l'hétérostructure, W_{HJ}. Le travail de sortie pour le métal de la grille a été fixé à 4.665 eV pour le type p et 4.625 eV pour le type n.

Le tunnel FET de type n atteint une pente sous le seuil d'environ 30 mV/dec, tandis que le type p présente une pente faiblement sous les 60 mV/dec. Concernant le courant à l'état ON avec une tension d'alimentation $V_{DD} = 0.3$ V, le type p produit une valeur relativement faible de 148 μA/μm. Le courant du type n atteint presque le double de cette valeur à l'état ON. Cette différence est principalement due à la grande dégénérescence dans la source du tunnel FET de type p qui est aussi responsable de la dégradation de la pente sous le seuil comparée au type n.

Hétérostructures avec les Hamiltoniens de type DFT

L'approche développée dans la partie précédente peut être étendue aux Hamiltoniens DFT. En effet, pour des matériaux récemment synthétisés, comme les matériaux 2D, les paramètres empiriques pour les Hamiltoniens EPM ne sont pas connus. Nous sommes donc obligés d'utiliser une méthode *ab-initio*, tel que la DFT, pour obtenir leur Hamiltonien. Comme dans le cas précédant, nous utilisons des Hamiltoniens de type pseudopotentiel développés sur une base d'ondes planes, ce qui nous permet d'utiliser une approche similaire à la méthode des EPM pour décrire les hétérostructures. La principale différence réside dans le terme de couplage à l'interface des deux matériaux. Dans le cas des Hamiltoniens EPM, ce terme de couplage était indépendant des matériaux utilisés du fait de la nature locale du Hamiltonien ce qui n'est plus le cas pour les Hamiltoniens DFT.

Pour estimer ce terme de couplage, nous proposons deux méthodes, la première consiste à utiliser le terme de couplage d'un des matériaux isolés qui compose l'hétérostructure; la deuxième méthode est basée sur l'idée que nous pouvons extraire le terme de couplage entre les deux matériaux dans une matrice Hamiltonien qui décrit une super cellule incluant l'interface. Cette dernière approche devrait produire une meilleure approximation comparée à la première méthode comme le terme de couplage est directement calculé par la DFT. Pour tester les deux approches, nous étudions deux systèmes différents. Le premier système est un MOSFET composé de l'hétérostructure entre une monocouche et une bicouche de $PtSe_2$. Ce dispositif a pour objectif d'exploiter la propriété particulière de ce matériau qui présente une transition semi-conducteur-métal lorsque nous augmentons le nombre des couches. En

utilisant la bicouche du $PtSe_2$ dans la source et le drain, nous devrions obtenir des résistances de contact plus faibles que ce qui est atteint avec les autres matériaux 2D actuels.

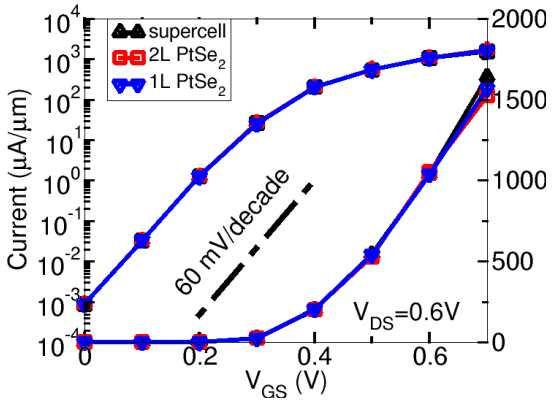

Figure B.3: Caracteristiques de transfert pour le MOSFET fait de multicouche de $PtSe_2$ calculées à $V_{DS} = 0.6$ V en utilisant le terme de couplage des matériaux isolé (1L $PtSe_2$ et 2L $PtSe_2$) et celui extrait de la super cellule.

Les caractéristiques I-V du MOSFET sont présentées dans la figure B.3. Le courant a été calculé avec la méthode du matériau isolé, en utilisant le terme de couplage du monocouche (1L $PtSe_2$) dans un cas et de la bicouche (2L $PtSe_2$) dans le second cas, et aussi avec la méthode de la super cellule. Bien que simple, la première méthode se révèle être proche de la deuxième pour ce système. Dans ce cas précis, les deux méthodes donnent des résultats similaires, ce qui semble être une conséquence du fait que les deux matériaux sont composés des mêmes types d'atomes et que le courant est essentiellement latéral et implique seulement les atomes directement connectés à l'interface.

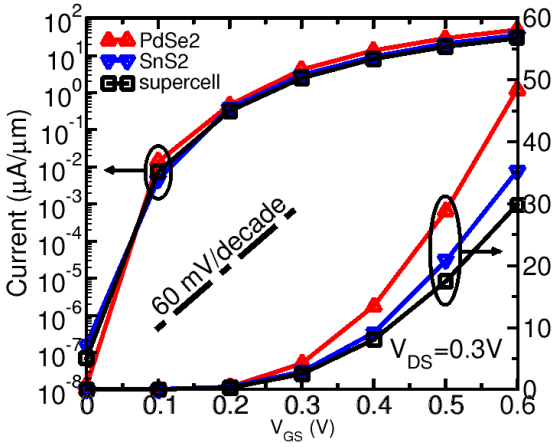

Figure B.4: Caracteristiques de transfert pour le MOSFET réalisé par le multicouche de $PtSe_2$ calculées à $V_{DS} = 0.3$ V en utilisant le terme de couplage des matériaux isolé ($PdSe_2$ et SnS_2) et celui extrait de la super cellule fait avec l'hétérostructure $PdSe_2/SnS_2$

Le deuxième système étudié est une hétérostructure latérale composée du monocouche de $PdSe_2$ et du monocouche de SnS_2 utilisée pour concevoir un tunnel FET. Les caractéristiques I-V du tunnel FET sont présentées dans la figure B.4. Dans ce cas, les résultats des deux approches diffèrent. La simulation réalisée avec le terme de couplage extrait

de la super cellule présente le courant à l'état ON le plus faible. Dans ce cas, l'origine de la différence est directement liée au terme de couplage, comme une inspection des profils de bande et de la densité d'état locale le montre. Les résultats obtenus dans cette partie permettent de montrer les performances prometteuses du MOSFET en $PtSe_2$. Pour le tunnel FET, les performances électriques sont bien plus mauvaises que celles obtenues avec le TFET en III-V de la partie précédente. Cependant, il est également important de rappeler que ces résultats ont été obtenus dans une approximation balistique, ce qui correspond au scénario idéal pour les performances du composant.

Transport assisté par les phonons dans les hétérostructure de matériaux 2D

Une option intéressante afin d'améliorer le courant ON des tunnel FETs est d'utiliser les hétérostructure de Van der Waals (vdW) comme il est possible d'y fermer la bande interdite. Pour que cela soit possible avec un champ électrique réaliste, il faut trouver une hétérostructure de vdW possédant un alignement des bandes de type II, proche du type III. Nous avons identifié un couple de matériaux 2D correspondant à cet alignement des bandes. Il s'agit du monocouche 1T-$HfSe_2$ avec le monocouche 1T-SnS_2. Les calculs DFT indiquent que l'inversion des bandes entre la plus basse bande de conduction du monocouche 1T-SnS_2 et la plus haute bande de valence du monocouche 1T-$HfSe_2$ sous l'action d'un champ électrique extérieur est réalisable.

Sachant également que l'interaction électron-phonon (él-ph) a un impact non-négligeable sur ce type de composant, nous avons complété notre modèle en les prenant en compte sous la forme de self-energies décrites dans l'approximation du potentiel de déformation qui décrit le couplage des électrons (et des trous) avec les phonons acoustiques et optiques. Les résultats indiquent que pour les deux matériaux, le phonon polaire est dominant parmi les phonons optiques et sera donc le seul mode pris en compte dans notre simulation pour les phonons optiques. Le modèle de l'interaction él-ph est simplifié en faisant plusieurs approximations: les potentiels de déformation sont supposés indépendant du vecteur d'onde du phonon considéré, les bandes des phonons optiques sont supposées non dispersives et celles des phonons acoustiques sont supposées comme linéaires. Pour les deux matériaux, les propriétés des phonons et les potentiels de déformation él-ph sont calculées *ab initio* à partir de la théorie de la fonctionnelle de la densité perturbée (DFPT).

Pour étudier l'intérêt de l'hétérostructure de vdW, nous simulons deux architectures de tunnel FETs à base de monocouche de $HfSe_2$ et monocouche de SnS_2. Un tunnel FET vertical, qui est composé de l'hétérostructure verticale des monocouches $HfSe_2$ et SnS_2 et un second qui est composé de l'hétérostructure latérale des mêmes matériaux. L'impact des phonons est également étudié sur les deux dispositifs pour lesquels nous avons réalisé des simulations avec des valeurs differentes du couplage él-ph en jouant sur le taux de diffusion des électrons.

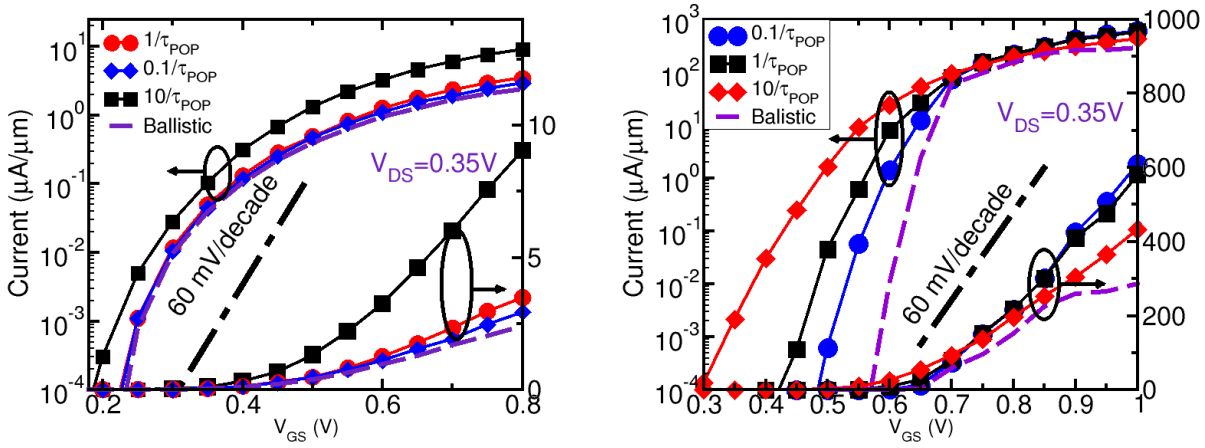

Figure B.5: (à gauche) Caractéristiques de transfert du tunnel FET avec l'hétérostructure latérale pour l'approximation balistique et différent taux de diffusion pour les électrons à $V_{DS} = 0.35$ V. (à droite) La même simulation mais pour le tunnel FET verticale.

Les résultats de nos simulations pour les deux dispositifs sont présentés dans la figure B.5. Ils suggèrent que le tunnel FET vertical possède de meilleures propriétés de transport que le tunnel FET latéral bien qu'il soit plus sensible au couplage él-ph. Il est intéressant de remarquer que l'interaction él-ph produit des effets inattendus sur les dispositifs. Nous observons que pour toutes les simulations où les phonons sont pris en compte le courant est plus important que dans l'approximation balistique. L'analyse de ces simulations révèle que l'origine de cet effet est une transmission inter-vallée qui est induite par l'interaction él-ph. Cela permet au courant à l'état ON d'atteindre des valeurs relativement élevées pour le tunnel FET vertical avec un courant de 580 μA/μm à $V_{DS} = 0.35$ V. Pour le tunnel FET latéral, le courant à l'état ON est aussi augmenté par le couplage avec les phonons, cependant, il reste faible avec une valeur de 3 μA/μm à $V_{DS} = 0.35$ V. Cette différence de courant entre les composants provient du parcours relativement long que doivent effectuer les porteurs de charge par effet tunnel dans le tunnel FET latéral. En effet, dans le cas du tunnel FET vertical, les porteurs n'ont pas de barrière à franchir dans l'état ON grâce à l'inversion entre bande de valence et bande de conduction ce qui améliore grandement la transmission. Au niveau de la pente sous le seuil, dans l'approximation balistique, les deux composants démontrent une pente sous le seuil très faible avec 15 mV/dec. Cependant, quand le couplage él-ph est inclus, la pente est dégradée. Pour le tunnel FET latéral, la dégradation est légère avec une pente de 24 mV/dec, mais pour le tunnel FET vertical, nous obtenons une pente de 40 mV/dec. Bien que le tunnel FET vertical soit plus impacté par les phonons, il reste le dispositif le plus prometteur en dépassant les recommandations de l'IRDS qui préconise un courant à l'état ON supérieur à 500 μA/μm et un SS inférieur à 65 mV/dec.

Calculs *ab-initio* de l'interaction él-ph pour le transport quantique

L'objectif de cette dernière partie est d'améliorer le modèle utilisé pour décrire l'interaction él-ph dans nos simulations du transport. En effet, les approximations faites dans le modèle précédent telles que l'indépendance du potentiel de déformation par rapport au vecteur d'onde du phonon considéré et la structure de bandes simplifiée peuvent aboutir à une représentation inexacte de l'interaction. Pour aller au delà du modèle des potentiels de déformation, nous exploitons le fait que notre base composée de fonctions de Bloch restreintes à la cellule unitaire est parfaitement adaptée pour inclure les éléments de matrice él-ph calculés avec la DFPT. Utiliser ces éléments de matrice pour calculer la self-energy est l'une des méthodes les plus précises pour modéliser cette interaction. Dans ce modèle, la self-energy obtenue est dépendante du vecteur d'onde de l'électron et du phonon et du mode phononique. Même la fréquence phononique garde la dépendance sur le vecteur d'onde et sur le mode. Pour tester ce modèle, nous avons choisis de l'appliquer au MoS_2 monocouche, qui est l'un des matériaux 2D les plus étudiés dans la littérature.

La première étude que nous avons réalisée consiste à calculer la mobilité limitée par les phonons pour le monocouche du MoS_2.

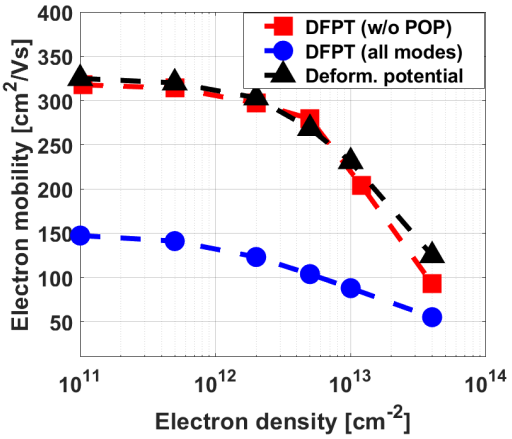

Figure B.6: Mobilité limitée par les phonons pour le monocouche du MoS_2 calculée en considérant (carrés) tous les modes de phonons sauf le mode polaire (POP), tous les modes de phonon (cercles) et calculée avec les potentiels de déformation (triangles).

Le calcul de la mobilité est illustré par la figure B.6. La figure montre que le mode du phonon polaire dans nos simulations induit une forte réduction de la mobilité à faible densité de porteurs de charge comparé au cas où il n'est pas pris en compte. Nos résultats sans le mode de phonon polaire sont proches des résultats que nous retrouvons dans la littérature, où le mode du phonon polaire est absent, avec une mobilité proche de 300 $cm^2/(V.s)$ avec une concentration de porteurs de charge de 2×10^{12} cm^{-2}. Lorsque nous incluons le mode polaire, nos résultats chutent à 111 $cm^2/(V.s)$ avec une concentration de porteurs de charge de 2×10^{12} cm^{-2}. Cependant, ce résultat est cohérent avec les calculs de mobilité où les éléments de matrice él-ph calculés avec la DFPT sont utilisés.

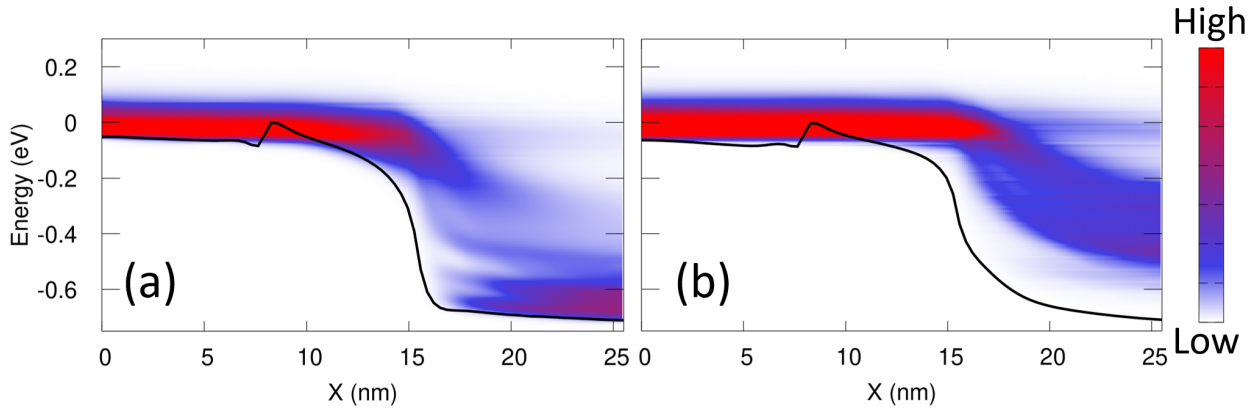

Figure B.7: (Couleur) Courant spectral selon dans la direction du transport à $V_{GS} = 0.7$ V et $V_{DS} = 0.68$ V pour le FET à base de monocouche de MoS_2. a) simulations utilisant la self-energy calculée à partir des éléments de matrice él-ph; b) simulations employant l'approximation des potentiels de déformation.

Nous avons ensuite réalisé une comparaison avec la méthode des potentiels de déformation pour un dispositif MOSFET à base de monocouche de MoS_2. Le résultat du courant spectral est illustré par la figure B.7 et montre que le modèle utilisant les éléments de matrice él-ph produit un couplage plus important entre les électrons et les phonons. Il en résulte en une relaxation en énergie plus complexe comparé à la simulation avec les potentiels de déformation.

Finalement, nous terminons notre étude en utilisant la flexibilité de notre modèle en simulant un dispositif FET technologiquement pertinent où nous utilisons le monocouche de MoS_2 pour le FET de type n et le monocouche de WSe_2 pour le FET de type p. La self-energy calculée avec les éléments de matrice él-ph est utilisée pour les deux dispositifs et une étude sur l'effet de la longueur du canal est également réalisée.

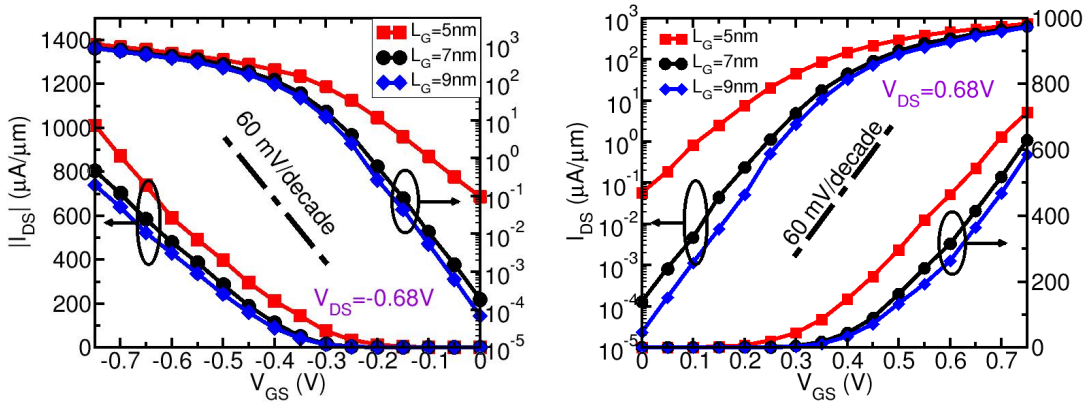

Figure B.8: (à gauche) Caractéristiques de transfert à V_{DS} = -0.68 V pour le FET de type p à base de monocouche de WSe_2 comportant une longueur de grille $L_G \approx 5$ nm (carrés), $L_G \approx 7$ nm (cercles) and $L_G \approx 9$ nm (losanges). (à droite) La même figure mais à $V_{DS} = 0.68$ V pour le FET de type n à base de monocouche de MoS_2.

Les caractéristiques de transfert des deux FET sont présentées dans la figure B.8. Il en résulte que même en présence des phonons, les deux dispositifs, avec une longueur de canal de 7 nm, démontrent des performances

encourageantes qui satisfont les préconisations de l'IRDS pour 2028 avec un courant à l'état ON de 717 μA/μm pour le WSe_2 et 518 μA/μm pour le MoS_2. Un excellent contrôle électrostatique est également observé pour les deux dispositifs avec un *SS* de 63 mV/dec pour le monocouche de WSe_2 et 65 mV/dec pour le monocouche de MoS_2. Ces résultats permettent de mettre en avant le potentiel de ces matériaux pour une prochaine génération de composants électroniques.

Conclusion

Les travaux réalisés lors de cette thèse ont permis de développer des méthodes de modélisation des dispositifs électroniques à partir des Hamiltoniens décrits par les pseudopotentiels. Une attention particulière a été portée à la description des hétérostructures et à l'interaction él-ph. Les développements réalisés ont permis d'étudier plus précisément les interfaces de ces systèmes grâce à une approche microscopique qui permet de décrire une transition entre les différents matériaux de manière réaliste comme nous l'avons vu avec la partie sur les hétérostructures des matériaux III-V. A l'aide de ce modèle, nous avons pu mettre en avant l'intérêt technologique des tunnel FETs réalisé à base des hétérostructures de matériaux III-V. L'étude des hétérostructures a été étendue aux matériaux 2D grâce aux Hamiltoniens DFT. Suite à ces travaux, nous avons complexifié nos simulations en ajoutant l'interaction él-ph modélisée par des potentiels de déformation calculés par des méthodes *ab-initio* (DFPT). Nos simulations ont ainsi pu démontrer l'importance de prendre en compte le couplage avec les phonons pour les tunnel FETs étudiés ainsi que la supériorité des hétérostructures de vdW comparé aux hétérostructures latérales. Les premières sont une option prometteuse pour les prochaines générations de composant électronique. Enfin, nous avons amélioré le modèle de l'interaction él-ph dans nos simulations en utilisant les éléments de la matrice de couplage él-ph pour décrire les self-energies dans le formalisme NEGF. Les simulations avec ce modèle ont démontré l'intérêt des matériaux 2D pour concevoir des dispositifs CMOS capables de satisfaire les métriques de performance des futurs dispositifs électroniques.

Comme perspective de ces travaux, nous pourrons améliorer le modèle de l'interaction él-ph en prenant en compte les effets non locaux de cette interaction. En effet, les phonons polaires produisent une interaction dont le caractère non-local a été négligé dans nos simulations. Cependant, des méthodes existent pour permettre de les prendre en compte efficacement dans les calculs. D'un point de vue du dispositif, il serait intéressant d'appliquer la méthode développée pour décrire l'interaction él-ph dans la dernière partie aux tunnel FET étudiés lors de cette thèse et en particulier au tunnel FET vertical. Cela permettrait de vérifier, avec une méthode plus sophistiquée, si le dispositif reste performant même en présence de phénomène dissipatifs et d'autres types de défauts. Par exemple, il est connu que les pièges dans les tunnel FETs sont responsables d'une importante dégradation de leur performance à cause du phénomène du tunneling assisté par les pièges et par des processus multi-phonons. Il serait donc également intéressant de rajouter ce type de perturbation pour évaluer encore plus finement ces dispositifs prometteurs.