CORA Cork Open Research Archive
Cartlann Taighde Oscailte Chorcaí

Title	Computational modeling of defects in nanoscale device materials
Authors	Greene-Diniz, Gabriel
Publication date	2014
Original Citation	Greene-Diniz, G. F. 2014. Computational modeling of defects in nanoscale device materials. PhD Thesis, University College Cork.
Type of publication	Doctoral thesis
Rights	© 2014, Gabriel F. Greene-Diniz. - http://creativecommons.org/licenses/by-nc-nd/3.0/
Download date	2026-11-16 10:00:02
Item downloaded from	https://hdl.handle.net/10468/2046



UCC
University College Cork, Ireland
Coláiste na hOllscoile Corcaigh



ucc
Coláiste na hOllscoile Corcaigh, Éire
University College Cork, Ireland



Computational Modeling of Defects in Nanoscale Device Materials

Gabriel Francis Greene-Diniz

Thesis submitted for the degree of Doctor of Philosophy

October 2014

Supervisor: Prof. James C. Greer

NATIONAL UNIVERSITY OF IRELAND, CORK

TYNDALL NATIONAL INSTITUTE

Abstract

This PhD thesis concerns the computational modeling of the electronic and atomic structure of point defects in technologically relevant materials. Identifying the atomistic origin of defects observed in the electrical characteristics of electronic devices has been a long-term goal of first-principles methods. First principles simulations are performed in this thesis, consisting of density functional theory (DFT) supplemented with many body perturbation theory (MBPT) methods, of native defects in bulk and slab models of $In_{0.53}Ga_{0.47}As$. The latter consist of (100) - oriented surfaces passivated with Al_2O_3. Our results indicate that the experimentally extracted midgap interface state density (D_{it}) peaks are not the result of defects directly at the semiconductor/oxide interface, but originate from defects in a more bulk-like chemical environment. This conclusion is reached by considering the energy of charge transition levels for defects at the interface as a function of distance from the oxide. Our work provides insight into the types of defects responsible for the observed departure from ideal electrical behaviour in III-V metal-oxide-semiconductor (MOS) capacitors.

In addition, the formation energetics and electron scattering properties of point defects in carbon nanotubes (CNTs) are studied using DFT in conjunction with Green's function based techniques. The latter are applied to evaluate the low-temperature, low-bias Landauer conductance spectrum from which mesoscopic transport properties such as the elastic mean free path and localization length of technologically relevant CNT sizes can be estimated from computationally tractable CNT models. Our calculations show that at CNT diameters pertinent to interconnect applications, the 555777 divacancy defect results in increased scattering and hence higher electrical resistance for electron transport near the Fermi level.

Acknowledgements

Firstly, I would like to express my sincere gratitude to Prof. Jim Greer, not only for the opportunity to conduct this research, but for the consistent and reliable support, patience, and encouragement over the last 5 years. I am very fortunate to be guided by such an experienced and knowledgeable educator of science.

I would like to thank my thesis committee, Dr. Giorgos Fagas and Prof. Stephen Fahy, for very helpful and fruitful discussions during each project review. A special thank you is extended to Giorgos for advice and guidance throughout the PhD project - on the TiMeS code and electronic transport in general. Very big thanks to all of my friends in the Materials Modelling for Devices Group and Electronics Theory Group. Everyone here is always eager to help and enthusiastic for discussion. This fosters a friendly working atmosphere that is invaluable to any postgraduate experience. In particular, I would like to thank Dr. Sarah Jones for her collaboration and expertise.

I also thank the high-*k* group at Tyndall for their collaboration and helpful discussions at regular meetings, especially Dr. Paul Hurley, Dr. Karim Cherkaoui, Dr. Vladimir Djara, Dr. Scott Monaghan and Dr. Éamon O'Connor. I thank the Irish Research Council for Science, Engineering, and Technology for financial support and also the Irish Centre for High-End Computing for provision of computational resources. I am also very grateful to the Intel Corporation for financial support and for the fantastic opportunity to undertake an internship at Santa Clara, California. The experience gained during that internship will undoubtedly prove extremely useful throughout my career; thank you very much to Dr. Sadasivan Shankar, Dr. Mike Haverty, Dr. Carlos Martinez, Dr. Christopher George, Dr. Patrick Theofanis, and Dr. Kelin Kuhn.

Last but certainly not least I wish to convey my sincerest thanks and appreciation to my family and friends, in particular my parents Michael and Irene to whom this thesis is dedicated. This work would not have been possible without their love, support, encouragement, reassurances, and guidance throughout my time at Tyndall and throughout my life.

Publications

G. Greene-Diniz, S. L. T. Jones, G. Fagas, M. Haverty, C. Martinez Lacambra, S. Shankar, J. C. Greer, "Divacancies in carbon nanotubes and their influence on electron scattering", *J. Phys.: Condens. Matter* ***26*** *(2014) 045303 (8pp).*

S. L. T. Jones, **G. Greene-Diniz**, M. Haverty, S. Shankar, J. C. Greer, "Effect of Structure on Electronic Properties of the Iron-Carbon Nanotube Interface", *Chem. Phys. Lett.* ***615*** *(2014) 11-15.*

In preparation:

G. Greene-Diniz, J. C. Greer, "Electronic structure of bulk $In_{0.53}Ga_{0.47}As$ within the *GW* Approximation", *in preparation, Applied Physics Letters (2014).*

G. Greene-Diniz, J. C. Greer, "First principles modeling of defects in Al_2O_3 passivated $In_{0.53}Ga_{0.47}As$", *in preparation, Physical Review B (2014).*

Conferences, Workshops, and Internships

1. *Theory and Practical Use of Quantum Espresso*. Tyndall National Institute and the Irish Centre for High-End Computing (ICHEC). June, 2010.

2. **G. Greene-Diniz**, S. L. T. Jones, L. Ansari, B. Feldman, G. Fagas and J. C. Greer in collaboration with Intel, "Computational Modeling of Electron Transport at Carbon Nanotube/Metal Junctions", *Intel European Research and Innovation Conference (ERIC)*, Leixslip, Ireland. October, 2010. (Poster).

3. **G. Greene-Diniz**, S. L. T. Jones, L. Ansari, B. Feldman, G. Fagas and J. C. Greer in collaboration with Intel, "Electron Scattering through Defective Carbon Nanotubes", *Directed Self-Assembly of Materials*, Materials Research Society, Nashville, Tennessee, USA. September, 2011. (Poster).

4. **G. Greene-Diniz**, S. L. T. Jones, L. Ansari, B. Feldman, G. Fagas and J. C. Greer in collaboration with Intel, "Electron Scattering through Defective Carbon Nanotubes", *Intel European Research and Innovation Conference (ERIC)*, Leixslip, Ireland. October, 2011. (Poster).

5. Internship with Intel at Santa Clara, California, USA. October 2012 - April 2013.

6. **G. Greene-Diniz**, P. K. Hurley, F. Ferdousi, R. Rios, K. Kuhn, and J. C. Greer, "First Principles Modeling of Defects in Al_2O_3 passivated InGaAs", *Workshop on Dielectrics in Microelectronics (WoDiM)*, Kinsale, Ireland. June 2014. (Oral).

7. Presentation of results at EU project networking meeting by P. K. Hurley, *3D integration of InGaAs nMOSFETs above SiGe pMOSFETS*, December 2014.

Declaration

I, Gabriel Francis Greene-Diniz, certify that this thesis is my own work and I have not obtained a degree in this university or elsewhere on the basis of the work submitted in this thesis.

Contents

1. Introduction 1

2. Motivation 4

2.1 Introduction 4

2.2 Device Physics and Defects 4

2.3 Indium Gallium Arsenide 13
2.3.1 Recent Experimental Work 16
2.3.2 Computational Modeling 17

2.4 Carbon Nanotubes 20

3. Theory and Methods 26

3.1 Introduction 26

3.2 Overview of First-principles Methods 28

3.3 Density Functional Theory 31
3.3.1 Hohenberg-Kohn DFT 31
3.3.2 Kohn-Sham DFT 33
3.3.3 Geometry Optimization 36
3.3.4 The DFT Self-consistency Loop 37
3.3.5 Limitations to Approximate DFT 39
3.3.6 Application 40

3.4 The GW Method 45
3.4.1 Photoemission Spectroscopy 45
3.4.2 Quasiparticles and Green's Functions 47
3.4.3 The GW Approximation of the Self-energy 50
3.4.4 Qualitative Features of the Electronic Self-energy 52
3.4.5 Strengths and Limitations 54

3.5 Electronic Transport from the Green's Function Method 55
3.5.1 The Scattering Matrix 56
3.5.2 Calculation of the Green's Function 58
3.5.3 Localization 60

3.6 Charge Transition Levels 62
3.6.1 Application of the GW Method 67

3.7 CNT Vacancy Formation 70

3.8 Electron Transport in CNTs 71
3.8.1 DFT + Green's Functions 71
3.8.2 Mean Free Path 73

4. Bulk InGaAs 75

4.1 Introduction 75

4.2 Electronic Structure 75
4.2.1 The Virtual Crystal Approximation 75
4.2.2 A 64-Atom Supercell Model 83

4.3 Defects - V_{Ga}, Ga_{As} and As_{Ga} 86
4.3.1 Methods 86
4.3.2 Structural Properties 87
4.3.3 Charge Transition Levels for Bulk Defects 90
4.3.4 Conclusions 97

5. InGaAs Surface Passivated with Al_2O_3 101

5.1 Introduction 101

5.2 Passivation Mechanisms 102

5.3 Defects - As_2, Ga_{As} and As_{Ga} 105
5.3.1 Methods 106
5.3.2 Results 109
5.3.3 Conclusions 124

6. Divacancies in Carbon Nanotubes and their Influence on Electron Scattering 130

6.1 Introduction 130

6.2 Methods 131
6.2.1 Geometry and Electronic Structure 131

6.3 Results 132
6.3.1 Vacancy Formation 132
6.3.2 Transport Properties 136

6.3 Conclusions 146

7. Summary and Conclusions 148

References 154

Chapter 1
Introduction

Over the past few decades, considerable scientific interest has been focused on alternative design strategies aimed at future complimentary metal-oxide-semiconductor (CMOS) technologies which are anticipated to overcome the present limitations to Moore's law scaling [1], *i.e.* the reduction in device dimensions in order to increase functionality per unit area while simultaneously decreasing manufacturing costs, resulting in a doubling of performance every 18 months. These strategies include proposals for novel nanoscale devices which could replace the present-day interconnects between transistors on an integrated circuit (IC), as well as the channel inside a metal-oxide-semiconductor field-effect transistor (MOSFET) - the basic building block of processors and memory devices. Specifically, a large proportion of recent research has been focused on materials which present promising electrical characteristics such as carbon nanotubes (CNTs) and ternary III-V compounds. In reality, such device materials invariably exhibit crystalline defects which dramatically affect their optical and electrical properties. For example, consider the semiconductor channel in a MOSFET; surface preparation techniques, interfaces with oxides, dopant implantation, annealing for dopant activation, the resulting ultra-sharp doping gradients, as well as strain, lattice-mismatch, and various other processing techniques can all result in lattice sites that are miscoordinated relative to the bulk crystal, in many cases resulting in electrically active defects such as recombination centers or compensating defects. Thus a large effort has been undertaken to understand and control precisely the nature of defects in technologically relevant materials [2-5].

In this thesis, the effect of crystalline disorder on the properties of such device materials is studied using computational models derived from *first-principles* methods for calculating electronic structure and transport, namely density functional theory (DFT) supplemented with Green's functions. Such

computational models provide an important complement to experiments used for defect characterization [6, 7] by aiding in the identification of defects responsible for observed measurements.

In certain situations, defects are intentionally induced to modify or tune the salient chemical/electrical properties for the purposes of sensor applications; *e.g.* vacancies in CNTs significantly increase their sensitivity to certain gases [8]. Undesired defects, on the other hand, lead to degraded electrical properties which prevent the large-scale integration of the disordered material, *e.g.* electrically active defects in indium gallium arsenide (InGaAs). In both cases, a thorough understanding of the atomic-scale dependence of device properties on defects is crucial to utilizing the materials for nanotechnological applications most effectively.

The structure of this thesis is as follows. In chapter 2, the motivation for this work is discussed, including the significance of defects with regard to device physics followed by the technological benefits and challenges associated with the materials studied. Chapter 3 contains a brief survey of first-principles methods and their use in atomic-scale defect characterization. This involves a background to the DFT formalism along with an outline of Green's function techniques. Specifically, the different contexts in which Green's functions can be applied are highlighted, such as the calculation of the electron self-energy via the GW approximation thereby yielding accurate bandgaps and defect levels, and the scattering problem for evaluating transport. The remaining chapters concern specific applications of these methods. In Chapter 4, the utility of the DFT+GW approach for calculating bandstructures is tested on a simple 2-atom model of $In_{0.53}Ga_{0.47}As$ in which the virtual crystal approximation is used for the cation sublattice. The DFT+GW approach is then extended to a more rigorous model of bulk $In_{0.53}Ga_{0.47}As$ in the presence of native point defects. The transitions between defect charge states as a function of the Fermi energy can be used to identify the defect species responsible for the experimentally inferred density of interface states (D_{it}) derived from capacitance-voltage (CV) measurements. However, performing a direct comparison between defects calculated within a 3-

dimensional periodic bulk simulation supercell and those measured in a MOS capacitor is incomplete without a consistent treatment of the semiconductor/oxide interface. To overcome this limitation, a surface model of $In_{0.53}Ga_{0.47}As$ is presented in Chapter 5. Using the model of Robertson *et al* [9] for Al_2O_3 passivation of $In_{0.53}Ga_{0.47}As$, a stable surface termination is constructed which adheres to electron counting rules. This allows for a consistent comparison between defects in bulk- and surface-like environments. In Chapter 6, localized basis sets are used to study the transport properties of defective CNTs as a function of diameter and chirality. Our results are validated against a more rigorous formalism [7, 10, 11] in which the onset of strong Anderson localization upon the introduction of divacancies is shown. This validation allows us to estimate the localization length and other salient transport features in CNTs of more technologically relevant diameters. The conclusions and an overall summary of the thesis are given in chapter 7.

Previous theoretical studies of defects in these device materials have either used methods that restrict the size of the system being modeled [11, 12] (in which CNT diameters larger than those of (10, 10) chirality were not studied) thereby preventing direct comparison with technologically relevant diameters, or use a model in which direct comparison with experiment is incomplete, for example computing properties of a bulk model which are then used to compare with oxide terminated surfaces [13]. In this thesis, first principles models of defects in InGaAs and CNTs are constructed and the electronic properties thereof evaluated, allowing for a comparison with, and extension of, previous computational studies using explicit, unparameterized computational methods for direct comparison to experiment.

Chapter 2
Motivation

2.1 Introduction

This chapter starts with an account of the role atomic-scale defects play in nanoelectronic devices, and their characterization using a combination of experimental and theoretical methods. The following sections give an outline of the importance of the materials studied in this work, namely indium gallium arsenide ($In_xGa_{1-x}As$) and carbon nanotubes (CNTs). The former represents an alternative to Si-based metal-oxide field effect transistors (MOSFETs) which is relatively compatible with existing CMOS technology. Carbon nanotubes, on the other hand, offer potentially tremendous improvements in IC performance, yet their large-scale integration poses a much more difficult challenge in terms of fabrication and processing. Thus, these potential device materials represent two ends of a spectrum of proposed solutions to continued Moore's Law scaling; short-term solutions which are readily adoptable to existing CMOS platforms on one end, and on the other end long-term solutions which likely require significant changes to the manufacturing process. While being less disruptive to the standard fabrication techniques, materials representing short-term solutions offer only incremental improvements in performance whereas materials corresponding to long-term solutions may open the gateway to “beyond-CMOS” technology platforms, once the problems associated with their large-scale integration are addressed.

2.2 Device Physics and Defects

The properties of crystalline solids change dramatically upon the introduction of defects and impurities. The idealized pure crystalline solid, with periodic boundary conditions which eliminate the disruption of symmetry at

surfaces, permits the theoretical derivation of, in principle, all intrinsic crystal properties. The latter are crucial to the determination of performance when employed in an electronic or photonic device. As real materials are never found with perfect crystallinity, but in fact contain interruptions to the regular pattern of lattice sites, for example interfaces with other materials, interfaces with vacuum, and defects, device performance can be significantly influenced by the *lack* of ideal crystallinity.

As device dimensions shrink, the influence of defects on device performance further increases. The level of precision that is necessary when designing modern nanoelectronic devices and in maintaining desired performance is highlighted if one considers the number of transistors on a single microprocessor over the past few decades. The exponential increase in this number reflects the increasing functionality and decreasing cost per integrated circuit.

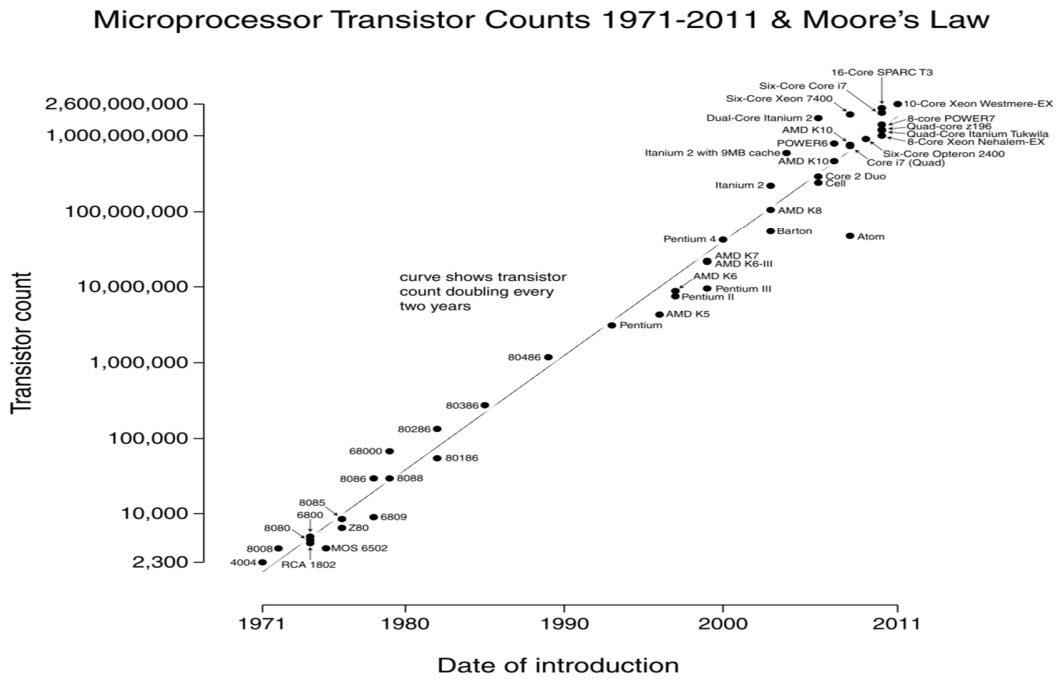


Figure 2.1 - Graphical depiction of Moore's law; exponential increase in the number of transistors in a microprocessor over time. Taken from ref. [14].

As the number of transistors on a single chip reaches the tens of billions, transistor miniaturization enters the atomic-scale regime - consider the International Technology Roadmap for Semiconductors (ITRS) prediction of sub-10nm MOSFET dimensions by 2018 [15] - with all the associated performance benefits of scaling and associated challenges related to the differences between a given material's macroscopic and microscopic properties. To continue the ~50 year trend of device miniaturization, highly sophisticated fabrication and characterization techniques have been developed to achieve the atomistic level of control required by today's nanoelectronic manufacturing standards [16, 17]. These developments reflect the fact that variations in a material at an *atomic*-scale are by no means irrelevant and in fact can have a large impact on the resultant device performance, particularly in the modern nanoelectronics era.

Atomic scale or "point" defects are isolated imperfections in a material's crystal structure. These include native point defects such as missing atoms or vacancies, and atoms not occupying a lattice site or interstitials. Other common native defects include miscoordinated sites such as dangling bonds and dimers. Certain point defects such as antisites are specific to compound semiconductors and consist of atoms native to the material occupying the wrong lattice site; *e.g.* cations occupying anion sites or vice versa in III-V semiconductors. Foreign impurities on the other hand consist of species not native to the host material, which may or may not occupy a lattice site of the host crystal. In the context of transistor physics, point defects can play a significant role in electrical characteristics. For example, they can trap charge carriers at a semiconductor-oxide interface, scatter electron waves propagating through a conductor, as well as pin the Fermi-level which hinders the effect of doping and/or gate voltage. All of these effects can be detrimental to the operation and reliability of MOSFETs as well as their interconnections. Therefore, a wide range of techniques exist for detecting the presence of defects and measuring their effect on device characteristics.

Point defects can be observed as distortions resulting in contrasts in high

resolution transmission electron microscopy (TEM) or scanning tunneling microscopy (STM) images. STM imaging performed by Ebert [18, 19] has revealed a wealth of information concerning defects on (110) oriented III-V compound semiconductor surfaces. Even certain aspects of the nature of bulk defects were inferred from those investigations. However, bulk defects remain difficult to observe and to identify experimentally; this is one area where computational modeling can make important contributions. Also, in ref. [19] Ebert concentrated on the (110) surface, *i.e.* nonpolar, III-V surfaces. For the (100) oriented III-V surface however, the polar nature of anion-cation bonds results in reconstructions that depend on whether the surface is anion or cation terminated. In either case electron-counting rules must be satisfied to ensure the Fermi level lies within the bandgap, possibly resulting in the creation of defects in the absence of a stable passivating oxide [9] such as Al_2O_3. To identify the nature of defects at the semiconductor-oxide interface in these cases, electrical properties can be measured and compared to results from atomistic computational modeling. Before discussing how the effect of defects may be analyzed experimentally, it is useful to discuss the MOSFET as an example and the impact defects can have on its operation.

Consider the operation of a planar n-channel MOSFET, a cross-section of which is shown in figure 2.2. A metal electrode (gate) is deposited onto an insulating oxide, beneath which lies a p-type semiconductor which provides a channel for minority carrier conduction. The semiconducting channel lies between two highly doped n-type electrodes that are labelled source and drain. An electrostatic field resulting from an applied positive gate voltage attracts minority charge carriers from the semiconductor bulk to the semiconductor/oxide interface. As the voltage is increased, the carrier concentration at the semiconductor surface eventually becomes inverted relative to the p-type doped substrate and current flows through the channel from source to drain.

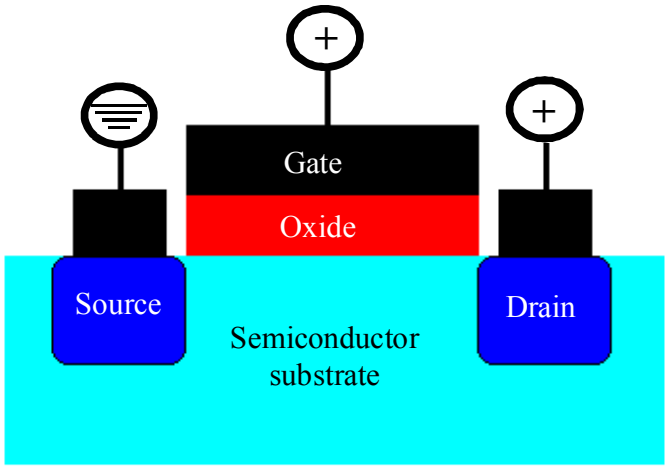


Figure 2.2 - General structure of an n-type MOSFET. Semiconductor is p-type, source and drain are (highly) n-type, so applying a positive voltage above a threshold to the gate creates a conducting channel of electrons underneath the oxide.

The quality and operation of the MOS stack is crucial to the desired performance of the MOSFET. More specifically, the behaviour of the oxide-semiconductor layer should be controlled by the gate potential rather than by unwanted charges present at the semiconductor-oxide interface. In a semiconductor, the valence and conduction bands are separated by the bandgap E_{gap}, and at 0 K all electronic states up to the valence band maximum (VBM) are filled and all conduction band states are unoccupied (even at room temperature, typically very few electrons are promoted to the conduction band in intrinsic silicon). In a MOSFET, minority carrier conduction along the semiconductor surface is generated by a bending of the band edges below the Fermi level (ε_f). As the gate voltage V_{gate} is increased, the semiconductor conduction band minimum (CBM) eventually crosses ε_f and a thin layer of conducting electrons forms at the semiconductor-oxide interface, which may be described as a 2-dimensional electron gas.

The performance of a MOSFET depends heavily on the absence of electronic states within the bandgap; for example, the fact that significant conduction occurs only above threshold, the mobility of the conducting channel, the lifetime of minority carriers, and the related rate of increase of current with

respect to V_{gate}. Atomic-scale defects at the semiconductor surface can dramatically alter the electrical characteristics of a device as they can contribute a finite density of states within the bandgap of the semiconductor. Such "midgap" defects can provide charge carriers at energies lower than E_{gap} and so offset the effect of the gate voltage, for example; in this situation, electrostatic control of the channel by the gate potential is degraded, which compromises subthreshold-slope (SS) and threshold voltage V_{th} stability. Midgap states can also "trap" minority carriers and reduce carrier mobility by scattering, which in turn reduces the current ON/OFF ratio.

The specific electrical properties of defects depend on the position of their energy levels relative to the electronic band edges. "Shallow" defects have energy levels within a few k_BT (k_B is Boltzmann's constant, T is temperature) of the conduction or valence band edges. These defects can donate/accept electrons at energy levels close to the band edges and these shallow defect levels are usually characterized by extended wavefunctions. Thus, they usually have a smaller effect on charge carrier mobility than deep defects. Deep level defects are represented by levels that are close to the center of a bandgap. They can be the result of a defect which produces extensive distortion of the lattice, and correspondingly exhibit localized wavefunctions. Deep level defects can be effective recombination centers. Consider the electron-hole recombination process (figure 2.4), which in the absence of midgap levels requires the emission of a photon (assuming a direct bandgap and neglecting Auger processes) with frequency $f = E_{gap}/h$, where h is Planck's constant. A midgap level provides a means for electrons and holes to recombine via Shockley-Read-Hall (SRH) recombination [20]. The energy loss of the conduction electron is either released as light or transferred to the lattice by a phonon. A deep defect level is considered as a "trap" that captures electrons and holes, reducing their lifetime.

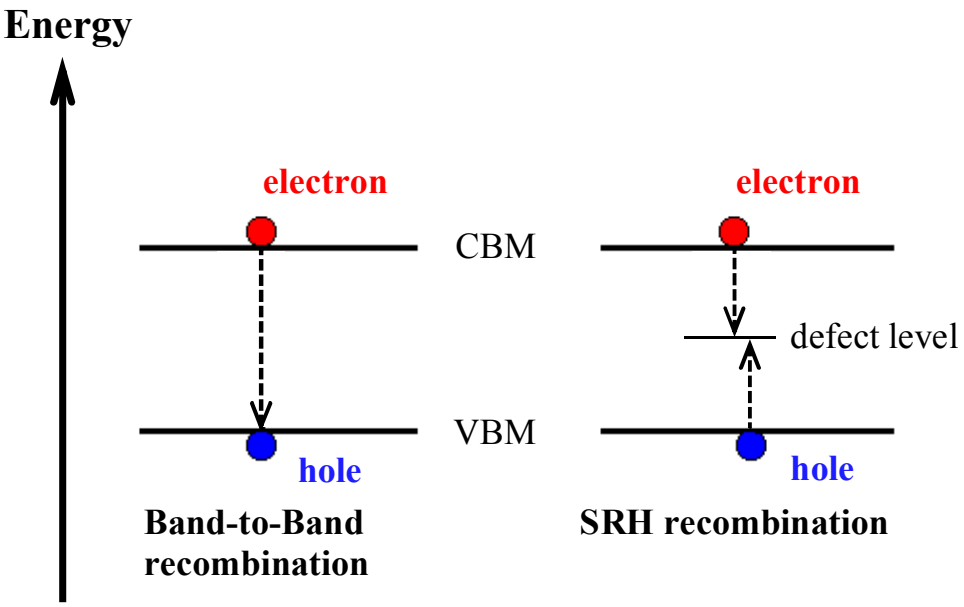


Figure 2.4 - Illustration of electron-hole recombination with and without midgap defects. As the defect level approaches the center of the bandgap, SRH recombination rate approaches a maximum [20].

As the recombination rate U (charge carrier recombination per volume per time) depends on the defect level E_D and the intrinsic Fermi level (*i.e.* center of bandgap) E_i as [21],

$$U \propto \left(\exp\left[\frac{(E_D - E_i}{k_B T}\right]\right)^{-1} \tag{2.1}$$

then the closer the defect level is to midgap, the larger the recombination rate, with a maximum when $E_D = E_i$. Thus deep level defects are the most efficient recombination centers for SRH recombination.

The nature of semiconductor defects in a MOSFET can be studied using a number of experimental methods developed in recent decades. A broad range of defect properties such as their concentration, density of states, and their influence on MOSFET performance, can be analyzed by comparing the measured and theoretical capacitance-voltage (CV) curves [6, 21]. The integrated density of interface states D_{it} can be inferred by measuring the

flatband and threshold voltages and comparing to the theoretical ideal [6]. The disagreement between measurement and theory can be used to study the nature of defects present at the semiconductor-oxide interface. The extraction of voltage shifts is illustrated by the following results recently published by Djara *et al* [6], who measured the CV characteristic of an n-channel MOSFET consisting of an Al_2O_3 gate oxide deposited onto an $In_{0.53}Ga_{0.47}As$ substrate. This system will be considered in detail later in this thesis.

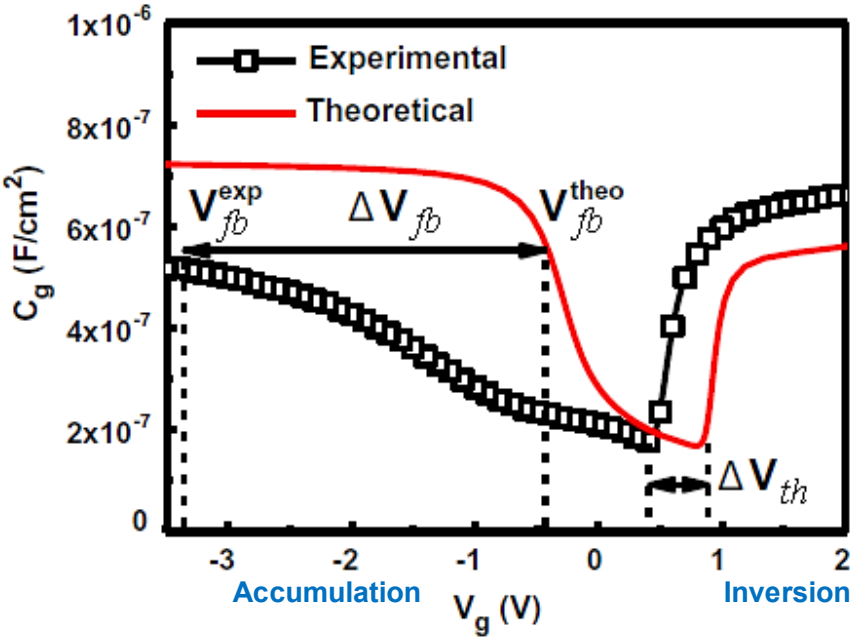


Figure 2.5 - Measured and simulated CV response of a MOS stack, reproduced from [6]. Theoretical CV curve obtained from a self-consistent Poisson-Schrödinger solver. The inversion and accumulation regions are indicated below the voltage axis. The difference between experimental and theoretical curves highlights the impact of defects and is explained in the text.

The presence of interface defects explains the difference between the experimental and theoretical CV curves; the capacitance measured at inversion is higher than the theoretical capacitance in the same region, which indicates the presence of extra states above the CBM, *i.e.* interface defect states resonant with the conduction band. A positive fixed charge in the oxide is also suggested by the difference in V_{th} between theory and as measured.

Electrical characterization techniques in general lack the ability to identify the atomic structure of the defect responsible for an observed deviation from ideal electrical behaviour. More specifically, defect levels from different species in different charge states can be located at similar places in a bandgap, and so an ambiguity exists that is particularly pronounced for compound semiconductors in the physical origin of the salient features of D_{it} profiles. It is precisely this ambiguity that *ab-initio* simulations can help address. Within this approach, *ab-initio* or first-principles methods are making increasingly important contributions to the understanding of defects in semiconductors, their influence on material properties, and their impact on device performance [2, 5, 7, 13, 22]. From these methods, essential information concerning the defect and host system can be calculated, such as charge densities, electrostatic potentials, electronic wavefunctions, density of states, relaxed atomic structures, and formation energies. The latter are highly relevant for a theoretical description; formation energies allow one to calculate the charge transition levels [5, 13, 23] of electrically active defects, which correspond to electron chemical potentials at which two charge states of a given defect center are equally likely to form. These thermodynamic quantities provide an important connection between first-principles theory and experimental electrical data.

Once the experimentally inferred D_{it} profile is obtained, atomistic modeling can be employed to help identify the defect(s). By deriving from first-principles the self-consistent electronic structure and equilibrium geometries of a number of isolated defects incorporated into the semiconductor host, formation energies and hence charge transition levels can be evaluated and compared to values inferred from experiment; details of the calculation of these quantities can be found in chapter 3.

Experimentally accessible fingerprints of defects can also be investigated by a first-principles account of electronic transport through a defective material. By calculating the electronic transmission probability, which directly relates to the low-temperature, linear-response conductance [24], through the defective (or scattering) region contacted to charge carrier reservoirs, the scattering of

electron waves by the defect can be quantitatively assessed. The results of such calculations could be directly compared to experimental transport spectroscopy data, allowing for the identification of a defect, or since the scattering cross section of the defect will depend on its arrangement within the host lattice, specific geometric properties of defects such as bonding arrangement and reconstruction can be distinguished by this approach. When the length of the conductor is reduced to below the elastic mean free path, interesting phenomena arises in the presence of defects such as the possibility of extended electron waves becoming localized by the disruption of the lattice periodicity. In this case, conduction becomes exponentially dependent on the characteristic localization length, which can also be inferred experimentally [7] and compared to theoretical predictions. A background on electron transport and its evaluation within an *ab-initio* framework is given in Chapter 3.

To summarize section 2.2, as device miniaturization continues on into the atomic-scale regime, point defects have an increasingly significant impact on device performance. First-principles computational modeling can then play an important complementary role to the experimental methods employed to study defects in technologically relevant materials. In the following sections, an account of the motivation for the materials studied in this work will be given.

2.3 Indium Gallium Arsenide

To appreciate the importance of $In_xGa_{1-x}As$ and other III-V semiconductors it is instructive to consider some of the present-day limitations to continued Moore's law scaling for transistors. In 1974, R. Dennard *et al* published their seminal paper on linear scaling which described techniques to minimize short-channel effects due to encroaching depletion layers from source-channel and channel-drain *p-n* junctions. This allowed for improvements in transistor density, switching speed and power dissipation simultaneously through linear scaling [25]. They concluded that the key device parameters such as

dimension, operating voltage, and doping concentration should all be scaled by a constant factor. The table below, reproduced from [25], highlights the benefits of this scaling strategy; as a transistor shrinks in size, it can switch faster and dissipate less power, so by increasing transistor density, performance of the integrated circuit can increase without increasing power usage; this provides a technological basis for Moore's law.

Device or Circuit Parameter	***Scaling Factor***
Device dimension t_{ox}, L, W	*1/k*
Doping concentration N_a	*k*
Voltage V	*1/k*
Current I	*1/k*
Capacitance $\varepsilon(L\times W)/t_{ox}$	*1/k*
Delay time VC/I	*1/k*
Power dissipation VI	*$1/k^2$*
Power density VI/(L×W)	*1*

Table 2.1 - Scaling results of Dennard *et al*. Gate oxide thickness, channel length, and width are denoted by t_{ox}, L, and W respectively. ε is the dielectric constant. Until recent years, $k \sim 1.4$ yielding 70% reduction in linear dimension, or ~2x increase in transistor density. This is sometimes referred to as "happy scaling".

Unfortunately, this scaling strategy relied on a number of key assumptions which are no longer valid at modern MOSFET dimensions. Scaling of the oxide thickness and consistent scaling of the threshold voltage V_{th} with the operating voltage are two examples. The latter assumption neglects the influence of sub-threshold leakage on power consumption, which was a negligible effect in the micron-scale MOSFETs of the 1970s. Combining this with the fact that further reductions in operating voltage will compromise performance by reducing switching speed has recently forced the CMOS industry into a regime of "power-constrained scaling" [26, 27]. Power density cannot increase any further

without incurring unfeasibly large manufacturing and operating costs associated with packaging and cooling.

An improvement would be a channel material with higher charge carrier mobility than Si, thereby allowing for a reduction in operating voltage without a slower switching speed. Compound semiconductors composed of elements from column III and column V of the periodic table exhibit such desirable electron mobilities. As discussed in the previous section, crystalline defects can create states inside the bandgap of the host semiconductor. Such defects can degrade key electrical properties such as minority carrier lifetime, subthreshold slope, and ON/OFF ratio, to which III-V compounds are prone.

The lack of a stable native III-V oxide, in contrast to SiO_2 for Si, has led to problems in this regard; direct oxidation of a III-V surface results in Fermi level pinning [28] resulting in the inability of the gate to control the electrostatic potential inside the semiconductor due to a large concentration of surface defects. Progress was made with atomic-layer-deposition (ALD) of Al_2O_3 for the gate oxide, yielding a III-V/high-*k* oxide interface with a manageable density of interface states (D_{it}). It has also been shown that increasing the indium concentration can improve transistor characteristics compared to GaAs [26, 29, 30]. Meanwhile, recent computational studies indicate the interface states arising from As-dimers and corresponding group III dangling bonds - which contribute to midgap D_{it} in GaAs - lie well inside the conduction band of $In_{0.53}Ga_{0.47}As$ [31, 32]. This could explain the improvement in performance with increasing In content, as the lowering of the band gap with increasing In content may push defect states out of midgap, assuming the position of the defect level relative to the valence band edge does not change with In content [30]. In addition, gate-all-around (GAA) $In_{0.53}Ga_{0.47}As$ nanowire transistors [33], and $In_{0.7}Ga_{0.3}As$ quantum well FETs [34] have in recent years been fabricated and in cases shown to perform better than their Si counterparts. While there are still many challenges to the large scale integration of III-V materials on a silicon platform such as minimizing defect densities at the high-*k* oxide interface, these findings result in

$In_{0.53}Ga_{0.47}As$ based MOSFETs being currently considered the most promising of the III-V family for replacing silicon in n-type MOS transistors [26].

2.3.1 Recent Experimental Work

Experimental electrical characterization of III-V based MOS capacitors has been performed by a number of research groups, including but not limited to refs. [35-38]. A recent work reported full gate capacitance measurements of the $Al_2O_3/In_{0.53}Ga_{0.47}As$ system combined with Poisson-Schrödinger modeling to extract the interface state density across the semiconductor bandgap [6]. The results indicate the presence of defects giving rise to midgap states associated with a (net) neutral to positive (donor-like) charge transition, as well as neutral to negative (acceptor-like) interface traps with a D_{it} distribution extending into the conduction band. The donor-like feature of the interface trap density peaks at 0.36 eV above the valence band edge with a peak height of $\sim 1.5\times10^{13}$ cm^2 eV^{-1}. It is not possible to distinguish the location of these traps as coming from either the semiconductor or oxide or both [6]. However, other groups have reported comparable D_{it} profiles using different high-*k* oxides [35, 38] suggesting that these defects are native to $In_{0.53}Ga_{0.47}As$. None of the experiments to date are capable of addressing the physical nature of the defects responsible for the observed measurements.

The atomic structure of point defects on III-V semiconductor surfaces has been investigated in the recent past using scanning tunneling microscopy (STM). While these studies have revealed a large amount of useful information concerning III-V defects, identifying STM images with the atomic defect geometry can sometimes be questionable. For example, LaBella *et al* noticed that at large filled-state sample biases an electronic tip-sample convolution effect can hinder the imaging of trenches between As-dimer rows in GaAs (2×4) reconstructed surfaces [39]. Thus, simulations of the electronic structure can be very useful in this regard particularly when interpreting experimental findings. In the case of ref. [39], the combination of theory and data significantly

strengthened their conclusions regarding the mechanism behind the inability of the STM tip to image dimer trenches at large bias.

2.3.2 Computational Modeling

The field of first-principles computational modeling of defects in III-V materials is now several decades old [2]. Recent density functional theory (DFT) investigations include predictions of the stability and amphoteric nature of a range of native defects using hybrid functionals [40-42] in bulk and oxide-terminated surface models [32], studies of bonding mechanisms [9], and defect states at III-V/oxide interfaces [43], as well as predictions of charge transition levels of defects in bulk Si [44] and III-V surface models [22] using DFT+MBPT (many-body perturbation theory) methods. The latter eliminate the parameterized nature of the hybrid-DFT calculated electronic structure. Also, recent studies by Guo and Robertson of the chemical trends of III-V/Al_2O_3 interfaces [45] show that As dimers have antibonding states near the conduction band minimum of GaAs, while the same defect state is located well inside the conduction band of InAs. The authors use these findings to explain why $In_{0.53}Ga_{0.47}As$ based devices undergo less Fermi-level pinning than GaAs based devices, which was also suggested previously [43]. The position of the acceptor-like charge transition level of the As dimer calculated by Miceli and Pasquarello [32] is consistent with this explanation. Interestingly, in ref. [45] nitrogen is found to be an effective passivant as it destabilizes dimers and pushes their associated states deep into the semiconductor bands; this is a possible explanation of the reduction of D_{it} near the conduction band at $In_{0.53}Ga_{0.47}As$/high-*k* oxide interfaces after nitrogen forming gas anneals [46]. Once again, one can see examples of the important role played by computational modeling in the interpretation of experimental data.

The calculations by Pasquarello and coworkers of bulk III-Vs [13, 40-42] indicate that the Ga_{As} and As_{Ga} antisites have transition levels that correlate with the observed D_{it} peaks from a range of experimental groups [37, 47, 48]. The

figure below shows the charge transition levels calculated using hybrid-DFT functionals [40-42], in which a fraction of non-local Fock exchange is incorporated into the exchange-correlation (XC) functional to correct the bandgap error inherent to DFT; *i.e.* the XC functional is parameterized to reproduce the experimental bandgap.

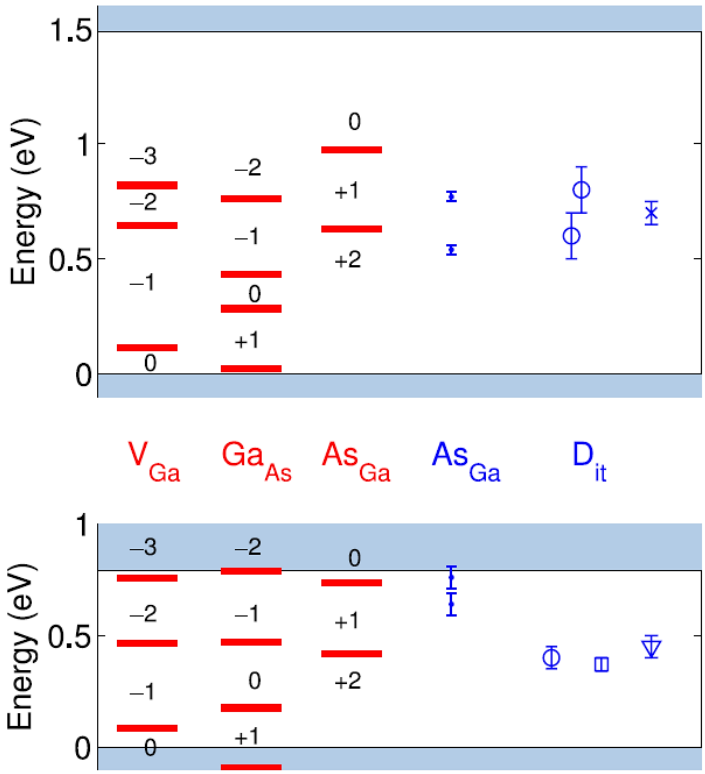


Figure 2.6 – Reproduced from ref. [13]. Red markers denote calculated charge transition levels of native defects in GaAs (top) and $In_{0.5}Ga_{0.5}As$ (bottom) using hybrid functionals. Blue markers indicate experimental results measured for GaAs (top) and $In_{0.53}Ga_{0.47}As$ (bottom) from refs. [37, 47, 48].

This study consisted of calculating charge transition levels using formation energy expressions which reflect two limiting conditions of As-rich and Ga-rich growth. The energy to form a neutral defect center from a pristine bulk lattice [5] is given by

$$E_{form} = E_D - E_{bulk} - \sum_i n_i \mu_i \,, \tag{2.3}$$

where E_D and E_{bulk} are the total energies of a supercell containing the defect(s) and the pristine bulk, respectively. In the work of ref. [13], the two growth conditions are obtained by modifying the form of the chemical potential μ_i of added ($n_i > 0$) or removed ($n_i < 0$) species i in the expression for the defect formation energy. For As-rich growth conditions, the total energy of the As_4 molecule is used to obtain the As chemical potential, while the equilibrium condition of GaAs, $\mu_{GaAs} = \mu_{Ga} + \mu_{As}$ yields the Ga chemical potential. For Ga-rich growth conditions, bulk Ga metal is used to obtain the Ga chemical potential, and the equilibrium condition with GaAs yields the As chemical potential. By studying the formation energetics as a function of growth conditions, Komsa and Pasquarello suggest the As_{Ga} antisite is favored [13].

As in any computational simulation, however, a few simplifying assumptions are made in these models. Firstly, defects calculated in a bulk model yield energy levels that can significantly shift when calculated along an oxide terminated surface. The termination of 3-dimensional periodicity by a surface results in a perturbation of the bulk charge balance and band bending towards the interface/vacuum. In addition, bonding coordination of the bulk is broken at the surface giving rise to new states in the band structure. These and other effects can cause surface defect levels to move with respect to the bulk values, rendering a direct comparison between charge transition levels calculated in a bulk model and those measured near a surface or interface inadequate.

Hybrid density functional calculations such as those carried out for the bulk simulations in ref [13] vary the amount of Fock-exchange until the experimental bandgap is recovered. The parameterization usually stops at this point and no other considerations are given to the remaining E-k dispersion beyond the Γ-point or to the position of defect-induced midgap states [40-42, 49]. The resulting electronic structure is assumed to be accurate once the band edges are corrected. It is believed that *ab-initio* calculations, which accurately describe the bandstructure without empirical corrections, may provide a more consistent or "even-handed" treatment of the electronic structure throughout the Brillouin zone. This is the motivation for our use of the GW approximation

(GWA), and in chapter 4 it is shown that this method makes a substantial difference compared to hybrid functional calculations when applied to midgap defect levels and their associated charge transition levels. The GWA is derived from many-body perturbation theory and provides a correction to the Kohn-Sham (KS) eigenvalues. This correction is obtained from a first-order approximation to the self-energy Σ, *i.e.* truncating the expansion of Σ in the Green's function G and screened Coulomb interaction W after the first term [50]. This correction to the KS eigenvalues tends to increase the separation between conduction and valence bands as a result of the detailed treatment of local field effects, as will be discussed in chapter 3. Thus, this method provides a framework for a consistent treatment of the electronic structure, as all states and eigenvalues included in the DFT simulation are included in the GW correction, not just the band edges. The DFT+GW method has been successfully applied in the past to bulk and surface semiconductors including III-Vs [22, 51, 52], yet the method has not to date been applied to bulk $In_{0.53}Ga_{0.47}As$ and oxide-terminated $In_{0.53}Ga_{0.47}As$ surfaces in the presence of defects.

2.4 Carbon Nanotubes

Carbon nanotubes (CNTs) may be considered as monolayers of graphene rolled into a cylinder. Their electronic properties are highly dependent on diameter and chirality, the latter being defined by the axis of the graphene plane along which the tube is "rolled" up. CNTs belong to a class of nanoscale materials regarded as quasi-one-dimensional; increasing scientific and technological interest is being focused on such materials which can be attributed to their large aspect ratios, surface chemistry and quantum effects. These properties can be harnessed for device integration and multi-functional uses, *e.g.* in spintronic, nano-electromechanical, and memristive devices [53, 54]. Their almost defect free structure and resulting ballistic conduction accounts for the intense effort in utilizing CNTs as ideal materials in nanoelectronics as both a transistor channel

material and as an interconnect [55, 56]. Despite these promising features many problems still exist with regard to their large scale integration. The lack of chirality control during CNT growth effectively precludes their manufacturability, and none of the proposed solutions to chirality separation achieve the purity required by very large scale integration (VLSI) standards [57]. These and other issues lead to the fact that CNTs are likely not to be utilized in VLSI applications in the near future. However, their exciting properties provide a unique "playground" for low-dimensional physics, and if manufacturing difficulties are overcome, CNTs could be expected to yield very large performance improvements over current Si-based devices [57].

As a transistor channel material, semiconducting CNTs offer a number of promising advantages over conventional semiconductor materials. Consider the figure below [57], which shows mobility versus bandgap for CNTs compared to standard semiconductors.

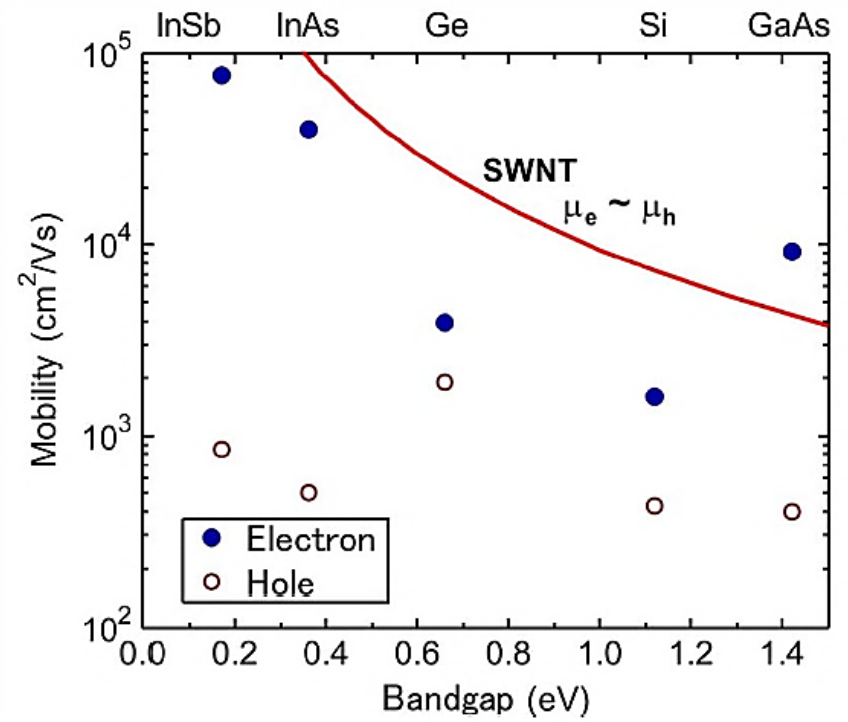


Figure 2.7 – Carrier mobilities of semiconducting single-walled CNTs (SWNTs) versus bandgap, compared to other semiconductors. Reproduced from ref. [57].

For standard semiconductor materials, mobility is limited due to the heavier effective mass of holes compared to electrons. On the other hand, the symmetry of the CNT bandstructure near the valence and conduction band edges equates to roughly similar mobilities for electrons and holes [58]. The correspondence of electron and hole mobilities in CNTs would avoid the necessity of different doping materials for *n*- and *p*-channel transistors. The high mobility of CNTs is related to a large mean free path for carrier scattering, resulting in nearly scattering-free or quasi-ballistic transport for typical nanotube lengths. In addition to desirable mobilities, the thin quasi-one-dimensional structure of CNTs offers another advantage over competing semiconductors. Electrostatic short channel effects (SCEs) plague conventional MOSFETs; as the channel length shortens, the channel potential becomes dependent not just on gate voltage but also on source-drain voltage (V_{SD}) resulting in a drain current that no longer saturates with increasing V_{SD}, as well as an increased OFF current. SCEs can be suppressed by increasing the channel doping concentration but this compromises mobility, and so a trade-off exists between mobility and SCE suppression. The electrostatic profile of the channel can also be improved by decreasing its thickness; therefore the quasi-one-dimensional structure of SWNTs means they are expected to be an ideal material for SCE suppression [59].

Copper has been employed as the interconnect material of choice for VLSI architectures since 1997 due to its low resistivity and high thermal stability compared to aluminium. As dimensions continue to shrink however, interconnect current densities increase. If continued indefinitely, the Cu interconnect will eventually become unusable due to problems associated with high resistance and electromigration [60]. Interconnect resistivity rises with decreasing circuit dimensions due to increases in scattering from surface roughness and at grain-boundaries. This can have serious consequences for IC performance in terms of speed due to elevated interconnect delay. In addition to this, Cu has a tendency to diffuse through the surrounding dielectric and possibly into the transistor channels, disrupting the electrical properties. A barrier layer is

thus needed between Cu and the oxide, potentially increasing resistance as well as process complexity.

CNTs are good candidates for mitigating all of these problems. Their outstanding thermal, mechanical, and electrical properties have attracted significant interest in their application as future interconnects. Stable current carrying capacities on the order of 10^{10} A cm^{-2} even at temperatures as high as 250°C have been observed [61]. Such properties suggest that CNTs could be a reliable solution to the problems of electromigration associated with Cu interconnects, either by replacing Cu entirely or using Cu/CNT composites [62]. Ballistic transport could be a solution to scaled-down interconnects, alleviating the problems of surface and grain boundary scattering. Also, CNTs do not need barrier layers to prevent the diffusion of atoms between the transistors and interconnects.

Regardless of the application, the low-dimensionality of CNTs results in a dramatic change of chemical and physical properties with the occurrence of structural defects. Defects have also been shown to modify the mechanical strength of CNTs [63]. The electronic response can be affected and tailored [64] either by the natural occurrence of defects or through their introduction at post-processing, by using ion irradiation for example. Charge carriers in CNTs are strongly confined and propagation occurs predominantly along a tube axis. As a result, scattering at defects can be pronounced due to the small transverse surface area seen by the propagating charge carriers, resulting in a proportionately large scattering cross section.

Previous transport studies of defective CNTs have shown that for sufficiently high defect densities the resistance increases exponentially with tube length [11, 65]. This occurs at device lengths greater than multiples of the mean free path associated with scattering from defects, and is a consequence of the fundamental effect of disorder on electron transport in low-dimensional systems; namely the metal-insulator transition due to Anderson localization [66]. These findings demonstrate the critical importance of defects in determining the conductivity of single-wall CNTs.

The most commonly observed defects in CNTs under ion irradiation are monovacancies and divacancies [67] leading to a reconstruction of the CNT bonding resulting from the removal of a single C atom or two neighboring C atoms, respectively. The latter have been found to be more stable relative to isolated monovacancy defects [12, 68]. A detailed investigation of the energetics of divacancy formation has determined the relative stabilities for different divacancy configurations [68]. In that work, explicit atomic-scale calculations from first-principles were applied to study nanotubes up to a maximum diameter of 1.4 nm. For larger nanotube radii, an empirical molecular mechanics Hamiltonian was used to describe the carbon bonds, with parameters fitted to the *ab-initio* results obtained from the smaller radii CNTs. Based on these approximations, a crossover between the so-called 585 (a reconstruction of the hexagonal bonding arrangement resulting in an octagon sandwiched between two pentagons) and 555777 (three pentagons interconnected with three heptagons) divacancy defects was predicted to occur at armchair tubes with diameters of approximately 4.0 nm.

In this thesis, formation energetics and electron transport are analyzed as a function of CNT diameter and chirality. The methods for this study are described in chapter 3. The results, presented in chapter 6, suggest that the 555777 defect indeed becomes more stable than the 585 configuration at large diameters. Regarding transport, we show that the localization length of the disordered system may be estimated by invoking the independent scattering approximation (ISA), whereby defects in an array are presumed to scatter charge carriers independently resulting in a simple expression relating the elastic mean free path to the localization length.

To summarize section 2.4, the remarkable properties of CNTs make them one of the most important emerging research materials for both front-end and back-end electronics applications [57]. Their utilization, however, depends on a number of issues associated with processing and fabrication. For example, FETs with a CNT as the channel (CNTFETs) require the CNT to be uniformly coated by the

insulator, desirably a high-*k* oxide formed by ALD [57]. The chemical inertness of the CNT wall results in difficulty of absorption of ALD precursors, and so a functional layer is needed to modify the CNT surface appropriately. Before VLSI applications of CNTFETs can be realized, a functionalization technique which does not degrade electrical current is necessary. More basic fabrication problems also still exist, such as difficulties in separating metallic and semiconducting nanotubes. Most of the proposed chirality separating techniques do not yet achieve the purity required by modern VLSI standards [69-71], as well as possibly introducing transistor hysteresis and other contamination problems [72]. Clearly, materials which have greater compatibility with existing Si-based technology are much more likely to be utilized before CNTs. Nevertheless, due to their remarkable electrical properties CNTs constitute a promising candidate for future generations of VLSI circuits.

Chapter 3
Theory and Methods

3.1 Introduction

This chapter describes the electronic structure approximations applied in this work for the purposes of modeling native defects in technologically relevant materials. The theoretical background is described at the beginning of this chapter, followed by an account of the computational methods. The latter involves describing the methods implemented in software developed [73-76] to perform first-principles atomistic calculations for materials and transport properties.

Density Functional Theory (DFT) lies at the heart of the theoretical machinery employed in this work. The central idea behind DFT is that all ground state properties are functionals of the ground state electronic density. However, if one wishes for an accurate description of properties beyond the ground state (excited states), DFT typically needs to be supplemented with other methods. To this end, Green's functions can be applied in different ways; for example within many-body perturbation theory (MBPT) Green's functions can be used to approximate the "electron self-energy" via the GW approximation (where G represents a Green's function and W stands for the screened Coulomb interaction) allowing one to assess the response of a system to the addition or removal of an electron. Using this method, it will be demonstrated that one may obtain accurate defect levels involved in transitions between charge states, which in turn may be used to complement experimental data related to electrically active defects. Section 3.4 discusses Green's functions in the context of MBPT and specifically, the utility of the GW approximation in obtaining the quasiparticle energy spectrum.

Electronic transport can also be evaluated using Green's functions to calculate the quantum mechanical scattering matrix. In section 3.5, the relevant

concepts of electronic transport theory are discussed, including how the Landauer conductance spectrum may be computed using an approach which combines DFT with Green's function methods. This provides a methodological basis for the results presented in chapter 6.

The calculations of defect formation energetics using total energy differences are described in subsequent sections. Two techniques are used in this work: charge state formation energies for the case of $In_{0.53}Ga_{0.47}As$, and the formation energies of neutral point-defects for the case of CNTs. When considering charged defects, the supercell method results in a spurious Coulomb interaction between periodic images. Many previous studies of charged defects include an explicit correction term that is obtained by extrapolating the formation energy to negligible defect image interactions, which tends to be computationally expensive due to the large simulation cells required. The expressions for charge state formation used in this work benefit from a cancelling of the electrostatic interaction term, in other words an error cancellation. The correction turns out to be unnecessary for the charged defects in $In_xGa_{1-x}As$ to within the overall accuracy of the calculations. Once charge state formation energies are obtained, charge transition levels may be calculated. The compensation for DFT errors with the GW approximation for the purposes of predicting charge transition levels has been successfully applied in the past, both to bulk solids [44] and III-V surfaces [22], and this technique is applied to determine charge transition energies in bulk $In_{0.53}Ga_{0.47}As$ and near $In_{0.53}Ga_{0.47}As$/oxide interfaces.

Section 3.8 describes methods for computing electronic transmission using Green's functions in which DFT-determined matrices representing the electronic structure are used as a starting point. Once a transmission spectrum is calculated, the localization length of the disordered system may be estimated by invoking the independent scattering approximation (ISA). The ISA, which will be validated for the applications in this thesis, is described in this section.

3.2 Overview of First-principles Methods

Following the development of quantum mechanics in the early 20th century, it was quickly realized that some of the difficulty associated with computing the energy and wavefunction of a molecule could be alleviated by a separation of the electronic and nuclear degrees of freedom [77]

$$\psi_{total} \cong \psi_{electronic} \times \psi_{nuclear}\ , \tag{3.1}$$

thereby permitting the solution of an electronic Schrödinger equation in which only the motion of electrons is taken into account. This is known as the Born-Oppenheimer approximation, and is achieved by subtracting out the nuclear kinetic energy operator and inserting the nuclear positions as scalar parameters to the electronic Hamiltonian. The electrons still "feel" the Coloumbic interaction with the nuclei, but the latter are held at fixed positions. The validity of this approximation is rooted in the high ratio between nuclear and electronic mass, such that electrons are assumed to respond instantaneously to a change in the nuclear positions $\{\boldsymbol{R}_{ion}\}$. Calculation of the electronic energy eigenvalue for a range of $\{\boldsymbol{R}_{ion}\}$ defines a potential energy surface (PES), such that minimizing the electronic energy with respect to the nuclear positions to obtain the equilibrium geometry is equivalent to finding the minimum on the PES. If the nuclear kinetic energy is ignored, the Hamiltonian of a system consisting of electrons (denoted by i, with mass m_e and charge e) and ions (denoted by a) can be written in atomic units ($\hbar = m_e = e = 4\pi\epsilon_0 = 1$) as

$$\hat{H} = -\frac{1}{2}\sum_{i}^{N_e} \boldsymbol{\nabla}_i^2 + \frac{1}{2}\sum_{i\neq j}^{N_e} \frac{1}{|\boldsymbol{r}_i - \boldsymbol{r}_j|} - \sum_{i,a}^{N_e,\, N_n} \frac{Z_a}{|\boldsymbol{r}_i - \boldsymbol{R}_{ion,a}|} + \frac{1}{2}\sum_{a\neq b}^{N_n} \frac{Z_a Z_b}{|\boldsymbol{R}_{ion,a} - \boldsymbol{R}_{ion,b}|} \tag{3.2}$$

where N_e is the number of electrons, N_n is the number of nuclei, and Z_a is the atomic number. The Laplacian $\boldsymbol{\nabla}^2$ is taken with respect to the electronic coordinates $\boldsymbol{r_i}$. The first term on the right-hand-side is the electronic kinetic energy operator $\hat{T}$, and is followed, respectively, by the Coloumbic interaction between electrons $\hat{V}_{el-el}$, the attraction of electrons to ionic cores which is usually considered as a fixed, external potential energy $\hat{V}_{ext}$ acting on electrons, and the repulsion between ions V_{ion}. The latter can be regarded as an additive constant to the total energy for any point in the PES. Solution of the time-independent many-electron Schrödinger equation

$$\hat{H}|\psi\rangle = E|\psi\rangle \tag{3.3}$$

yields the many-electron wavefunction, which for N_e electrons must be anti-symmetric with respect to an exchange of electron coordinates in order to satisfy the Pauli-exclusion principle

$$\psi(\boldsymbol{r}_1, \boldsymbol{r}_2, \dots, \boldsymbol{r}_{N_e}) = -\psi(\boldsymbol{r}_2, \boldsymbol{r}_1, \dots, \boldsymbol{r}_{N_e})\,. \tag{3.4}$$

With the ground state wavefunction at hand, one can calculate ground state properties such as electron density (n) and total energy (E) as expectation values of their respective operators [78],

$$n(\boldsymbol{r}) = \frac{\langle\psi|\hat{n}(\boldsymbol{r})|\psi\rangle}{\langle\psi|\psi\rangle}$$

$$= N_e \frac{\int d^3r_2 \int d^3r_3 \dots \int d^3r_{N_e} \left|\psi(\boldsymbol{r}, \boldsymbol{r}_2, \boldsymbol{r}_3, \dots, \boldsymbol{r}_{N_e})\right|^2}{\int d^3r_1 \int d^3r_2 \dots \int d^3r_{N_e} \left|\psi(\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_3 \dots, \boldsymbol{r}_{N_e})\right|^2} \tag{3.5}$$

$$E = \frac{\langle\psi|H|\psi\rangle}{\langle\psi|\psi\rangle} = \langle\hat{T}\rangle + \langle\,\hat{V}_{el-el}\rangle + \langle\hat{V}_{ext}\,\rangle + V_{ion}\,. \tag{3.6}$$

The ground state wavefunction ψ_0 corresponds to the state with the lowest energy, which can be determined by variationally minimizing the total energy with respect to all parameters of $\psi(\{\boldsymbol{r}_i\})$, with the constraints of anti-symmetry (equation 3.4) and orthonormality

$$\langle\psi|\psi\rangle = 1\,. \tag{3.7}$$

For an N_e-particle system ψ is a complex function in $3N_e$-dimensional configuration space. With this in mind, one can appreciate that the space of possible wavefunctions, and thus the cost of calculating expectation values of observables grows extremely rapidly with the number of (interacting) particles [79]. This is the central problem of electronic structure theory; electrons form an interacting many-body system described by a wavefunction that depends not only on the individual positions of electrons, but also on all of their mutual interactions. As Walter Kohn pointed out in his Nobel lecture, wavefunction based methods are tractable for "sufficiently small molecules" highlighting the application by James and Coolidge to H_2 [80]. However, for practical calculations of realistic systems extreme difficulty is encountered when tackling the wavefunction "head-on".

The first step towards circumventing this issue and rendering computations tractable is to make the approximation that electrons are non-interacting and form a homogenous gas. The Thomas-Fermi (TF) model [81, 82] constitutes an early attempt of applying this idea. This model shows how all terms in the Hamiltonian can be written as *functionals* of the density leading to the Thomas-Fermi expression for the total electronic energy,

$$E_{TF}[n] = \frac{3h^2}{10m_e}\left(\frac{3}{8\pi}\right)^{2/3}\int d^3\,r\,n^{5/3}(\boldsymbol{r}) + \int d^3r\,\hat{V}_{ext}(\boldsymbol{r})n(\boldsymbol{r}) + \frac{1}{2}\int d^3r d^3r'\frac{n(\boldsymbol{r})n(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|} \tag{3.8}$$

The first term on the right-hand-side is the TF kinetic energy functional, the last term is known as the Hartree potential, which is the classical electrostatic interaction between electrons due to their mutual repulsion. $\hat{V}_{ext}(\boldsymbol{r})$ is the external potential arising from the interaction of electrons with ionic cores. Due to the inaccuracy of the approximation to the kinetic energy functional and the neglect of any exchange or correlation between electrons, the TF model leads to *qualitatively* inaccurate conclusions, for example failing to reproduce binding and shell structures. It does, however, illustrate how properties of a system can be described as functionals of the density, which reduces the $3N_e$-dimensional problem of wave-function based methods to a 3-dimensional problem in *real* space; thus the role of the wavefunction ψ as the key variable is replaced by the electron density $n(\boldsymbol{r})$. The ability of this method to dramatically reduce the size of the many-electron problem prompted Hohenberg and Kohn to develop a theory in which the ground state electronic structure could be *exactly* described in terms of the electronic density.

3.3 Density Functional Theory

3.3.1 Hohenberg-Kohn DFT

Modern density functional theory (DFT) began from the hypothesis that "*a knowledge of the ground state density of n(r), for any electronic system (with or without interactions), uniquely determines the system"* [79]. Starting with this idea, Hohenberg and Kohn proved two theorems which formed the basis of DFT:

Theorem 1 states that for a system of interacting electrons moving in some external potential $V_{ext}(\boldsymbol{r})$, the ground-state electronic density $n(\boldsymbol{r})$ uniquely determines the external potential up to an additive constant. The immediate implication is that the Hamiltonian is fully determined (up to a constant), and consequently all properties derivable from the Hamiltonian through a solution of the Schrödinger equation are determined, given only the ground state electronic

density.

Theorem 2, also known as the Hohenberg-Kohn variational principle states that since the total energy can be defined as a functional of the density, then it follows that the density that minimizes this functional *is* the ground state density of the system.

Theorem 1 suggests a one-to-one correspondence between density and external potential, *i.e.* the Hohenberg-Kohn total energy functional can be defined for densities that can be generated by some external potential (this is known as "V-representability"). **Theorem 2** provides the first step towards a useful calculation and follows from noting that all (electronic) contributions to the energy are uniquely determined once $n(\boldsymbol{r})$ is specified. Then each such contribution can be seen as a functional of $n(\boldsymbol{r})$, including the total energy functional

$$E[n] = F_{HK}[n] + \int d^3r\, V_{ext}(\boldsymbol{r}) n(\boldsymbol{r}) + V_{ion}\,, \tag{3.9}$$

where the Hohenberg-Kohn universal functional $F_{HK}[n]$ contains all internal kinetic and potential energies of the interacting electron system. Consider a system in which the ground state density $n_1(\boldsymbol{r})$ corresponds to an external potential $V_{ext,1}(\boldsymbol{r})$. Thus, the total energy functional is the expectation value of the Hamiltonian $\widehat{H}_1$ with wavefunction ψ_1,

$$E[n_1] = \langle \psi_1 | \widehat{H}_1 | \psi_1 \rangle\,. \tag{3.10}$$

With **Theorem 1** in mind, it follows that the ground state density $n_1(\boldsymbol{r})$ is unique to the external potential $V_{ext,1}(\boldsymbol{r})$, such that any other density, say $n_2(\boldsymbol{r})$ corresponding to wavefunction ψ_2, leads to an energy that is greater than $E[n_1]$ since

$$E[n_1] = \langle \psi_1 | \hat{H}_1 | \psi_1 \rangle < \langle \psi_2 | \hat{H}_1 | \psi_2 \rangle . \quad (3.11)$$

Thus, if the universal functional $F_{HK}[n]$ is known, then by minimizing the total energy with respect to variations in the density, the exact ground state density and energy can be found. At this point, it should be noted that no prescription is given to find the exact $F_{HK}[n]$ [78]. In practice, approximate functionals which have been constructed from the full many-body wavefunction that includes the effects of exchange and correlation between electrons need to be considered to develop accurate approximations to F_{HK}.

3.3.2 Kohn-Sham DFT

Kohn and Sham subsequently introduced the idea of replacing the original many-body problem with an auxiliary system of non-interacting electrons [83]. This *Ansatz* allows for approximations to be developed for the exchange-correlation (XC) functional defined as the difference between the *exact* universal functional and a functional containing only kinetic and electrostatic (Hartree) potential terms [79]

$$E_{xc}[n] = F_{HK}[n] - \left[T[n] + \frac{1}{2} \int d^3r d^3r' \frac{n(\boldsymbol{r}) n(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|} \right] . \quad (3.12)$$

Kohn and Sham made the approximation that the ground state density of the auxiliary non-interacting electron system is equal to that of the physical interacting system. This leads to a set of soluble independent-particle equations in which all the complicated many-body effects have been incorporated into $E_{xc}[n]$. Each of these equations obey a Schrödinger-like Kohn-Sham (KS) equation in which the auxiliary Hamiltonian contains a kinetic energy operator and an effective potential $V_{KS}(\boldsymbol{r})$ acting on the N electrons

$$\widehat{H}_{KS} = -\frac{\nabla^2}{2} + V_{KS}(\boldsymbol{r})\,, \tag{3.13}$$

and

$$V_{KS} = V_{ext}(\boldsymbol{r}) + V_{Hartree}(\boldsymbol{r}) + V_{xc}(\boldsymbol{r})\,. \tag{3.14}$$

The XC potential is defined as the functional derivative of the XC energy with respect to the density

$$V_{xc}(\boldsymbol{r}) = \frac{\partial E_{xc}(\boldsymbol{r})}{\partial n(\boldsymbol{r})}\ . \tag{3.15}$$

Solutions of the KS equation define the KS orbitals φ_i, with corresponding KS eigenvalues ε_i. The ground state density corresponds to the lowest eigenvalues of the auxiliary Hamiltonian and has one electron per (spin) orbital, such that the ground state density can be written as the sum of squares of the KS orbitals

$$n(\boldsymbol{r}) = \sum_{i}^{N_e} |\varphi_i|^2\ , \tag{3.16}$$

and the KS kinetic energy takes the form

$$T_{KS} = -\frac{1}{2}\sum_{i}^{N_e} \langle \varphi_i | \nabla^2 | \varphi_i \rangle\ . \tag{3.17}$$

This leads to a re-writing of the Hohenberg-Kohn energy into the KS total energy functional

$$E[n] = T_{KS}[n] + \int d^3r\, V_{ext}(\boldsymbol{r}) n(\boldsymbol{r}) + \frac{1}{2}\int d^3r d^3r' \frac{n(\boldsymbol{r}) n(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|} + E_{xc}[n] + V_{ion}\ . \tag{3.18}$$

Therefore, if the *exact* XC functional $E_{xc}[n]$ is known, including the correction to the KS kinetic energy term, then the exact ground state density and total energy of the true many-electron system can be found by solving the KS equations for non-interacting electrons. The exact form of $E_{xc}[n]$ remains unknown [78] but to the extent that reasonably approximate forms of $E_{xc}[n]$ can be developed, then KS-DFT provides in many instances a reasonably accurate method for computing many-body ground state properties.

By variationally minimizing the KS total energy functional, subject to the constraint of orthonormality of the KS orbitals,

$$\langle \varphi_i | \varphi_j \rangle = \delta_{ij} \tag{3.19}$$

one arrives at a Schrödinger-like equation

$$\left(\widehat{H}_{KS} - \varepsilon_i\right)\varphi_i = 0 \tag{3.20}$$

where the KS eigenvalues ε_i are introduced as Lagrange multipliers to enforce the orthonormality constraints [78]. Equations (3.13), (3.14) and (3.20) are the well-known Kohn-Sham equations [78]. From these, the total energy and ground state density can be obtained once a self-consistent relationship between the density and potential is established. This involves iteratively solving the equations until the potential calculated from the density at a given iteration equals the potential used to construct the density via the next iteration of the KS equations. Within KS-DFT, this iterative or self-consistent process requires an evaluation of the XC energy. To this end it is useful to note that, in many cases, solids can be treated as close to the limit of a homogenous electron gas and within this limit, exchange and correlation effects can be considered local in character [78]. Hence, the XC energy can be written as an integral over space of the product of density and XC energy per electron $\varepsilon_{xc}[n]$,

$$E_{xc}[n] = \int d\boldsymbol{r}\, \varepsilon_{xc}[n] n(\boldsymbol{r}) \tag{3.21}$$

where $\varepsilon_{xc}[n]$ equals the XC energy per electron of a homogenous electron gas with density n; at every point inside the solid, $\varepsilon_{xc}[n]$ is computed using the local density around that point. This is the local density approximation (LDA) for the XC functional. Ceperley and Alder computed $\varepsilon_{xc}[n]$ for a range of densities in the homogenous electron gas using Quantum Monte Carlo techniques [84]. As a first approximation, LDA reproduces many measureable chemical properties accurately, especially if variations in the density are small compared to the Fermi wave vector $k_f = (3\pi n)^{1/3}$ [85],

$$\frac{|\nabla n(\boldsymbol{r})|}{k_f n(\boldsymbol{r})} \ll 1\,. \tag{3.22}$$

3.3.3 Geometry Optimization

A theorem derived independently by Hellman and Feynman shows that the force on a nucleus can be specified entirely in terms of its interaction with the surrounding charge density [86]. It can be shown [78] that this force theorem also holds in density functional calculations. Once the self-consistent total energy is evaluated, the forces on the nuclei can be computed from the derivative of the total energy with respect to nuclear coordinates [78]. The equilibrium or "optimized" geometry is thus the set of nuclear positions $\{\boldsymbol{R}_{ion}\}$ that minimize these forces. In practice, a force may be computed once a self-consistent total energy is obtained, and this force can then be used to update the nuclear positions, which in turn yields a new KS self-consistency problem to be solved iteratively. Thus geometry optimization can be considered as an "outer loop" to the self-consistent solution of the KS equations.

3.3.4 The DFT Self-consistency Loop

The KS equations discussed in section 3.3.2 are solved subject to the condition of consistency between the KS potential V_{KS} (equation (3.14)) and electron density (equation (3.16)). The iterative solution to the KS equations, as well as the outer geometry optimization loop, is shown schematically in figure 3.1. An initial approximation to the charge density initializes the loop; this consists of a linear combination of atomic orbitals (LCAO) approximating the KS orbitals which yields an electronic charge density. This is then used to calculate a new KS potential. The KS potential is inserted into a KS Hamiltonian which yields a Schrodinger-like KS equation. Matrix diagonalization in the LCAO basis then leads to the solution of the KS equation, yielding a new set of KS orbitals which can be used to construct a new charge density. The input and output densities are "mixed" through a linear combination which assists in convergence, and this generates a new input charge density [87, 88] for the next iteration. This process is iterated until the input and output densities of a given iteration differ by less than a user-specified threshold, at which point the system is considered "self-consistent". The total energy can be calculated at self-consistency and hence the ionic forces can be determined, which again can be compared to a force-threshold. After calculating the force, the ionic coordinates are updated, typically using variants of the steepest descent method [78]. Once the force is minimized below the specified threshold, desired quantities such as the bandstructure and density of states can be obtained at the relaxed geometry.

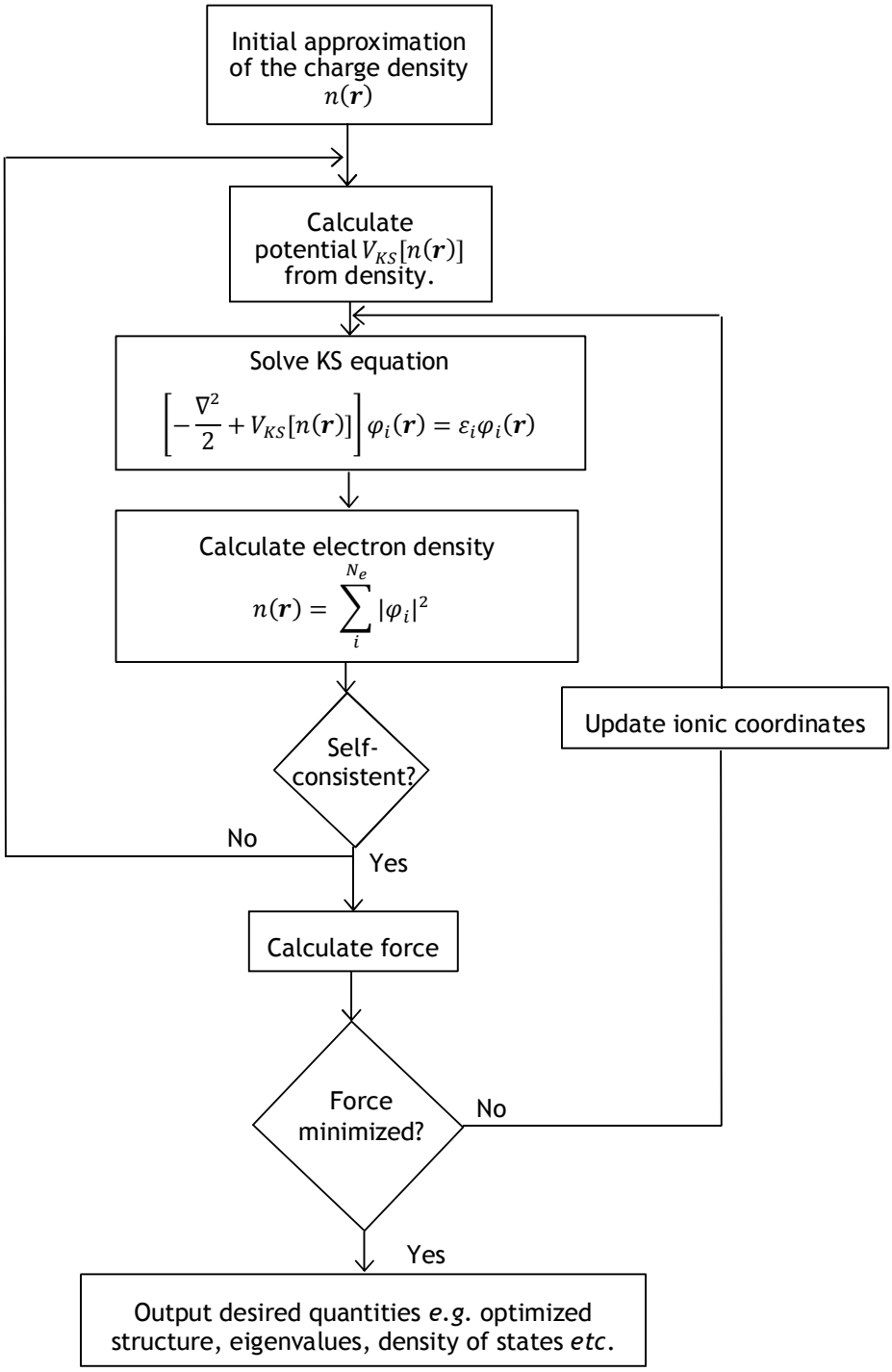


Figure 3.1 - Schematic representation of the iterative solution of the KS equations, yielding a self-consistent electronic structure, which in turn is used to calculate the ionic forces. The procedure is iterated until a minimum energy configuration or "optimized geometry" is obtained.

3.3.5 Limitations to Approximate DFT

Approximate DFT suffers from a number of weaknesses which lead to systematic errors in the computation of physical properties. There are two main sources of these limitations; self-interaction of electrons and the derivative discontinuity in XC functionals. Unlike in Hartree-Fock theory, within DFT the cancellation between the unphysical self-interaction term in the Hartree potential and the exchange interaction is only approximate. What remains is a spurious self-interaction term which causes the DFT eigenvalue magnitudes to be too large. Self-interaction-correction (SIC) can be incorporated into the LDA XC functional, for example through the parameterization by Perdew and Zunger [89], which leads to improvements. This correction alone, however, does not fully solve the problems of DFT and remains an *ad hoc* procedure.

The derivative discontinuity in the XC functional upon the addition or removal of an electron (Δ) can be defined as the difference between the exact quasiparticle (see section 3.4) bandgap E_{gap} and the bandgap obtained from KS eigenvalues E_{gap}^{KS} [90, 91].

$$\Delta = E_{gap} - E_{gap}^{KS} . \tag{3.23}$$

In other words, interpreting KS eigenvalues as physical excitation energies is not formally justified, apart from the highest occupied orbital which can be shown to correspond to the ionization energy [91]. These shortcomings lead to consistent underestimation of bandgaps in semiconductors such as Si, GaAs, and Ge; the errors are even more severe for more complicated systems such as transition-metal oxides. Thus, for an accurate treatment of *sp*-bonded systems such as the III-V semiconductors and CNTs studied in this work, an improvement over the electronic structure predicted by KS-DFT is required. Nevertheless, LDA-DFT often provides a good starting point for more accurate methods as the electronic structure obtained from DFT can be used as the input to more accurate methods.

The GW approximation to the self-energy, derived systematically from many-body perturbation theory, constitutes such a method and has been applied to various systems such as transition metals and semiconductors [51] with results in substantially better agreement with experiment than approximate DFT methods.

3.3.6 Application

3.3.6.1 The Supercell Method

The most popular approaches to DFT calculations applied to solids, surfaces and interfaces utilize the supercell method [2, 92], by which translational symmetry is imposed on a unit cell of a crystal in 3 dimensions. In the absence of defects and vacuum, the supercell approach allows one to predict properties of an infinite bulk crystal, yielding experimentally testable predictions for pure crystal properties. When a defect is introduced the calculation remains periodic, but the periodicity no longer corresponds to the pristine crystal, possibly resulting in electronic states in the fundamental energy gap for the case of insulators or semiconductors. However, due to the periodic boundary conditions (PBC), sufficiently large simulation cells are required to avoid defect-defect interactions as required to understand the properties of isolated defects.

3.3.6.2 Non-periodic systems

The most obvious objection to the use of PBC is that Bloch's theorem cannot be applied to a system that breaks translational symmetry, for example due to a defect or a surface. Bloch's theorem, and the advantages thereof, can only be utilized if a periodic supercell is used. Point defects and surfaces are therefore still modeled computationally using the supercell method, as long as the periodic images are spaced sufficiently far apart that their interactions can be neglected. For surfaces, the amount of vacuum introduced into a supercell should be chosen

to ensure that slab faces do not interact. Key electronic properties such as the valence and conduction band edges should not change as a function of vacuum spacing. For point defects, interactions between periodic images should be minimized which can be ensured if the atomic geometries (bond lengths and bond angles) located between the periodic defects images maintain their bulk values. The supercells for these 2 cases are shown schematically in figure 3.5. The interaction between charged defects is discussed in section 3.6.

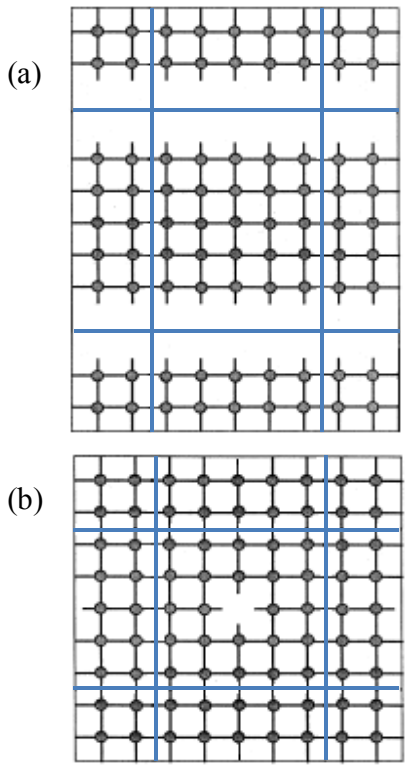


Figure 3.5 - Schematic representation of a slab surface (a), and a vacancy defect in a bulk solid (b), modeled in the supercell approach. The blue lines depict the boundaries of the supercell. Modified from ref. [92].

3.3.6.3 Plane waves versus localized orbitals

In this thesis, both plane wave basis sets and localized numerical atomic orbitals (NAO) are applied to study different problems. The advantages and disadvantages of their use reflect computational demands of the specific application in this work. Localized orbitals provide a chemically intuitive picture of the electronic states within a material, which is reflected in their widespread use in computational chemistry. The localized nature of the orbitals results in a finite interaction range, yielding relatively sparse matrices which can be utilized to generate "low-order" scaling approaches in which the computational cost scales better than the cube of the number of basis functions [93]. The locality of NAOs also means they are appropriate for Green's function based methods which treat the response of extended electronic states to a localized perturbation; such is the case when evaluating the transport properties of disordered CNTs. The requirement of non-interaction of the "leads" across the "scattering region" (see section 3.5) is easily satisfied by localized orbitals if a sufficiently large scattering region is specified. The desirable computational scaling of NAOs and their compatibility with the transport formalism are the primary reasons for choosing NAOs for the study of CNT vacancies. We wish to study electronic transport and formation energies of CNTs hosting vacancies in supercells consisting of up to 840 atoms (see chapter 6). Surprisingly, reasonable accuracy is maintained even with a minimal NAO basis set size. This is shown in chapter 6 by a comparison between transmission spectra obtained for CNTs using single- and double-ζ, in agreement with a comparable study by Ohfuchi *et al* [94]. Localized orbitals are also well suited for systems containing a network of sp^2-hybridized carbon atoms [78].

Localized basis sets are not without their disadvantages, however. They are dependent on the chemical environment; while NAO optimization [95] can lead to high accuracy with a relatively small number of basis functions for a given system, the resulting NAO then becomes highly tailored to that specific problem, limiting the transferability compared to the highly transferable,

spatially unbiased nature of plane waves. Plane wave basis sets are thus well suited to studying the variation of material properties between, for example, the bulk crystal and an oxide-passivated surface [78]. They are also especially suited to periodic supercells as they allow for the easy implementation of simple and efficient algorithms based on Fast Fourier Transforms (FFT). Finally, the systematic improvement in accuracy as a function of plane wave basis set size is a big advantage over localized orbitals; the latter can suffer from inconvenient effects such as over-completeness, in which accuracy of the resulting electronic structure is *not* a monotonic function of basis set size (see [96], where this effect is observed in BCC iron).

3.3.6.4 Pseudopotentials

In order to account for the rapid oscillations of the valence electron wavefunctions inside the core region, and also to expand the wavefunctions of the tightly bound core electrons, an extremely large basis set would be required. This makes all-electron calculations very computationally inconvenient, and often intractable. It is possible, however, to circumvent this problem by exploiting the fact that most of the chemically relevant physical properties of a material are primarily dependent upon the outer valence electrons and are not strongly correlated with core electrons. The pseudopotential approximation [97] removes the core electrons and replaces them and the ionic core potential with a pseudopotential that acts on pseudo-wavefunctions rather than real valence wave functions. Pseudopotentials are generated using an all-electron atomic radial Schrödinger equation. They are constructed so that the pseudo-wavefunctions have no radial nodes inside the core region, but outside a certain "cutoff radius" ($\boldsymbol{r_c}$ in figure below) the pseudo-wavefunctions and all-electron wavefunctions (and eigenvalues thereof) are equal. Increasing $\boldsymbol{r_c}$ leads to less transferable, but also more rapidly convergent, pseudopotentials.

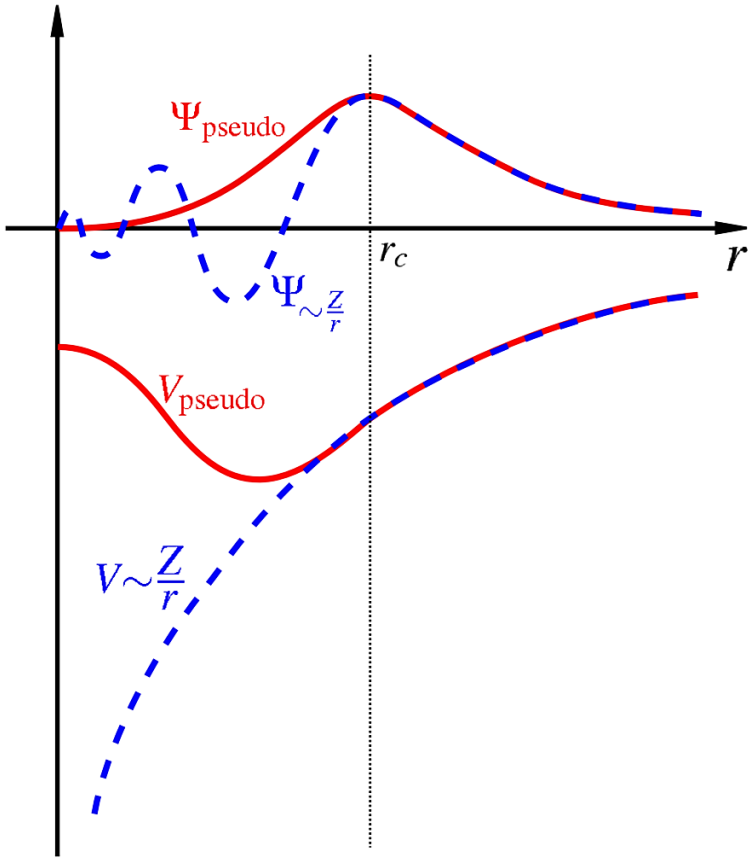


Figure 3.6 - Comparison of an all-electron potential and wavefunction (blue dashed curves) with a pseudopotential and pseudo-wavefunction (red solid curves). Note the lack of nodes of the pseudo-wavefunction inside r_c, and the equivalence of the all-electron and pseudo-quantities outside r_c. After ref. [98].

In this work, norm-conserving LDA pseudopotentials are used for all calculations. The LDA form of the XC functional proposed by Perdew and Zunger [89] which contains partial self-interaction correction as implemented in the Quantum Espresso [99] code is used for the InGaAs studies. The parameterization of Ceperley and Alder [84] is used for the study of disordered CNTs within the localized orbital periodic DFT code, OpenMX [76].

3.4 The *GW* Method

3.4.1 Photoemission Spectroscopy

Before considering the improvement of DFT-calculated properties via many-body perturbation theory, it is instructive to discuss the study of electron energy levels as determined by photoemission (or inverse photoemission) experiments, and in particular their relation to the concept of quasiparticles. In any spectroscopic experiment, the sample is perturbed and the response to the perturbation is measured. In a photoemission experiment, the system absorbs a photon of energy hf and an electron with kinetic energy E_k is ejected. Under the assumption that the ejected electron is decoupled from the system, energy conservation can be invoked to evaluate the energy level of the hole, the level that *was* occupied by the ejected electron. Since hf and E_k are measured, photoemission experiments infer the occupied density of states, corresponding to the valence bandstructure. Taking the highest occupied state as an example, its energy can be inferred from $E_k = hf - E_{HOMO}$ where E_{HOMO} is the energy of the highest occupied orbital, or the top of the valence band in an insulator. For inverse photoemission experiments, an electron is injected from vacuum, falls into an unoccupied state, and subsequently a photon of energy equal to the kinetic energy E_k of the absorbed electron minus the unoccupied state energy relative to the vacuum level E_{vac} is emitted, thus yielding information on the unoccupied density of states or conduction band.

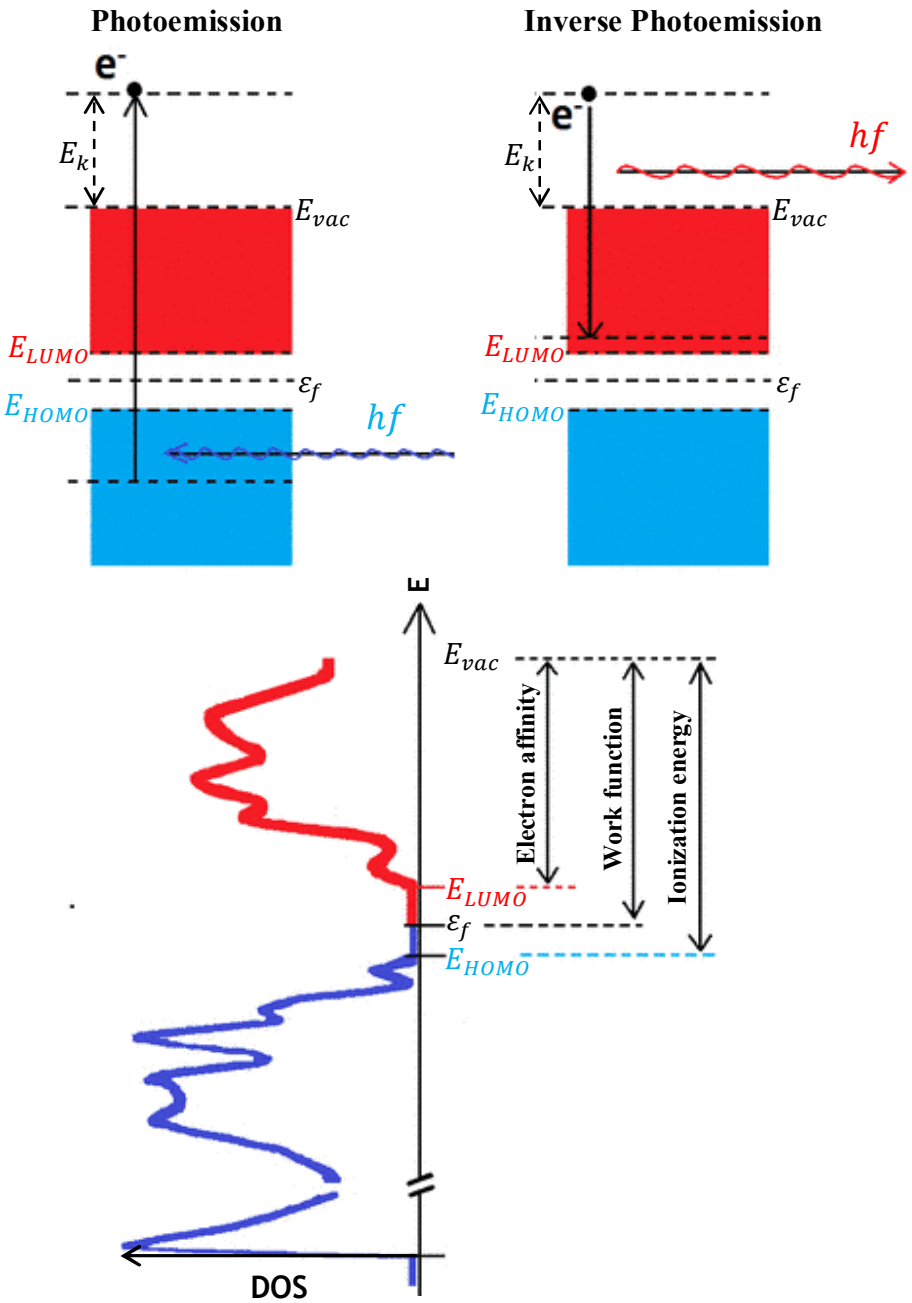


Figure 3.2 - Illustration of the photoemission and inverse photoemission processes, and the corresponding density of states probed by each experiment. Occupied (unoccupied) bands are represented in blue (red). The Fermi level is denoted by ε_f. This figure also serves as a graphical definition of the electron affinity and ionization energies, the former being the energy required to bring an electron from the vacuum level to the lowest unoccupied orbital (LUMO), the latter being the energy required to take an electron from the highest occupied orbital (HOMO) and bring it to the vacuum level. Modifed from ref. [100].

In photoemission experiments, the electronic energy levels become affected by the presence of an extra electron or hole which leads to a reorganization of charges and new electronic wavefunctions. This is particularly important when using energy differences to calculate addition and removal energies [101]. Thus, for an accurate theoretical account of such experiments, the approximation of independent electrons needs to be replaced by quasiparticles which occupy levels that contain the effects of other particles. The Coulomb repulsion between electrons results in a "cloud" of negative charge depletion surrounding a given electron, and the ensemble consisting of that electron and the positive charge screening its interactions with other particles forms the quasiparticle. The evolution of this quasiparticle can in many cases be described by a one-particle Schrödinger equation [102] but the Hamiltonian requires an effective potential that deals with the complicated interactions of the additional electron or hole with the rest of the system.

3.4.2 Quasiparticles and Green's Functions

Quasiparticles that interact weakly can often be used to describe the excitations of a system of strongly interacting particles [102]. As mentioned above, a bare electron inside a solid repels other electrons and becomes surrounded by a positively charged polarization cloud. In this instance, quasiparticles interact via a screened rather than the bare Coulomb potential. As quasiparticles are approximate eigenstates of the Hamiltonian, the quasiparticle lifetime is finite. The quasiparticle lifetime is inversely proportional to the imaginary part of a complex energy which arises due to the residual interaction between quasiparticles. This complex energy is the non-local, energy dependent, non-Hermitian self-energy Σ. This in principle accounts for all exchange and correlation effects and describes the energy difference between the quasiparticle and the bare particle [102, 103]. The quasiparticle equation governs the behaviour of quasiparticles and disregarding spin degrees of freedom can be written as

$$\left[-\frac{1}{2}\nabla^2 + V_{Hartree} + V_{ext}\right]\Phi_i(\boldsymbol{r}) + \int d\boldsymbol{r}'\,\Sigma(\boldsymbol{r},\boldsymbol{r}';E_i)\Phi_i(\boldsymbol{r}') = \Phi_i E_i\,. \quad (3.24)$$

All terms inside the square brackets have been defined in the previous section, Φ_i is the quasiparticle wavefunction, and E_i is the quasiparticle energy. The propagation of quasiparticles through space and time is described by the single particle Green's function G. From G one can obtain the quasiparticle excitation spectrum, life times, and expectation values of single particle operators in the ground state, such as the electronic density. Let $|N_0\rangle$ be the ground state of the N-electron Hamiltonian, $\widehat{\Theta}(\boldsymbol{r}\,t) = e^{i\hat{H}t}\widehat{\Theta}(\boldsymbol{r})e^{-i\hat{H}t}$ the Fermion annihilation operator in the Heisenberg representation [102] ($\widehat{\Theta}(\boldsymbol{r}\,t)$ destroys an electron at point $\boldsymbol{r}$ and time t), $\Theta^{\dagger}(\boldsymbol{r}\,t)$ the Fermion creation operator, and $\hat{\mathcal{T}}$ the time ordering operator; G then reads [103]

$$\begin{aligned} G(\boldsymbol{r}\,t,\boldsymbol{r}'\,t') &= -i\langle N_0|\hat{\mathcal{T}}\left[\Theta(\boldsymbol{r}\,t),\widehat{\Theta}^{\dagger}(\boldsymbol{r}'\,t')\right]|N_0\rangle \\ &= \begin{cases} -i\langle N_0|\widehat{\Theta}(\boldsymbol{r}\,t)\widehat{\Theta}^{\dagger}(\boldsymbol{r}'\,t')|N_0\rangle, & t > t' \\ i\langle N_0|\widehat{\Theta}^{\dagger}(\boldsymbol{r}'\,t')\widehat{\Theta}(\boldsymbol{r}\,t)|N_0\rangle, & t' > t \end{cases} \end{aligned} \quad (3.25)$$

If $t > t'$ ($t' > t$), G describes the propagation of an electron (hole) added to N-electron system described by $\widehat{H}$, *i.e.* the $N \rightarrow N$+1 ($N \rightarrow N$-1) electron addition (removal) process. A direct connection to experiment can be made by taking the imaginary part of G,

$$A = \pi^{-1}|\mathrm{Im}G(\boldsymbol{r},\boldsymbol{r}';E)|. \quad (3.26)$$

This is the single-particle spectral function which is related to the photoemission spectrum. The quasiparticle energy and inverse lifetime are determined, respectively, by the position and width of peaks in the interacting spectral function to be defined below. The area under the peak evaluates the quasiparticle weight. The physical significance of the Green's function becomes more apparent when it is expressed in terms of the quasiparticle wavefunctions and

energies. The quasiparticle wavefunctions satisfy a completeness relation [103] such that it is possible to introduce a complete set of eigenstates into the definition of G given above. Fourier transforming to energy space [102] then leads to the spectral representation of the Green's function, of which quasiparticles can be identified as poles

$$G(\boldsymbol{r},\boldsymbol{r}';E)=\sum_{i}\frac{\Phi_i(\boldsymbol{r})\Phi_i{}^{*}(\boldsymbol{r}')}{E-E_i}=\oint dE'\frac{A(\boldsymbol{r},\boldsymbol{r}';E')}{E-E'}\ . \qquad (3.27)$$

In practice an infinitesimal imaginary term is added to the denominator of the summation, and the contour integral runs infinitesimally below (above) the real energy axis for $E'\geq\varepsilon_f$ $(E'<\varepsilon_f)$, where ε_f is the Fermi level. The integral defines the interacting spectral function $A(\boldsymbol{r},\boldsymbol{r}';E)=\sum_i\Phi_i(\boldsymbol{r})\Phi_i{}^{*}(\boldsymbol{r}')\delta(E-E')$. As a starting point one can choose a suitable independent particle Hamiltonian, whose eigenfunctions are close to the quasiparticle wavefunctions, which allows for the evaluation of quasiparticle energies via first-order perturbation theory. The Kohn-Sham formalism yields such a Hamiltonian $\widehat{H}_{KS}=-\frac{\nabla^2}{2}+V_{Hartree}+V_{ext}+V_{XC}$ (see section 3.3.2) with KS orbitals φ_i as the corresponding independent particle eigenfunctions and KS eigenvalues ε_i. In this case, the independent particle Green's functions G_0 and the independent particle spectral function A_0 can then be constructed via

$$G_0(\boldsymbol{r},\boldsymbol{r}';E)=\sum_{\boldsymbol{i}}\frac{\varphi_i(\boldsymbol{r})\varphi^{*}{}_i(\boldsymbol{r}')}{E-\varepsilon_i}=\oint dE'\frac{A_0(\boldsymbol{r},\boldsymbol{r}';E')}{E-E'}\ . \qquad (3.28)$$

G_0 and the self-energy Σ can be used to construct the interacting Green's function from the Dyson equation [103]

$$\begin{aligned}&G(\boldsymbol{r},\boldsymbol{r}';E)\\&=G_0(\boldsymbol{r},\boldsymbol{r}';E)\ +\int d\boldsymbol{r}_1\,\boldsymbol{r}_2 G_0(\boldsymbol{r},\boldsymbol{r}_1;E)\Sigma(\boldsymbol{r}_1,\boldsymbol{r}_2;E)G(\boldsymbol{r_2},\boldsymbol{r}';E)\,. \qquad (3.29)\end{aligned}$$

Quasiparticle energies can be determined directly from the quasiparticle equation. The formal resemblance between equation (3.24) and the KS equation (equation (3.20)) provides a good reason for suggesting the use of the KS equation as a starting point for more accurate calculations beyond LDA-DFT [101]. Unlike KS DFT, however, quasiparticle wavefunctions and energies are actually physical and not just mathematical objects. Assuming $\langle\varphi_i|\Phi_i\rangle \cong 1$ (allowing for first-order perturbation theory) and letting V_{XC} equal the KS exchange-correlation potential, the quasiparticle energies can then be determined from [103]

$$E_i \cong \varepsilon_i + Z\langle\varphi_i|\Sigma(\varepsilon_i) - V_{XC}|\varphi_i\rangle\,, \tag{3.30}$$

where the quasiparticle weight is [103]

$$Z = \left[1 - \frac{\partial\Sigma_i(E)}{\partial E}\bigg|_{E=\,\varepsilon_i}\right]^{-1}, \tag{3.31}$$

and

$$\Sigma_i(E) = \langle\varphi_i|\Sigma(E)|\varphi_i\rangle \tag{3.32}$$

Most practical implementations compute quasiparticle energies using these equations. At this point however, it becomes necessary to find reasonable, in term of accuracy and computability, approximations to the self-energy.

3.4.3 The *GW* Approximation of the Self-energy

Hedin's equations [104] are a set of integro-differential equations which, by connecting the interacting Green's function G, the self-energy Σ, the polarizability P, and the vertex function, allows for the introduction of a dynamically screened interaction W to provide a formalism for the evaluation of

the self-energy when combined with the Dyson equation [103]. The determination of Σ requires iteration of Hedin's equations until G and W are self-consistent. This however, restricts the system size accessible due to computational expense. An approximation to the self-energy which balances accuracy with tractability is thus desired.

The GW approximation results from setting the vertex function arising in Hedin's equations to unity and is defined by the following equations [103]

$$\Sigma(\mathbf{1},\mathbf{2}) = iG(\mathbf{1},\mathbf{2})W(\mathbf{1}^+,\mathbf{2}), \tag{3.33}$$

$$W(\mathbf{1},\mathbf{2}) = \mathrm{v}(\mathbf{1},\mathbf{2}) + \int d\mathbf{3}d\mathbf{4}\, W(\mathbf{1},\mathbf{3})P(\mathbf{3},\mathbf{4})\mathrm{v}(\mathbf{4},\mathbf{2}), \tag{3.34}$$

$$P(\mathbf{1},\mathbf{2}) = -iG(\mathbf{1},\mathbf{2})G(\mathbf{2},\mathbf{1}). \tag{3.35}$$

Here, a compact form of the space-time arguments is used *e.g.* $\mathbf{1}$ represents $\boldsymbol{r}_1, t_1$, and $\mathbf{1}^+ = \mathbf{1} + i0^+$, $0^+ > 0$ is an infinitesimal. The bare Coulomb interaction is denoted by v. The energy dependence of the self-energy comes from W which gives rise to energy dependent correlation effects [101, 103]. A description of W in terms of the dielectric matrix and polarizability is now given.

Density fluctuations and many-body effects lead to screening in solids or local-field effects which can be accounted for by the inverse dielectric matrix ε^{-1}. The screened interaction W is related to $\varepsilon^{-1}(\boldsymbol{r},\boldsymbol{r}';E)$ as

$$W(\boldsymbol{r},\boldsymbol{r}';E) = \int d\boldsymbol{r}''\, \varepsilon^{-1}(\boldsymbol{r},\boldsymbol{r}'';E)\mathrm{v}(\boldsymbol{r}'',\boldsymbol{r}')\,, \tag{3.36}$$

where E represents energy (the conjugate variable obtained after Fourier transforming from the time domain). The off-diagonal elements of the reciprocal space representation of the dielectric matrix $\varepsilon_{\mathbf{GG}'}(\boldsymbol{q};E)$ account for the screening due to local fields, *i.e.* screening in the presence of an inhomogeneous density distribution [103]. The dielectric matrix can be calculated from the

irreducible polarizability P by

$$\varepsilon(\boldsymbol{r},\boldsymbol{r}';E)=\delta(\boldsymbol{r}-\boldsymbol{r}')-\int d\boldsymbol{r}''\,\mathrm{v}(\boldsymbol{r},\boldsymbol{r}'')P(\boldsymbol{r}'',\boldsymbol{r}';E)\,. \tag{3.37}$$

The random phase approximation allows for a determination of the polarizability from independent-particle wavefunctions and eigenvalues [102, 103]

$$P(\boldsymbol{r},\boldsymbol{r}';\,E)=\sum_{a,a'}\varphi_a^*(\boldsymbol{r})\varphi_{a'}(\boldsymbol{r})\varphi_{a'}^*(\boldsymbol{r}')\varphi_a(\boldsymbol{r}')\left[\frac{f_a(1-f_{a'})}{\varepsilon_a-\varepsilon_{a'}+E+i\Delta}+\frac{f_{a'}(1-f_a)}{\varepsilon_{a'}-\varepsilon_a-E+i\Delta}\right], \tag{3.38}$$

where φ_a and ε_a are respectively, the LDA-DFT eigenfunctions and eigenvalues as in section 3.3, and f_a is the Fermi distribution function determining the occupation of state a. The thrust of this section is that knowledge of the LDA spectrum can be used to evaluate the self-energy; by using the φ_a and ε_a to obtain P, the dielectric matrix and in turn the screened interaction can be calculated, enabling the calculation of Σ [52, 103].

3.4.4 Qualitative Features of the Electronic Self-energy

In the GW approximation, the self-energy, which is a function of the single-particle Green's function and the inverse dielectric function, can be decomposed in two parts; the exchange (ex) and correlation ($corr$) contributions

$$\Sigma(E)=\Sigma^{ex}(E)+\Sigma^{corr}(E)\,. \tag{3.39}$$

The exchange contribution arises from the Coulomb interaction and anti-

symmetry of the many-electron wavefunction; it is essentially the Hartree-Fock exchange for electrons. The correlation contribution accounts for the effect of self-energy beyond the Hartree-Fock approximation. It can be expressed using the independent particle Green's function [103]. A more physically intuitive way of expressing the components of the self-energy is to adopt the Coulomb-hole plus screened exchange (COHSEX) approximation, in which the real part of the self-energy is divided into the two terms

$$Re\ \Sigma(\boldsymbol{r},\boldsymbol{r}';E) = \Sigma_{COH}(\boldsymbol{r},\boldsymbol{r}';E) + \Sigma_{SEX}(\boldsymbol{r},\boldsymbol{r}';E). \tag{3.40}$$

Here, the first term on the right hand side arises from the poles of the screened interaction; in analogy to the spectral representation of the Green's function, a spectral function can also be used to express W [103]. The second term arises from the poles of the Green's function. In the static limit of this approximation, the interpretation of each term becomes clear: Σ_{SEX} is simply the Hartree-Fock exchange in which the Coulomb term is replaced by a statically screened interaction (the energy argument of W is set to 0)

$$\Sigma_{SEX}(\boldsymbol{r},\boldsymbol{r}') = -\sum_{a}^{occ} \varphi_a(\boldsymbol{r})\varphi_a^*(\boldsymbol{r}')W(\boldsymbol{r},\boldsymbol{r}';0)\ , \tag{3.41}$$

and the Coulomb-hole term equals the interaction energy of the quasiparticle with the surrounding Coulomb-hole, where the latter arises from to the rearrangement of electrons due to the additional particle [51, 103]

$$\Sigma_{COH}(\boldsymbol{r},\boldsymbol{r}') = \frac{1}{2}\delta(\boldsymbol{r}-\boldsymbol{r}')[W(\boldsymbol{r},\boldsymbol{r}';0) - \mathrm{v}(\boldsymbol{r},\boldsymbol{r}')]\ . \tag{3.42}$$

While the static COHSEX approximation has well-known failures (*e.g.* 20% overestimation of the self-energy in Si [51, 103]) it serves as an intuitive illustration of the various contributions to the self-energy.

In order to describe the qualitative features of the self-energy, Hybertsen and Louie [105] plotted both the Coulomb hole term and screened exchange in the (110) plane of Si. Their results show that the screened interaction is anisotropic and dependent on charge inhomogeneity. The build-up of charge along a Si-Si bond means that an electron added to this region will undergo more screening of the Coulomb interaction than an electron added to an interstitial region, for example. As a result, valence electrons which are mostly localized to the bond experience a deeper potential well than the conduction electrons localized to the vicinity of interstitial regions, leading to an enhanced separation of the band edges. Thus local field effects act to increase the bandgap compared to LDA [103, 105].

3.4.5 Strengths and Limitations

The GW approximation, corresponding to the first iteration of Hedin's equations, has been shown to exhibit a number of weaknesses in certain systems; particularly systems with strongly interacting/correlated particles, such as 3*d* or 4*f* materials. For example, *s* to *d* orbital promotion energies for Mn, Fe, Co, Ni, and Cu are overestimated by up to ~3eV [103]. Also, nearly degenerate single and double core holes lead to a breakdown of the quasiparticle picture and in turn incorrect predictions of core-level spectra [106], hence the attribution of strongly correlated electron or hole pairs to the failure of the GW approximation. For alkali metals, satellites in the photoemission spectra are also poorly reproduced by the GW approximation; since W represents the coupling of electrons to single plasmon excitations, the multiple plasmon excitations of alkali metals surpass the range of applicability of the GW approximation [51]. Spin interactions in transition metals are also problematic since there is no spin dependence in W as there is only purely Coulombic screening.

Despite these shortcomings, the GW approximation systematically improves the DFT eigenvalues of many bulk and surface models of semiconductors and insulators to which it has been applied [51], [103]. Much of

this improvement over DFT can be attributed to the anisotropic screening discussed in the previous section; local fields result in a self-energy that varies in strength from point to point in the unit cell, which in turn leads to more accurate relative positions of the band edges [51, 105]. In summary, the self-consistent Kohn-Sham orbitals and eigenvalues can be used as the starting point to build the Green's function, the dielectric response, the screened interaction, which yields the self-energy within the GW approximation and ultimately the quasiparticle eigenvalue spectrum. These *a posteriori* implementations of the GW method have been shown to be highly successful in terms of agreement with experiment for "prototypical" semiconductors and insulators including but not limited to bulk Si, Ge, GaAs, and diamond [103], [51] as well as their surfaces [22, 52].

3.5 Electronic Transport from the Green's Function Method

The scattering properties of carbon nanotubes (CNTs) are investigated in this thesis by evaluating electron transport in the presence of carbon vacancies. The electron transmission probability through a central scattering region, which is in contact with semi-infinite electron reservoirs or "leads" on either side is evaluated using the Green's function method applied to systems near equilibrium. The periodicity of the leads allows for the description of electrons far from the scattering region in terms of Bloch wavefunctions propagating along the lead axis or transport direction. For directions perpendicular to the lead axis, confinement results in a quantization of the electron momentum such that only a discrete number of channels are available for conduction; the number of channels is proportional to the width of the confinement, *i.e.* the dimensions of the leads perpendicular to the transport direction. Each channel has a probability T of scattering into the lead on the opposite side of the central region. The

transmission spectrum $T(E)$ is calculated as a trace over a matrix product involving a Green's function. In this context, the Green's Function essentially plays the role of the scattering matrix [24], which relates the incoming and outgoing electron states.

3.5.1 The Scattering Matrix

Consider the case of 1-dimensional leads each containing 2 transverse channels or modes; electrons, represented by extended plane waves and originating from each lead, encounter the non-periodic scattering region including the conductor which contains the defect or scatterer and are thus scattered with a finite probability of both being reflected back into the respective lead and transmitted through the central region to the lead on the other side [24]. A schematic of this process is shown below.

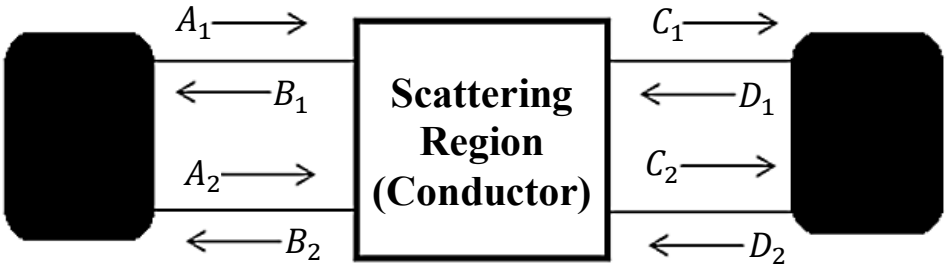


Figure 3.3 - Scattering of electron waves by a two conductance channel region connected to leads on either side; the leads act as electron reservoirs. $\boldsymbol{A}$ ($\boldsymbol{B}$) and $\boldsymbol{D}$ ($\boldsymbol{C}$) correspond to amplitudes of incoming (outgoing) waves. Modified from ref. [24].

The relation between amplitudes of incoming and outgoing modes can be described by the scattering matrix S. If r and t correspond to reflection and transmission coefficients of a mode incoming from the left lead, then S can be defined as

$$S = \begin{pmatrix} r & t' \\ t & r' \end{pmatrix}, \tag{3.43}$$

where r (r') and t (t') are $N \times N$ matrices (where N equals the number of modes in each lead, $N = 2$ in figure 3.3) corresponding to the reflection and transmission amplitudes of an incoming mode from the left (right) lead. For a more detailed description of the scattering matrix, see ref. [24]. The left and right lead modes $\varphi_n^{left,right}$ are linear combinations of normalized plane waves multiplied by a column vector χ_n which describes the n^{th} mode's sub-bands [107]

$$\varphi_n^{left}(x) = \chi_n \left[\frac{A_n}{\sqrt{v_n}} \exp(ik_n x) + \frac{B_n}{\sqrt{v_n}} \exp(-ik_n x) \right], \tag{3.44}$$

$$\varphi_n^{right}(x) = \chi_n \left[\frac{C_n}{\sqrt{v_n}} \exp(ik_n x) + \frac{D_n}{\sqrt{v_n}} \exp(-ik_n x) \right] \tag{3.45}$$

where v_n, k_n, x, and are, respectively, the group velocity [24], wave vector, and position along the lead axis. For a given mode, A (B) and D (C) correspond to amplitudes of incoming (outgoing) waves, as in figure 3.3. Normalization of the plane waves ensures unitarity of the scattering matrix *i.e.* $S^{\dagger}S = 1$, which in turn ensures current conservation and satisfaction of the sum rule for conductance when describing conductance in terms of transmission [24]. By bearing in mind the definition of the scattering matrix and the lead wavefunctions, the relation between outgoing and incoming mode amplitudes can be expressed as [107]

$$|Out\rangle = S|In\rangle \rightarrow \begin{pmatrix} B_1 \\ \vdots \\ B_N \\ C_1 \\ \vdots \\ C_N \end{pmatrix} = S \begin{pmatrix} A_1 \\ \vdots \\ A_N \\ D_1 \\ \vdots \\ D_N \end{pmatrix}. \tag{3.46}$$

The scattering matrix can be determined from the Green's function once a relationship between the Green's function and wavefunctions is obtained [107]. Transmission and reflection coefficients can then be calculated, which in turn

allows for the calculation of the transmission spectrum as

$$T(E) = trace(t^{\dagger}t)\,. \tag{3.47}$$

3.5.2 Calculation of the Green's Function

A schematic of the system geometry studied using the Green's function method is shown below.

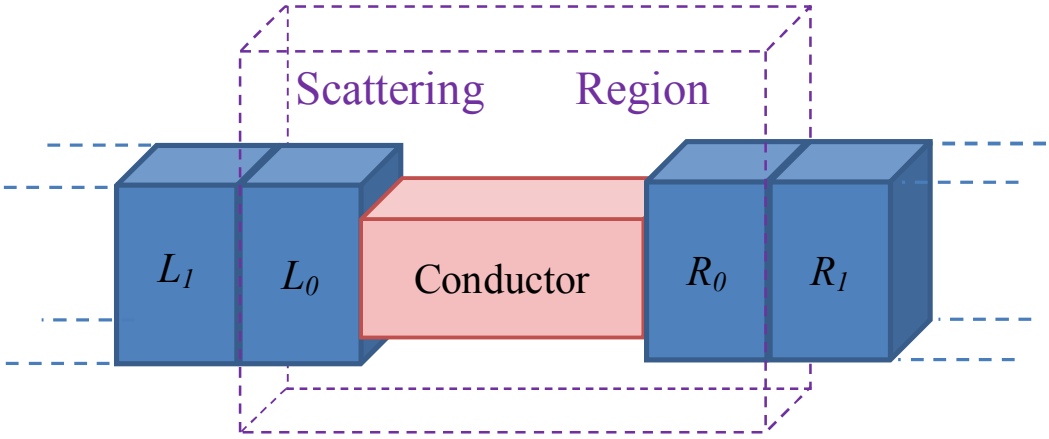


Figure 3.4 - Schematic illustration of the system geometry of which the transmission spectrum is calculated using Green's functions. The left (L) and right (R) leads are repeated periodically in their respective directions as indicated by the dashed lines. The scattering region consists of the central region containing the scatterer (conductor) and a copy of the principal layer of each lead.

Two conditions are imposed when calculating the Green's function for such a system; each lead repeat cell or principal layer interacts only with nearest-neighbor cells, and the conductor region interacts only with the L_0 and R_0 principal layers. The employment of localized basis functions allows for adherence to these conditions if sufficiently large lead principal layers are chosen. Once these conditions are satisfied, the system Hamiltonian H takes on a block tridiagonal form (the main diagonal contains the $N \times N$ Hamiltonians of each region, the upper and lower diagonals account for their interactions; N is the number of basis functions) and the entire system's Green's function can be

expressed as

$$G(E) = (E\mathcal{O} - H)^{-1}\,. \tag{3.48}$$

Here, E is a complex energy and $\mathcal{O}$ is the overlap matrix. Figure 3.4 depicts an open system connected to leads that extend to infinity in their respective directions, and therefore calculating G involves inversion of an infinite matrix [24]. What is needed is a way to truncate the matrix without ending up with a closed system with fully reflecting boundaries. To this end, one may express the Green's function of the scattering region G_S using an effective energy-dependant term which arises from the interaction of the leads with the conductor; this is the self-energy due to the leads $\Sigma_{(L,R)}$. This self-energy describes the interaction of the scattering region with the semi-infinite leads, and should not be confused with the electron self-energy in section 3.4 which describes the single electron interactions beyond the mean field approximation. By expressing G_S in this way, the entire effect of the semi-infinite leads is taken into account without approximation [24, 76, 108],

$$G_S(E) = (E\mathcal{O}_S - H_S - \Sigma_L(E) - \Sigma_R(E))^{-1} \tag{3.49}$$

$$\Sigma_{(L,R)}(E) = \left(E\mathcal{O}_{S(L_1,R_1)} - H_{S(L_1,R_1)}\right)G_{(L,R)}(E)\left(E\mathcal{O}_{(L_1,R_1)S} - H_{(L_1,R_1)S}\right). \tag{3.50}$$

where alternating order of the subscripts for $\mathcal{O}$ and H implies Hermitian conjugation. S denotes the scattering region and $G_{(L,R)}$ are the lead surface Green's functions, calculated by direct inversion of the corresponding lead Hamiltonian and overlap matrices,

$$G_{(L,R)}(E) = \left(E\mathcal{O}_{LL,RR} - H_{LL,RR}\right)^{-1}. \tag{3.51}$$

With G_S at hand, the transmission spectrum $T(E)$ or equivalently the Landauer

conductance spectrum in units of $2e^2/h$ can be calculated from [24]

$$T(E) = Tr[\Gamma_L G_S(E)^\dagger \Gamma_R G_S(E)]\,. \tag{3.52}$$

The function $\Gamma_{L,R} = i\left[\Sigma_{(L,R)}(E) - \Sigma_{(L,R)}(E)^\dagger\right]$ describes the coupling of lead modes to electronic states inside the conductor.

3.5.3 Localization

Anderson localization refers to the lack of diffusion or "trapping" of waves when propagating through a disordered medium. As a result of interaction with randomly distributed defects, electronic wavefunctions are no longer extended as in a pristine lattice but may become localized and the material exhibits significant deviations from Ohm's law [109]. Under such conditions the system is said to be in the strong Anderson localization regime of electron transport. It is useful to explain how this relates to the diffusive and ballistic regimes. Each regime is distinguished from others in terms of the length of the conductor L relative to various characteristic lengths of a given material. All lengths considered here are greater than the de Broglie wavelength of electrons. The various transport regimes can be described as

- Ballistic – When L is less than the phase relaxation length L_θ (*i.e.* no phase-breaking processes such as electron-phonon or electron-electron interactions), and less than the elastic mean free path λ, *and* the localization length ξ, electrons may propagate ballistically through the conductor, without scattering. In this case the resistance reaches the lower quantum limit of $h/2e^2$ per channel and originates entirely from the interfaces between the conductor and the contacts, assuming a much larger number of transverse modes in the contacts [24].

- Diffusive – Ohmic behaviour is recovered when L is greater than L_θ, λ and ξ. Elastic and inelastic collisions dominate the transport process such that electronic momentum is randomized between scattering events. In this regime, transport can be described entirely in terms of a semi-classical model [110].

- Anderson localization regime – Strong Anderson localization dominates transport when $\xi < L < L_\theta$. In the presence of a large number of defects, electrons become localized to a region of the order of ξ as a result of interaction with the random potential of defects. The resultant resistance increases exponentially with L [109].

Beenakker has shown that in the absence of magnetic fields or spin-orbit coupling a linear relationship exists between the localization length and the mean free path [111]

$$\xi = (N_{ch} + 1)\lambda \ , \tag{3.53}$$

where N_{ch} is the number of channels. In section 3.8, an approximation which yields a linear relationship between the elastic mean free path and distance between defects is described, thus relating the localization length to the defect density for a 1-dimensional conductor. This follows from the assumption that scatterers in series do not interact, leading to an estimation of the elastic mean free path by considering only a single scatterer. This independent scattering approximation (ISA) will be used to study the strong Anderson localization regime in defective carbon nanotubes (chapter 6). As ξ exceeds 100 nm for typical defect densities in CNTs, which is evidenced by experiments and previous theoretical approaches [7, 11], studying the strong Anderson localization regime in a purely *ab-initio* framework seems at first intractable. However, the ISA provides a means to overcome this problem. To conclude this section, the problem of computing electrical conductance at nanometer scale

dimensions is reduced to a wave-scattering problem via utilization of the Green's function, which can be evaluated from the self-consistent Hamiltonian and overlap matrices obtained from the output of a (localized-orbital-based) DFT calculation. The implementation of these ideas for the study of defective CNTs is discussed in section 3.8.

3.6 Charge Transition Levels

Point defects in semiconductors typically give rise to states inside the fundamental gap of a material. Measuring these levels experimentally or measuring the levels which correspond to transitions between charge states of the defect center is a widely used method of defect characterization. Such charge transition levels can also be calculated from first-principles. Thus the combination of experimental and theoretical results allows for insight into the possible atomic structures of a defect.

The definition of a thermodynamic charge transition level $\varepsilon_{q/q'}$ is the Fermi level position where energetically competing charge states have equal formation energy. The formation energy of charge state q, relative to another charge state q', is given by

$$E_{form}(q/q') = E(q, R_q) - E(q', R_{q'}) + q\Delta\varepsilon_f \,, \qquad (3.54)$$

where $E(q, R_q)$ is the total energy of a supercell containing a single defect in charge state q with the supercell geometry relaxed to the configuration of the q^{th} charge state, denoted by R_q. In the final term $\Delta\varepsilon_f$ is the Fermi level position with respect to the valence-band maximum (VBM), and accounts for the transfer of electrons to and from a charge reservoir. The slope of $E_{form}(q/q')$ will be given by the coefficient of the $\Delta\varepsilon_f$ term, *i.e.* the final charge state in the $q' \rightarrow q$ addition/removal process as the Fermi level is swept from the valence to the conduction band. The linear dependence of E_{form} on the Fermi level can be used

to extract the charge transition level $\varepsilon_{q/q'}$; E_{form} can be plotted against $\Delta\varepsilon_f$ for different charge states, and the position of the Fermi level at which the E_{form} lines cross as a function of $\Delta\varepsilon_f$ corresponds to the charge transition level, $q' \rightarrow q$.

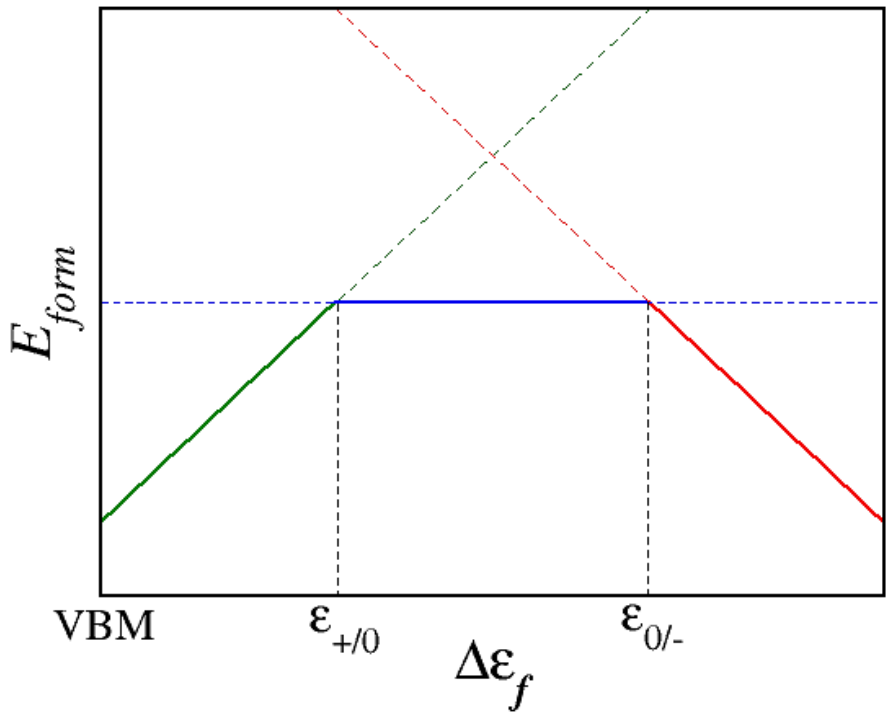


Figure 3.7 - Generic plot of formation energies (E_{form}) of positive (green line), neutral (blue line), and negative (red line) charge states of a given defect center, versus Fermi level position ($\Delta\varepsilon_f$). Charge transition levels correspond to values of the Fermi level at which the most stable charge state (*i.e.* the lowest formation energy) changes. Here, the positive to neutral ($\varepsilon_{+/0}$) and neutral to negative ($\varepsilon_{0/-}$) transitions are shown.

As their name suggests, charge transition levels involve calculations with different numbers of electrons occupying the defect level. Thus the well-known deficiencies of DFT, namely the lack of derivative discontinuity of the XC functional with respect to particle number, and an inappropriate treatment of the electron self-interaction are exacerbated when computing charge transition levels. For example, the underestimation of the fundamental gap which arises from the above-mentioned weaknesses, can lead to a qualitatively wrong picture

in which charge transition levels resonate with the host bands, whereas experiment, and more rigorous theoretical approaches would indicate they lie within the bandgap [44]. In order to overcome these problems, a new computational approach has been developed in recent years [22, 44], which combines many-body perturbation theory with DFT. This leads to a correction to the Kohn-Sham levels obtained from the electron self-energy calculated within the GW approximation.

To overcome the shortcomings of DFT, the formation of charged defects can be decomposed into 2 major contributions: the energetic cost of adding an electron, and the energy change arising from the relaxation of surrounding atoms upon addition of the electron. This is achieved by rewriting equation (3.54). Let $E(q, R_{q'})$ be the energy of a simulation cell with charge state q but with atomic positions optimized for charge state q'. Adding and subtracting $E(q, R_{q'})$ and grouping all terms appropriately yields

$$\begin{aligned} E_{form}(q/q') &= E(q, R_{q'}) - E(q', R_{q'}) + E(q, R_q) - E(q, R_{q'}) + q\Delta\varepsilon_f \\ &= A(q, q', R_{q'}) + \Delta(R_q, R_{q'}, q) + q\Delta\varepsilon_f \quad . \end{aligned} \tag{3.55}$$

The first term A represents the energy to add an electron, neglecting the rearrangement of surrounding atoms $E(q, R_{q'}) - E(q', R_{q'})$. The second term Δ arises from the relaxation of atoms in the presence of the extra charge, which in DFT translates to a re-minimization of the total energy $E(q, R_q) - E(q, R_{q'})$.

Notice that the charge state of the energy terms in Δ do not change; in charged supercells, a finite spurious interaction between periodic images exists that is typically accounted for by an electrostatic correction term in the formation energy. Makov and Payne proposed a correction for the spurious Coulomb interaction which takes the form of a Madelung-like energy with a third order term to account for the delocalized part of the defect charge distribution and its interaction with periodic images considered as a screened array of point-charges [112]. The correction is often obtained from an

extrapolation of the formation energy of the charged cell to infinite supercell volume [4], which requires a large computational effort. The fact that the Δ term in equation (3.55) contains an energy difference between cells in the same charge state suggests that much of the error due to spurious electrostatics may cancel in this term. In chapter 5 we investigate this suggestion by plotting the total energies $E(q, R_q)$ and $E(q, R_{q'})$ versus a linear expression of the cell volume. Our results indicate that these electrostatic interactions are practically negligible compared to elastic interactions between periodic images of defects, and the error due to elastic interactions is also very small in the cells we use to calculate the Δ term.

The representation of $E_{form}(q/q')$ given by equation (3.55) can be illustrated by use of configuration coordinate diagrams.

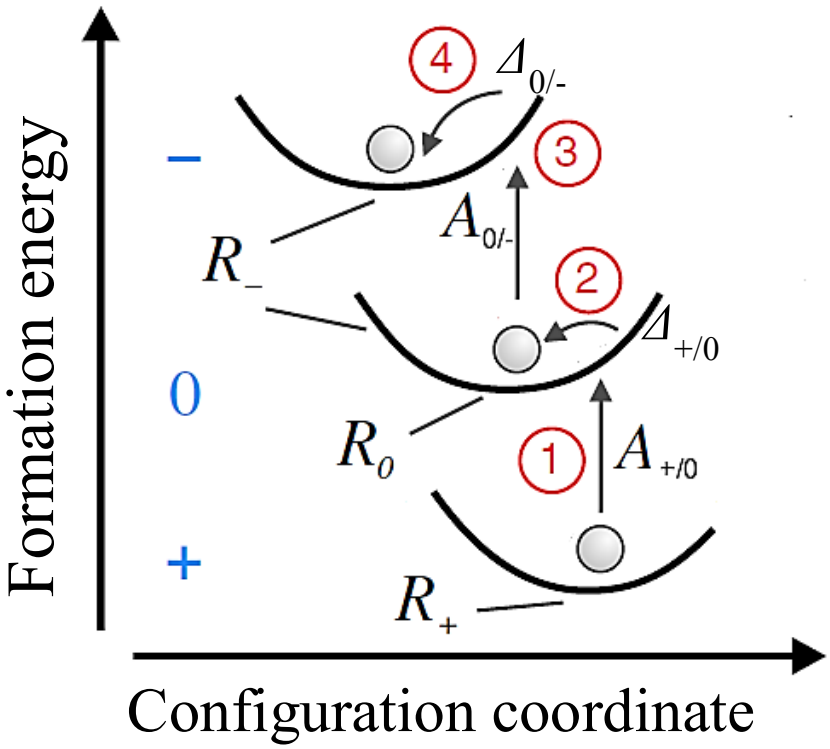


Figure 3.8 - Formation energy as a function of atomic configuration for different charge states. The process outlined by steps 1 to 4 depicts the neutral charge state of a given defect center formed from the positive, which in turn is used to form the negative charge state by adding an electron to the neutral. In each case the atomic positions are relaxed following the addition of charge. Modified after ref. [44].

As Δ can be computed without changing the number of electrons, the problem of

a discontinuous XC functional with respect to electron number does not apply, and DFT can be used to accurately calculate Δ. The opposite is true of A, such that the merit of this formalism can be seen in the identification of where approximate DFT is best applicable and where it is not. The GW approximation is used to calculate A. For electron addition ($q' \rightarrow q' - 1$) the vertical transition corresponds to the electron affinity of the defect; in this case the electron is absorbed from the surrounding reservoir into the defect level [113]. Figure 3.8 represents the formation of a charge state by addition of an electron and subsequent atomic relaxation, *e.g.* the formation of a negative charge state from the neutral state ($q = -1, q' = 0$ in equation (3.55)), or neutral from a positive state ($q = 0, q' = +1$ in equation (3.55)).

It is also possible to form charge states by removal of electrons. In this case it is assumed that the vertical electron removal energy (I) of q to q' equals the negative of the vertical addition energy ($-A$) of q' to q. This is depicted in figure 3.9. As can be seen, the relaxation energy terms are equal in magnitude. For electron removal ($q \rightarrow q + 1$), a downward vertical transition would correspond to the ionization energy of the defect, where the removed electron is transferred *from* the defect level *to* the electron reservoir.

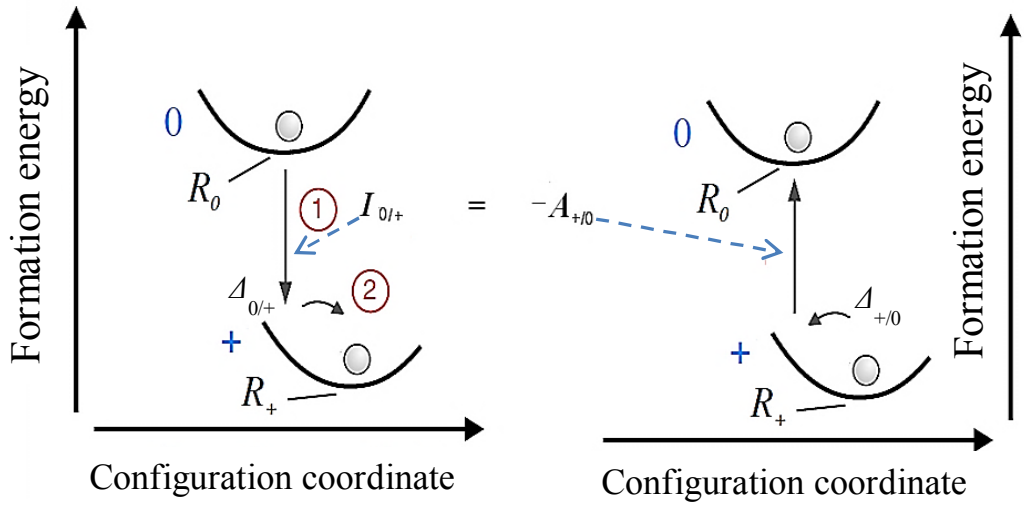


Figure 3.9 - Configuration coordinate diagram depicting the formation of a positive charge state from the neutral state (left). The energy to remove an electron from the neutral state in the neutral configuration (downward vertical transition on the left) is assumed to be equal to the negative of the energy to add an electron to the positive charge state in the neutral atomic configuration (upward vertical transition on the right).

The charge state formation energy can thus be re-expressed in terms of the energy to remove an electron plus the energy associated with the subsequent relaxation of atoms due to the added hole,

$$\begin{aligned} E_{form}(q'/q) &= E(q', R_{q'}) - E(q, R_q) + q'\Delta\varepsilon_f \\ &= E(q', R_q) - E(q, R_q) + E(q', R_{q'}) - E(q', R_q) + q'\Delta\varepsilon_f \\ &= I(q', q, R_q) + \Delta(R_{q'}, R_q, q') + q'\Delta\varepsilon_f \end{aligned} \tag{3.56}$$

where $I(q', q, R_q) = -A(q, q', R_q) = -[E(q, R_q) - E(q', R_q)]$. Hence, in this case the charge state q' is formed from q. In this work, the neutral charge state (formed from the positive charge state) and negative charge state (formed from the neutral) formation energies are calculated using equation (3.55), with the formation process described by figure 3.8. The positive charge state formation energies are calculated using equation (3.56), with the process described by figure 3.9.

3.6.1 Application of the *GW* Method

The GW method, as implemented in the many-body theory code YAMBO [75], is employed to calculate the A term of the formation energy expressions above. In this method, a first order correction to the Kohn-Sham eigenvalues E_{DFT} are obtained via many-body perturbation theory

$$E_{qp} = E_{DFT} + \langle \varphi_{DFT} | \Sigma - V_{xc} | \varphi_{DFT} \rangle \tag{3.57}$$

to yield quasiparticle levels E_{qp}. As is typical of the GW method applied to systems not involving strongly correlated d or f electrons, the quasiparticle weight Z (see section 3.4.2) is taken as unity [103]. The separation of the highest occupied and lowest unoccupied levels is taken as A. In order to calculate this correction, the self-energy is required, which in turn requires a convolution of

the Green's function G_0 and the screened Coulomb interaction W. The former is constructed using wave functions and band energies acquired from a self-consistent LDA-DFT calculation, while the latter depends on the dielectric response function that takes into account the local field effects and the dynamics of the screened interaction. In this work, we utilize the plasmon pole approximation (PPA) [50, 75] to describe the frequency dependence of the inverse dielectric function ε^{-1} required to calculate W. The PPA assumes the spectral function for the screened interaction to be a single narrow peak in frequency, corresponding to the plasmon frequency, and is thus justified if the imaginary part of the Fourier-transformed dielectric function is a peaked function in ω. Fourier transforming from real space to a wave-vector basis using the following convention

$$W(\boldsymbol{r}, \boldsymbol{r}'; \omega) = \sum_{\mathbf{G},\mathbf{G}',\mathbf{q}} e^{i(\mathbf{q}+\mathbf{G}).\boldsymbol{r}} W_{\mathbf{G},\mathbf{G}'}(\mathbf{q}, \omega) e^{-i(\mathbf{q}+\mathbf{G}').\boldsymbol{r}'} , \qquad (3.58)$$

leads to a simpler expression (compared to equation (3.36)) of the screened interaction in terms of the Fourier transformed dielectric function and Coulomb interaction $v(\mathbf{q}+\mathbf{G}')$ given by

$$W_{\mathbf{G},\mathbf{G}'}(\mathbf{q}, \omega) = \varepsilon_{\mathbf{G},\mathbf{G}'}{}^{-1}(\mathbf{q}, \omega) v(\mathbf{q}+\mathbf{G}') . \qquad (3.59)$$

Here, $\mathbf{q}$ is an arbitrary wave vector (not to be confused with the charge state symbol q used previously) while $\mathbf{G}$ is a reciprocal lattice vector. The local fields arise from the off-diagonal $\mathbf{G} \neq \mathbf{G}'$ elements. The connection back to DFT-calculated quantities is made from the relation between the static ($E = 0$) dielectric function and the non-interacting polarizability P, as the latter is obtained from LDA wavefunctions and eigenvalues as in equation (3.38)

$$\varepsilon_{\mathbf{G},\mathbf{G}'}(\mathbf{q}, E = 0) = \delta_{\mathbf{G},\mathbf{G}'} - v(\mathbf{q}+\mathbf{G}) P_{\mathbf{G},\mathbf{G}'}(\mathbf{q}, E = 0) . \qquad (3.60)$$

At this point, the PPA is employed to extend the dielectric function to finite frequencies. As stated above, the PPA assumes all the weight in the $\mathrm{Im}W_{\mathbf{G},\mathbf{G}'}(\mathbf{q},E)$ is contained in a single peak at the plasmon frequency, thus the simplest form of the PPA is given by

$$\mathrm{Im}\,\varepsilon^{-1}(\mathbf{q},E) = A_p\delta\big(E-E_p\big)\,. \tag{3.61}$$

The parameters A_p and E_p are obtained from constraints imposed by the Kramer-Kronig relation for the static dielectric matrix and the *f*-sum rule [50, 51]. The latter relates the imaginary part of the exact many-body dielectric function to the plasma frequency and the electronic density of the crystal [50]. In YAMBO, the single-pole function used to approximate ε^{-1} is expressed as [75]

$$\begin{aligned}{\varepsilon_{\mathbf{G},\mathbf{G}'}}^{-1}(\mathbf{q},E) \approx \delta_{\mathbf{G},\mathbf{G}'} + R_{\mathbf{G},\mathbf{G}'}(\mathbf{q})[&\big(E-\Omega_{\mathbf{G},\mathbf{G}'}(\mathbf{q})+i0^+\big)^{-1}\\ &-\big(E+\Omega_{\mathbf{G},\mathbf{G}'}(\mathbf{q})-i0^+\big)^{-1}]\end{aligned} \tag{3.62}$$

for each set of momentum components $\{\mathbf{G},\mathbf{G}',\mathbf{q}\}$. The YAMBO parameters $R_{\mathbf{G},\mathbf{G}'}$ and $\Omega_{\mathbf{G},\mathbf{G}'}$ are obtained by forcing the approximation to reproduce the exact many-body dielectric function at the static limit ($E=0$), and at $E=iE_{ppa}$, where E_{ppa} is defined in the input file [75]. For all calculations in this work, the default value of E_{ppa} = 1 Hartree is used.

That the imaginary part of the inverse dielectric function exhibits a single peak with respect to frequency is observed for a number of common semiconductors and insulators [50]. In these cases the PPA is justified since the evaluation of the self-energy involves an integral over ω so the fine details of the energy/frequency dependence should not be critical. When applied to certain metals like Cu [114], the PPA tends to provide too simple a picture as the behaviour of ${\varepsilon_{\mathbf{G},\mathbf{G}'}}^{-1}(\mathbf{q},E)$ in such systems is very different to that of a single pole function. When the PPA breaks down, the dielectric function must be explicitly computed throughout the full frequency axis, leading to large increases

in computational time and memory. Fortunately, a single peak in the dielectric function is observed for $In_{0.53}Ga_{0.47}As$ as studied in this thesis; both the bulk and surface models (presented in chapter 4 and 5, respectively) exhibit a single peak in the imaginary part of the inverse dielectric response function. This allows for the application of the GW method to supercells containing >150 atoms.

3.7 CNT Vacancy Formation

In this section, our approach to calculating the formation energetics of monovacancies and divacancies in CNTs is presented. The formation energy ΔE of each defect is calculated by considering the difference between the energy of the pristine tube, $E_{pristine}$ and the sum of the energy of the defective tube, E_{defect} plus that of an isolated, non-interacting carbon atom or twice this, for divacancies,

$$\Delta E = \left| E_{pristine} - E_{sum} \right| \tag{3.63}$$

where,

$$E_{sum} = E_{carbon} + E_{defect} \text{ (for monovacancies),}$$
$$E_{sum} = 2E_{carbon} + E_{defect} \text{ (for divacancies).}$$

Although energetically the pristine structure is always more favorable, the formation energy allows comparison of the relative stability of defects. Equation (3.63) implies that the smaller the formation energy, the greater the stability of a specific defect. This relation was used previously to assess the relative stabilities of mono- and divacancies as a function of armchair CNT diameter [12]. However, fragmentation processes in materials such as C_{60} or larger fullerenes typically proceed by desorbing a carbon dimer [115], thus a more accurate measure of divacancy formation is to assume the dissociated state consists of the defect tube and a free C_2 molecule. In this case, E_{sum} in equation (3.63) is

replaced by $E_{C_2} + E_{defect}$, where E_{C_2} is the energy of the carbon dimer.

To compare the energies associated with the orientation of the divacancy defects relative to the CNT axis, the quantity $d\Delta = |\Delta E_{vertdivac} - \Delta E_{latdivac}|$ is introduced; the monotonic reduction of this value as a function of tube diameter (see chapter 6) reflects the convergence to a graphene-like structure for larger CNT radii. We examined semiconducting and metallic CNTs of various sizes, allowing for the analysis of defect energetics as a function of CNT diameter and chirality. The formation energies and relative stabilities are tabulated for zigzag and armchair CNTs in chapter 6. A similar analysis was also performed for graphene. Where possible, the computational details (*e.g.*, basis sets, supercell dimensions) for the relaxation of graphene, as well as for the single-point energy calculation after relaxing, were chosen similar to those of the CNTs.

3.8 Electron Transport in CNTs

3.8.1 DFT + Green's Functions

To assess electronic transport across defective nanotubes, the transport simulator TiMeS (Transport In MEsoscopic Systems) [73, 116] was used to calculate the electron transmission spectrum [12, 68, 117] which is proportional to the low-temperature low-bias Landauer conductance [24, 118]. The transmission spectrum can also be used to extract the mean free path for impurity scattering of charge carriers, as applied previously [73]. Each CNT is

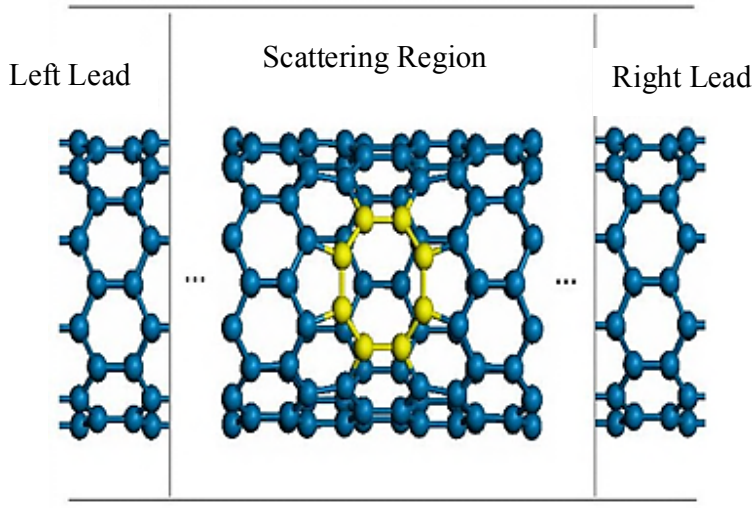


Figure 3.10 - Schematic partitioning of a CNT (chirality (10, 0) with a lateral divacancy shown here) into lead and scattering regions. The dots indicate the repetition of ideal principal layers in each lead.

divided into 3 regions: a scattering region in the center which contains the defective supercell, and left and right pristine leads on either side which act as electrodes (Figure 3.10). Electronic states from the leads are scattered as they propagate through the nanotube structure. The framework underlying the TiMeS calculation of the quantum-mechanical scattering matrix from which transport coefficients can be extracted is based on a Green's function implementation [24, 117, 118]. The matrix representations of the Hamiltonians in the localized numerical atomic orbital (NAO) [96] basis for the relaxed structures are obtained from OpenMX, and these Hamiltonians are transferred to TiMeS [116]. For this step a single-point calculation was performed using a single-ζ basis set to reduce the computational cost in computing transport properties, allowing us to extend our study to larger CNT radii. It is shown in chapter 6 that the use of a double-ζ basis does not affect transport significantly as may be expected for the simple sp^2 network of carbon nanotubes. For the single-point calculations in tubes of chirality (n, 0) and (n, n), the supercell contains 13 unit cells for $n \leq 10$, and 7 unit cells for $n \geq 20$, with the central cell containing the defect.

3.8.2 Mean Free Path

Once the transmission function $T(E)$ is obtained, the mean free path λ as a function of the distance between defects l_d can be calculated from the following considerations. In a diffusive one-dimensional conductor, the transmission T is proportional to the ratio of λ to the length of the conductor L, and to a first approximation this ratio is effectively the transmission probability of each channel. For N_{ch} channels, and assuming $L >> \lambda$, the transmission probability can be approximated as [24]

$$T \cong N_{ch} \frac{\lambda}{L} \ . \tag{3.64}$$

Clearly, the resistance of such a conductor is proportional to the inverse of this,

$$R = R_c \frac{L}{\lambda} , \tag{3.65}$$

where R_c is the contact resistance,

$$R_c = \frac{h}{2e^2} N_{ch}^{-1} \ . \tag{3.66}$$

According to the ISA, the resistance due to defect scattering will simply be the sum of the individual resistances of each defect, so that, assuming each defect has the same "scattering" resistance R_s, the total resistance is

$$R = R_c + R_s \frac{L}{l_d} \approx R_s \frac{L}{l_d} . \tag{3.67}$$

Here, L/l_d = number of defects, and we are assuming $R_c \ll R_s$ implying a diffusive regime for transport. Expressions (3.65) and (3.67) can be equated to solve for the mean free path

$$\lambda = \frac{R_c}{R_s} l_d \ , \tag{3.68}$$

where the scattering resistance is given by [119]

$$R_s = \frac{h}{2e^2} T^{-1} - \frac{h}{2e^2} N_{ch}^{-1} \ , \tag{3.69}$$

and T is the transmission. In chapter 6, λ is plotted against a range of values of l_d for all CNTs considered, and where possible, we show that this relationship can yield an experimentally verifiable relationship between the localization length and the defect density. The value used for the transmission in equation (3.69) is $T(\varepsilon_f)$ for metallic tubes and $T(E_{CBM} + 0.3\ eV)$ for semiconducting tubes, where ε_f is the Fermi level and E_{CBM} is the conduction band minimum.

Chapter 4
Bulk InGaAs

4.1 Introduction

In this chapter simulations of native point defects in bulk $In_{0.53}Ga_{0.47}As$ are presented. A number of methods exist for the modeling of the atomic structure of compound semiconductors. We first employ a computationally convenient approach which neglects short-range cation disorder. This allows us to evaluate the efficacy of the GW approximation in describing the bulk $In_{0.53}Ga_{0.47}As$ electronic structure. Following this, a more realistic model of bulk $In_{0.53}Ga_{0.47}As$ is given in which the cations are randomly distributed. Defects are then introduced into this explicit model and calculations of thermodynamic transition levels of energetically competing defect charge states are performed. These calculations involve a combination of DFT and the MBPT-derived GW approximation. The results of our defect calculations agree semi-quantitatively with previous methods involving hybrid density functionals, although important differences are found and these previous calculations are extended to new cases.

4.2 Electronic Structure

4.2.1 The Virtual Crystal Approximation

Application of the supercell method to the study of disordered alloys, in which large cells are utilized in order to mimic the distribution of local atomic arrangements of the alloy constituents, can be computationally demanding. A computationally convenient technique is to employ the virtual crystal approximation (VCA), in which the periodic crystal is constructed from a primitive unit cell containing one or more “virtual” or “averaged” atoms. This method has been applied to the study of electronic structure [120], atomic

structure and thermodynamics [121], and dielectric properties [122]. In the VCA, the virtual atoms are described by pseudopotentials which linearly interpolate between those of the alloy constituents, such that the resulting virtual pseudopotential achieves the desired alloy proportion. For example, the VCA-description of bulk $In_xGa_{1-x}As$ replaces the individual In and Ga pseudopotentials with a virtual "In_xGa_{1-x}" pseudopotential

$$V_{In_xGa_{1-x}} = xV_{In} + (1-x)V_{Ga} \quad , \tag{4.1}$$

which is then placed in a zinc blende primitive unit cell along with As and periodic boundary conditions are imposed. This unit cell is then treated in the usual first-principles pseudopotential approach, *i.e.* the Kohn-Sham (KS) equations are solved with the $V_{ext}(\boldsymbol{r})$ term of the Kohn-Sham potential containing contributions from $V_{In_xGa_{1-x}}$ and V_{As}. The purpose of the application of the VCA in this work is to demonstrate the qualitative features of the electronic structure calculated from the DFT+GW method. Specifically, the convergence of the quasiparticle gap with respect to YAMBO input parameters is achieved and subsequently the bandstructure is calculated, showing the qualitative agreement of the energy-momentum dispersion throughout the Brillouin Zone (BZ) with previous work based on empirical pseudopotentials. Quantitative comparisons show that our DFT+GW approach yields on average better agreement with a hybrid-DFT study than with the empirical pseudopotential method, although in general the overall agreement is good between all three methods.

4.2.1.1 Numerical Details

A 2-atom model of bulk $In_{0.53}Ga_{0.47}As$ is generated using the VCA by linearly mixing the local and non-local parts of the In and Ga pseudopotentials using equation (4.1). Norm-conserving pseudopotentials utilizing the Perdew-Zunger

form of LDA [89, 123] are employed for In, Ga, and As. These In and Ga pseudopotentials are used to generate the virtual In_xGa_{1-x} pseudopotential with x = 0.53. The use of these pseudopotentials results in 8 electrons per cell, 3 from '$In_{0.53}Ga_{0.47}$' and 5 from As. Zinc blende symmetry and periodic boundary conditions are imposed, and the lattice constant set to the experimental value of 5.87 Å [124]. The ground state electronic structure for the resulting VCA model of '$In_{0.53}Ga_{0.47}As$' is self-consistently calculated using the Quantum Espresso code [99]. Plane wave basis sets are used to expand the electronic wavefunction, with a kinetic energy cutoff of 60 Rydbergs. The self-consistent wavefunctions generated by Quantum Espresso are used as the input to YAMBO [75]. Initially, the convergence of the quasiparticle gap is studied with respect to calculation parameters including the number of unoccupied states, plane wave basis set size, size of dielectric response matrix, and number of points in the FFT grid. For all calculations of the GW correction, all bands included in the DFT run are included in the dielectric function, and in the correlation part of the self-energy [75]. The convergence tests also include the *k*-point grid which is increased to 16x16x16, at which point the quasiparticle gap changes by less than 1 meV compared to 12x12x12. 12x12x12 *k-points* are used to plot the quasiparticle bandstructure.

4.2.1.2 Results and Discussion

The LDA-DFT bandgap is significantly underestimated at 0.34 eV, compared to the low-temperature experimental value of 0.82 eV [125]. This underestimation is attributed to the limitations of DFT discussed in section 3.3. Upon correcting the KS eigenvalues with the GW method, the HOMO-LUMO gap dramatically increases. The plasmon pole approximation (PPA) is used to describe the dynamical effects of the GW approximation. The validity of the PPA is determined from a plot of the imaginary part of the inverse dielectric function versus energy, as described in the work of Hybertsen and Louie [50]. Using the DFT parameters mentioned in the previous section, along with 46 unoccupied

bands in the dielectric response function and neglecting local fields (no off-diagonal elements in the dielectric matrix), the real and imaginary parts of the dielectric matrix are plotted as a function of energy as shown in figure 4.1.

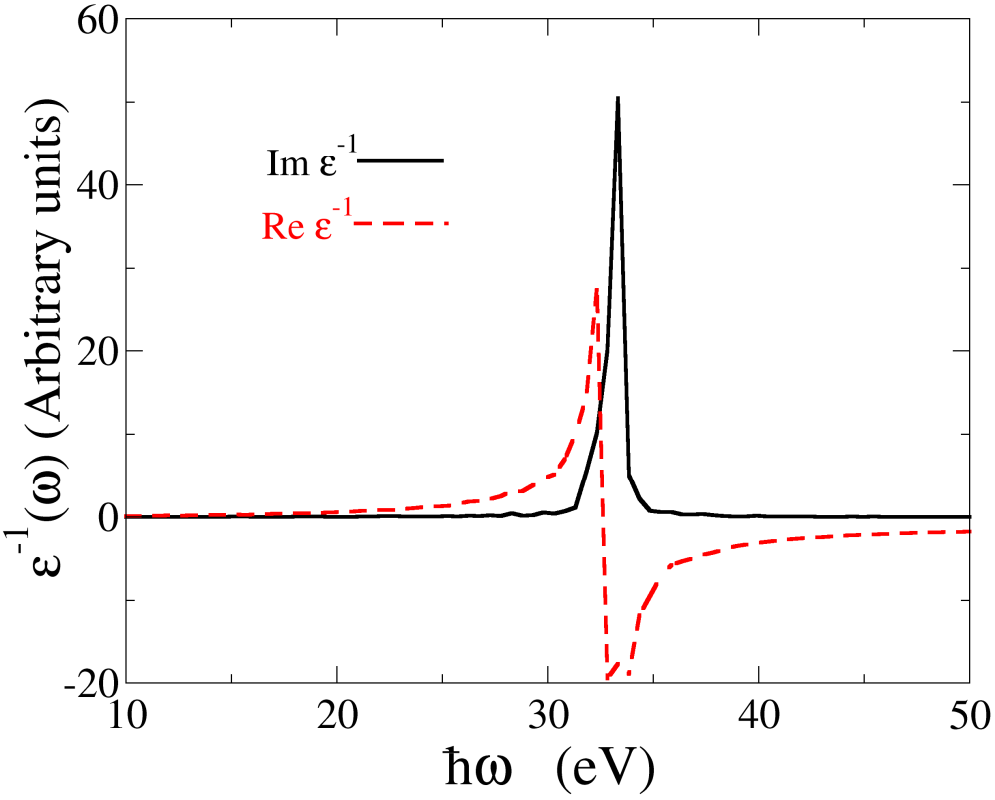


Figure 4.1 - Real and imaginary parts at the Γ-point component ($\mathbf{G} = \mathbf{G}' = \mathbf{0}$) of the inverse dielectric function, obtained from the VCA model of bulk $In_{0.53}Ga_{0.47}As$, plotted as a function of energy.

The single large peak in the spectrum of the dielectric function allows for an extension of the static dielectric matrix to finite frequencies using the PPA [50]. Hence, the PPA remains valid for $In_{0.53}Ga_{0.47}As$ even when the short range order of the cation sublattice is neglected. Using a plasmon-pole approximated dielectric function to calculate the self-energy, the resulting difference between highest occupied and lowest unoccupied *GW*-corrected DFT states is highly overestimated compared to the experimental value. This corresponds to including dynamical effects via the PPA, but with an inadequate description of

local fields. Adequate convergence is obtained once the size of the dielectric matrix $\varepsilon_{\mathbf{G},\mathbf{G}'}$ is increased to 181 × 181 where each element corresponds to a reciprocal lattice vector. Increasing $\varepsilon_{\mathbf{G},\mathbf{G}'}$ to 259 × 259 results in less than 1 meV change in the bandgap relative to 181 × 181, see figure 4.2. Numerical convergence is also achieved with respect to all other YAMBO calculation parameters, including the FFT integration mesh. Increasing the number of unoccupied states from 46 to 56 results in less than 1 meV difference to the *GW* bandgap relative to the converged value, and a similarly negligible change in the bandgap is observed upon increasing the plane wave cutoff energy to 120 Rydbergs.

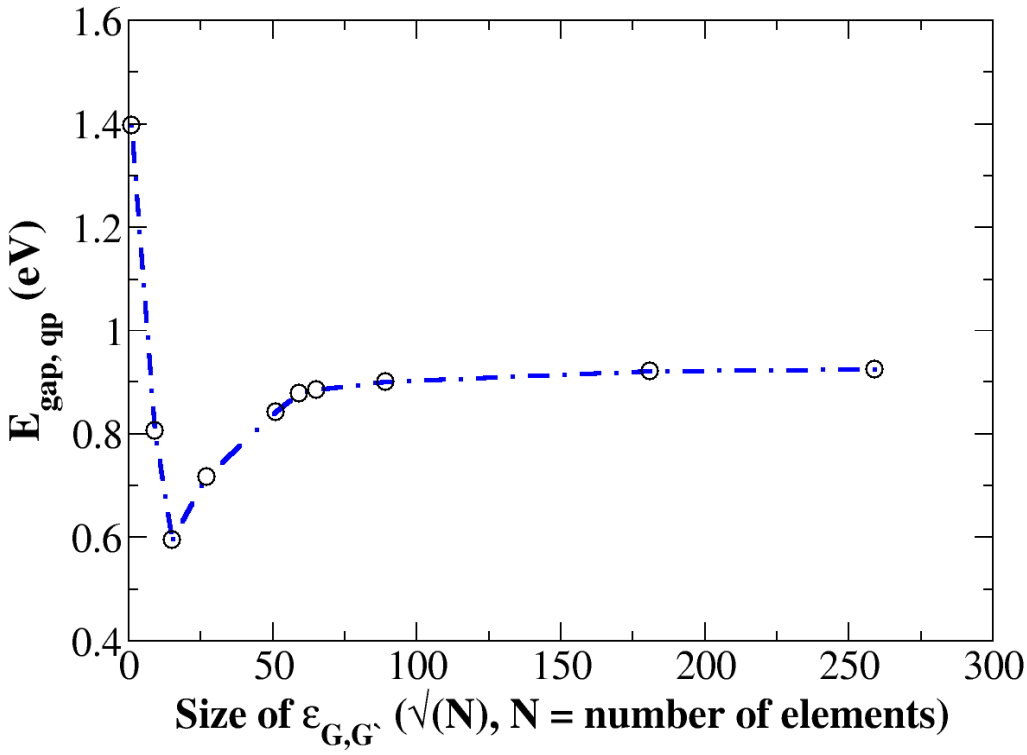


Figure 4.2 - Bandgap calculated from *GW*-corrected DFT states ($E_{gap,qp}$), versus the size of the dielectric matrix.

Using 56 unoccupied states, 1728 (72) *k*-points in the (irreducible) BZ, and 181 × 181 elements in $\varepsilon_{\mathbf{G},\mathbf{G}'}$, the *GW*-corrected bandstructure was plotted using a *k*-

path interpolated from the uniform grid. The YAMBO code follows the interpolation procedure of ref. [126].

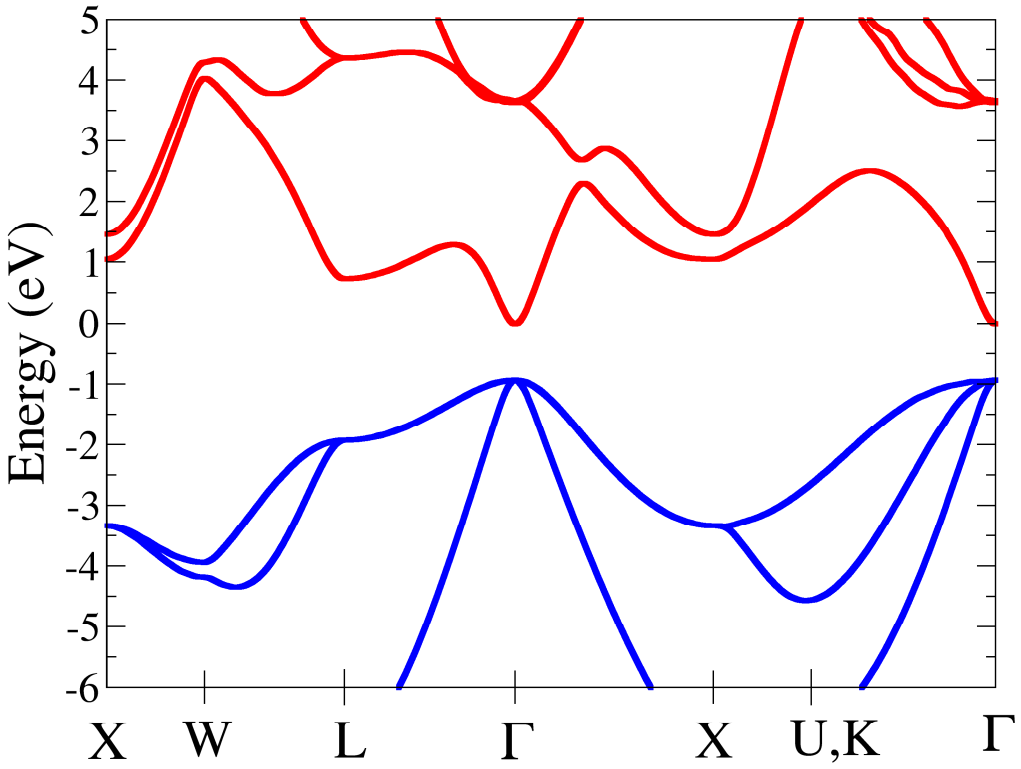


Figure 4.3 - Bandstructure calculated from *GW*-corrected Kohn-Sham states using a 2-atom VCA model of $In_{0.53}Ga_{0.47}As$. The *GW*-corrected bandgap at Γ is 0.92 eV.

The converged *GW*-corrected 0 K Γ-point bandgap is 0.92 eV, which somewhat overestimates the experimental value of 0.82 eV [125], but is in much better agreement with experiment than the uncorrected DFT bandgap of 0.34 eV. In addition, this is a comparable level of accuracy as that obtained for semiconductors (diamond, Si, Ge, and LiCl) studied by Hybertsen and Louie [50], more recent studies of bulk Si by Rinke *et al* [44], and bulk GaAs and InP calculations reported by Hedström, Schindlmayr and Scheffler [22]. A small overestimation of the low temperature experimental bandgap using a DFT+*GW* approach is a reoccurring theme throughout these works, and is on the order of a

few tens to 100 meV. In addition, recent studies of graphene nanoribbons within the GW approximation report accuracies in the quasiparticle energies of 0.1 eV-0.2 eV [127, 128]. Our results are well within the lower limits of this range of error.

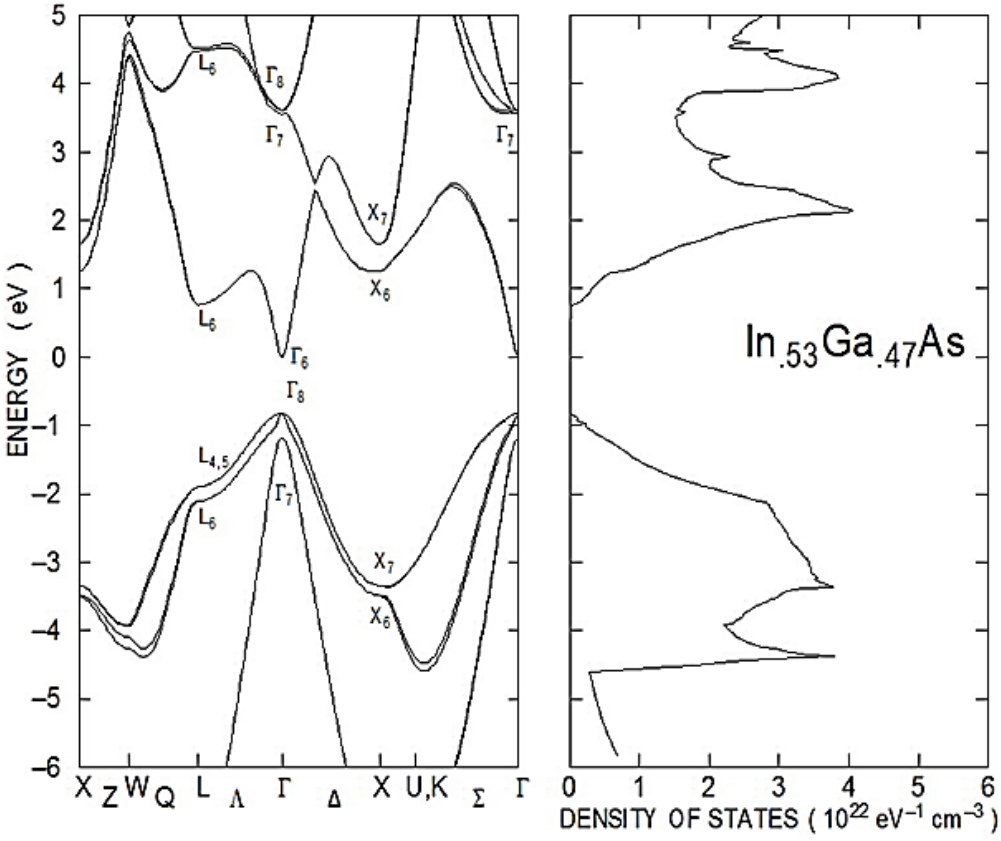


Figure 4.4 - Bandstructure of $In_{0.53}Ga_{0.47}As$ calculated from an empirical pseudopotential formalism. Private communication from M. Fischetti [129].

Comparing with a bandstructure calculated by M. Fischetti (figure 4.4, unpublished) who used empirical pseudopotentials, we find overall qualitative agreement in the dispersion of occupied and unoccupied states. To quantitatively assess the comparison, we compare our satellite valley energy separations at high symmetry points (L and X) to those calculated using empirical pseudopotentials, shown in table 4.1 below. Satellite valleys are important figures of merit which play a role in the build-up of electron concentration at the

semiconductor surface as a function of gate voltage in MOS devices. They are also used as input device parameters in empirical transport simulations [130]. In addition to our DFT+*GW* results, we also include satellite valleys obtained from a hybrid functional calculation applied to bulk $In_{0.53}Ga_{0.47}As$. The Heyd-Scuseria-Ernzerhof (HSE) hybrid functional is employed [131]. When the proportion of Hartree-Fock exchange, denoted a, is set to 0.27, a Γ-point bandgap of 0.92eV is obtained, *i.e.* reproducing our DFT+*GW* result.

Γ/eV	***L/eV***	***X/eV***	***Method***
0.73	1.49	1.98	Fischetti (empirical pseudopotentials) [130]
0.92	1.64	1.96	This work (*GW*+DFT)
0.92	1.77	2.05	This work (hybrid-DFT, $a = 0.27$)
0.82	1.68	1.99	This work (hybrid-DFT, $a = 0.24$)

Table 4.1 - High symmetry satellite valley energy separations relative to the VBM obtained from M. Fischetti who used empirical pseudopotentials, and compared to our values obtained using a DFT+*GW* approach and hybrid functionals.

Using $a = 0.27$ in order to investigate the difference of the satellite valley energies between the DFT+*GW* and hybrid functional schemes, we find better agreement between DFT+*GW* and hybrid-DFT than with the empirical pseudopotentials for the L-valley. The X-valley values are closer for all three methods, and for both satellite valleys hybrid-DFT yields the highest estimate amongst all these approaches. To investigate the effect of parameterizing the hybrid-DFT calculation to fit DFT+*GW*, a was also adjusted to reproduce the low temperature experimental bandgap 0.82 eV ($a = 0.24$); in this case, the X-valley decreases by 60 meV compared to $a = 0.27$, and again the L-valley is closer to the DFT+ *GW* value than to the empirical pseudopotential value.

Regarding the *GW*-corrected bandgap, it is interesting to note that the VCA approach has led to larger bandgaps relative to explicit tight-binding models in previous studies [120, 132]. Such findings suggest that the difference between the *GW*-corrected bandgap calculated using the VCA and the experimental value may not be entirely due to the *GW* approximation. In fact,

we find a lowering of the bandgap when the individual cations are taken into account, as shown in the next section. In any case, a large improvement over DFT in the description of the electronic structure is obtained using the GW approximation. Considering the slightly larger DFT+GW-calculated bandgaps relative to experiment commonly reported in the literature [22, 44, 50] and the fact that the VCA method also tends to overestimate the bandgap, our results seem to be well within the accepted limits of accuracy and significantly improve the DFT bandstructure. This gives us confidence to move onto systems consisting of an explicit description of the cation arrangement in which point defects may be studied.

4.2.2 A 64-Atom Supercell Model

4.2.2.1 Numerical Details

We use norm-conserving pseudopotentials utilizing the Perdew-Zunger form of LDA [89, 123]. A kinetic energy cutoff of 60 Rydbergs is used for the plane wave basis set. 2x2x2 *k*-point meshes are used for geometry optimization. To facilitate comparison with recent experimental results [6], an alloy proportion of 53 % In and 47 % Ga is desired. This was achieved by randomly distributing 17 In and 15 Ga in a 64-atom supercell. In order to compare with the VCA study of the previous section, atomic geometries are relaxed at the experimental lattice constant. This will allow for an investigation of the importance of explicitly taking into account the distribution of indium and gallium sites. Quantum Espresso implements a Quasi-Newton optimization algorithm [78], which is utilized here for all geometry relaxations. Regarding the iterative loop for calculating a self-consistent electronic structure, the charge density mixing scheme mentioned in section 3.3.4 is used.

4.2.2.2 Results and Discussion

Figures 4.4 and 4.5 show the convergence of the *GW*-corrected gap calculated at the Γ-point, $E_{gap,qp}$, as a function of the number of unoccupied bands and dielectric matrix size, respectively. Both the dielectric matrix and the correlation part of the self-energy (Coulomb-hole self-energy term in the static limit) require a summation over unoccupied states. Compared to the 2-atom VCA study, a much larger number of unoccupied states are required to achieve adequate convergence. This is attributed to the larger number of occupied states (128) in the 64-atom case, which leads to a much larger band summation in the evaluation of the self-energy operator [51].

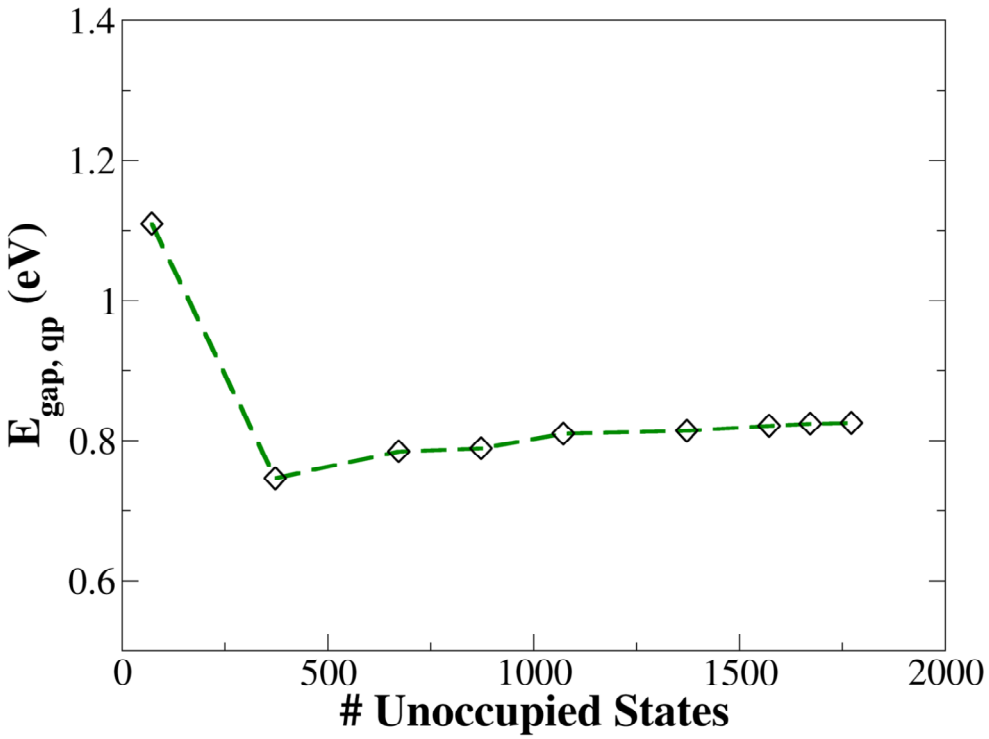


Figure 4.4 - Bandgap calculated from *GW*-corrected DFT states ($E_{gap,qp}$), versus the number of unoccupied states used to calculate the self-energy.

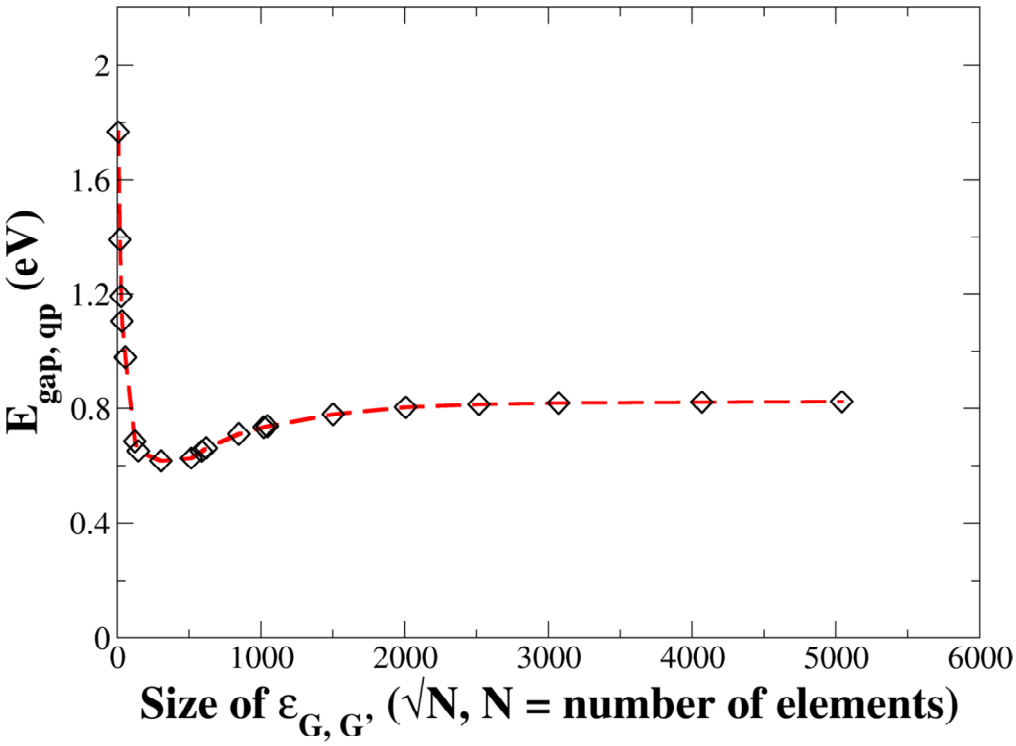


Figure 4.5 - Bandgap calculated from *GW*-corrected DFT states ($E_{gap,qp}$), versus the number of elements (reciprocal lattice vectors) used in the dielectric matrix.

As expected, the primary effect of correcting DFT with the *GW* approximation is a shifting upwards in energy for conduction band states and a lowering of energy for valence band states. The bandgap from LDA-DFT is 0.31 eV for the 64-atom cell, while the converged *GW*-corrected bandgap $E_{gap,qp}$ = 0.83 eV is in excellent agreement with the low temperature experimental bandgap of 0.82 eV [125]. Comparing with the quasiparticle gap calculated using the VCA (0.92 eV), the importance of explicitly accounting for the differing cation potentials and positions becomes apparent.

4.3 Defects - V_{Ga}, Ga_{As} and As_{Ga}

The purpose of this section is to discuss the study of native defects in bulk InGaAs. Native defects have been invoked in the past to explain the observation of midgap peaks in the distribution of interface states for $In_{0.53}Ga_{0.47}As$ as determined by electrical spectroscopy. The exact nature of the defect(s) responsible for the peaks remains elusive, which motivates first-principles computational studies. In this work, the following defects are considered for the bulk simulation cell: two antisites, Ga_{As} (Ga on an As site) and As_{Ga} (As on a Ga site), and a Ga vacancy, denoted V_{Ga}. These defects have been previously studied using hybrid-DFT functionals [13, 41] in the bulk, and shown to exhibit charge transition levels lying within the bandgap. Here, we study the same defects to provide a comparison between hybrid functionals and the DFT+GW approach which is free of empirical parameterization. Following the introduction of defects, the GW correction is used to evaluate the electron addition and removal energies of a range of charge states of each defect, yielding the charge transition levels using the formalism described in chapter 3.

These calculations serve to "calibrate" the hybrid-DFT and DFT+GW methods against each other, and to consider the calculations against what is known experimentally. This lays the groundwork for calculations which explicitly include a semiconductor/oxide interface and the effects on defects thereof considered in chapter 5.

4.3.1 Methods

Using the 64-atom model of bulk InGaAs described in section 4.2.2, full structural relaxation of the supercell plus atomic geometries is performed while maintaining a defect free cell. Relaxation of the lattice vectors for the defect free cell circumvents any possible errors due to non-equilibrium supercell volumes and unconverged supercell energies, as discussed in [5]. Following the full

structural relaxation of the pristine cell, point defects are then inserted and the atomic geometries re-optimized. This is repeated for each point defect, i.e. one defect per 64-atom cell. Carrying out this procedure for various charge states of each defect allows for the determination of formation energies as described in section 3.6. For charged simulation cells, Quantum Espresso applies a compensating background charge density; this prevents divergences due to Coulomb interaction between image charges arising from periodic boundary conditions and ensures well defined total energies. Regarding the numerical details of the DFT calculations and subsequent GW corrections, similar convergence criteria to the pristine case were applied to defective cells; specifically a 60 Rydbergs kinetic energy cutoff, 2x2x2 k-point meshes for geometry optimization, and 1772 unoccupied bands along with $\sqrt{\mathrm{N}}$ = 5041 where N is the number of elements in the dielectric matrix used to calculate the self-energy for the GW correction. This value of N was chosen based on the convergence of the bandgap of the pristine cell, see figure 4.5.

4.3.2 Structural Properties

Three point defects in the bulk simulation cell are studied separately, V_{Ga}, Ga_{As}, and As_{Ga}. In the following, displacements for neutral defects are measured relative to pristine cell, while displacements for charged defects are measured with respect to the relaxed neutral defects. Focusing on the Ga vacancy first, upon removal of a Ga atom the four surrounding As atoms relax by moving in towards the vacancy site by 0.5 Å on average. The tendency of surrounding anions to form weak bonds across a cation vacancy has been indicated by previous computational studies of III-Vs [133]. The structural effect of charging this defect by adding a single electron is a very slight (~0.01 Å displacement) further inward movement of the As atoms. Removing an electron from the neutral state and relaxing to bring the vacancy to a positively charged configuration results in a similar small degree of movement of the surrounding

As atoms, but in the opposite direction to relaxation found in the negative state. Looking to the Ga_{As} antisite, which is bonded to two Ga and two In atoms, we find that the bonds between the neutral defect and the surrounding Ga atoms shorten by 0.08 Å relative to the unrelaxed substitutional site, that is relative to an As site in the pristine cell, while the bonds between the defect and the nearest-neighbor In atoms elongate by 0.04 Å relative to the unrelaxed position. Thus, the Ga_{As} remains in a four-fold coordinated arrangement but moves to an off-center position with respect the non-defective As site (see figure 4.6 (b)). Charging the Ga_{As} site to the positive state pushes the defect center up toward the surrounding Ga atoms (see figure 4.6) by 0.04 Å compared to the neutral position. Charging this antisite to the negative state results in a slight downward (see figure 4.6) movement of the defect center by 0.03 Å compared to the neutral position, in the opposite direction to the positive state relaxation. For the neutral As_{Ga} antisite, the surrounding anion bonds undergo larger changes compared to the cation bonds of the Ga_{As} antisite: an outward relaxation of 0.16 Å relative to the pristine bonds is found. Removing an electron causes the As-As_{Ga} bonds to contract by 0.05 Å relative to the neutral case, while the defect center remains in the same position as the relaxed neutral charge state. This behaviour continues for the doubly charged state of the As_{Ga} antisite; an average of 0.06 Å contraction of the anion-antisite bond lengths relative to the +1 state, and the As_{Ga} defect center remains in the same position to within 0.01 Å. For all defects, we find that atoms located along atomic planes which bisect the spacing between periodic images of the defect center undergo displacements of less than 0.01 Å on average relative to the corresponding bonds in the pristine cell, indicating that the surrounding environment approaches that of an isolated point defect. The periodic supercells containing each neutral defect (highlighted) after relaxation are shown in figure 4.6.

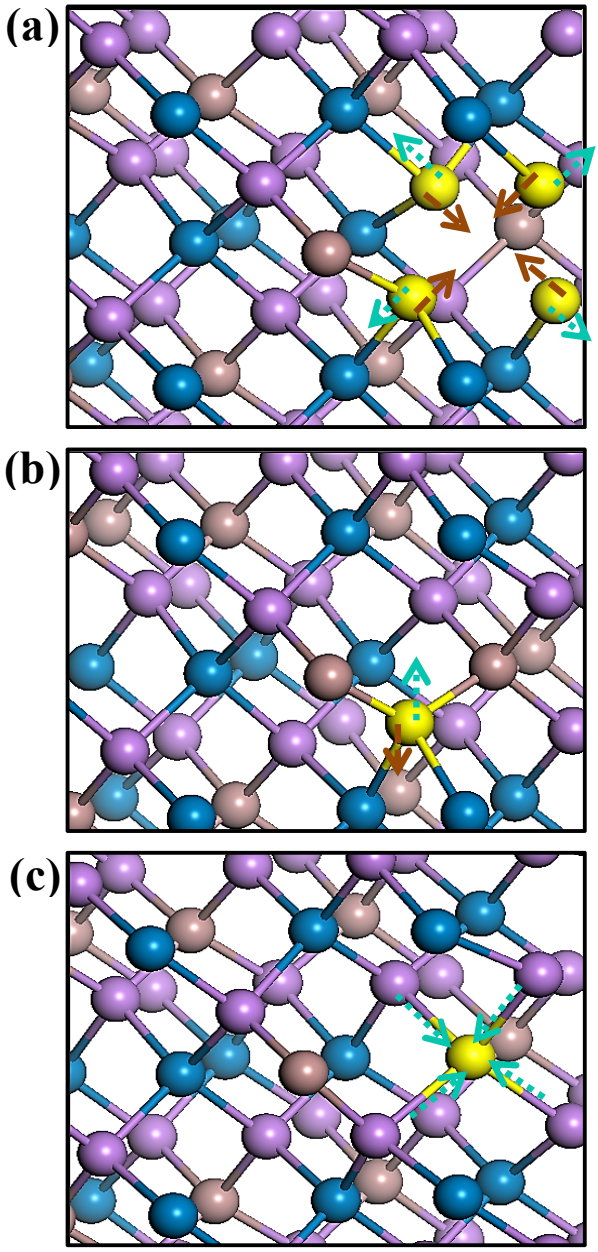


Figure 4.6 - Relaxed atomic structures of neutral point defects in $In_{0.53}Ga_{0.47}As$. Blue, brown and, pink spheres represent In, Ga, and As, respectively. For (a), the As atoms surrounding the Ga vacancy are highlighted. For (b) and (c), the Ga antisite and the As antisite, respectively, are highlighted. Arrows indicate the direction of movement of atoms for relaxation due to positive (turquoise dotted) and negative (orange dashed) charges.

4.3.3 Charge Transition Levels for Bulk Defects

Charge transition levels of each defect were calculated as values of the Fermi energy in which charge states have equal formation energy, as outlined in chapter 3. We first discuss the recently published charge transition levels from Komsa and Pasquarello obtained within hybrid-DFT [13] for bulk defects. For the V_{Ga}, they calculate a neutral to negative transition, denoted here as $\varepsilon_{0/-1}$, at ~0.08 eV above the valence band maximum (VBM). They obtain a $\varepsilon_{+1/0}$ transition for the As_{Ga} antisite ~0.74 eV above the VBM, while the $\varepsilon_{+2/+1}$ As_{Ga} transition lies close to midgap at ~0.42 eV above the VBM. The latter is ~0.15 eV lower than the $\varepsilon_{0/-1}$ transition for Ga_{As}, while the $\varepsilon_{+1/0}$ Ga_{As} transition lies slightly above the $\varepsilon_{0/-1}$ V_{Ga} transition from their hybrid-DFT calculations. In their work, a study of the defect formation energies as a function of approximated growth conditions is also performed [13], as described in section 2.3.2. Based on this study and on the proximity of the $\varepsilon_{+2/+1}$ As_{Ga} transition to the midgap D_{it} peak measured at InGaAs/oxide interfaces by electrical spectroscopy, they assign the As_{Ga} antisite as the sole defect responsible for the observed measurements. We calculate the same defects within the DFT+GW approach and compare with these hybrid-DFT results; overall qualitative agreement is obtained.

A different approach to the calculation of charge transition energies is taken in this work (section 3.6). The formation energies at zero gate bias are calculated by adding two terms $A + \Delta$. The former is the electron addition energy, while the latter is associated with the structural relaxation due to charge in the defect. We consider the formation of the neutral V_{Ga} first, obtained by adding an electron to the positive charge state of V_{Ga}, and subsequently relaxing the geometry of the simulation supercell. Taking the total energy difference between the neutral cell in its relaxed geometry and the neutral cell with the atomic positions fixed to those of the +1 charge state as the relaxation energy term Δ, we find Δ = -0.01 eV (negative relaxation energy terms are expected when moving "downward" in the configuration coordinate diagram, see figure 3.8). The DFT+GW approach yields a GW-corrected HOMO-LUMO separation

of 0.07 eV, bringing the formation energy of the neutral V_{Ga} defect to 0.06 eV. The -1 charge state is formed by adding an electron to the neutral defect. The DFT-calculated relaxation energy plus the GW correction to the electron addition energy brings $E_{form}(q = \text{-}1)$ to 0.21 eV for the V_{Ga} defect. Both the charge state relaxation energies for this defect are relatively low, which is anticipated by the correspondingly small shift of atomic positions between each relaxed charge state.

A brief explanation of the slope of $E_{form}(q, \Delta\varepsilon_f)$ now follows. For acceptor-like transitions in which, for example, a neutral defect acquires an electron from the surrounding bulk material or reservoir characterized by the Fermi level ε_f, the energy to form the bound state, or the work done in moving the electron from ε_f up to the lowest unoccupied defect level decreases with increasing ε_f. A similar line of reasoning explains why the formation energy of positive states (*i.e.* donor-like transitions from the neutral) increases with ε_f, and in general why the slope of $E_{form}(q, \Delta\varepsilon_f)$ is determined by the final charge state [113]. Plotting the formation energies against a linear scale which represents the change in ε_f due to a gate bias as in figure 4.7 (pages 94-95), we see that charge transition levels can be expressed as

$$\varepsilon_{q/q'} = \frac{E_{form}(q', \Delta\varepsilon_f = 0) - E_{form}(q, \Delta\varepsilon_f = 0)}{q - q'}, \qquad (4.2)$$

since the transition occurs for the value of ε_f at which the crossing of the energies occurs. For the V_{Ga} case $q = 0$ and $q' = \text{-}1$; hence, we find a charge transition between the neutral and negative states of V_{Ga} at 0.15 eV above the VBM. These findings suggest the neutral Ga vacancy exhibits acceptor-like behaviour in III-V materials, in qualitative agreement with previous computational studies [13, 41, 133].

For the +1 charge state of the Ga_{As} antisite in the relaxed configuration, an electron addition energy of $A = 0.11$ eV is obtained within the GW

approximation. Adding this to the structural relaxation energy of the neutral charge state results in 0.06 eV for the neutral formation energy of the Ga_{As} antisite at zero gate bias. Forming the negative and positive charge states by adding and removing an electron to and from the neutral as described in section 3.6, we find charge transition levels $\varepsilon_{+1/0}$ and $\varepsilon_{0/-1}$ at 0.19 eV and 0.31 eV above the VBM, respectively.

For the As_{Ga} antisite the +1 and +2 states are both formed by successive removal of electrons starting from $q = 0$. We find GW-corrected charge transition energies for As_{Ga} slightly below midgap for the +2/+1 transition and close to the experimental conduction band edge for the +1/0 transition. The magnitude of Δ reflects the degree of structural rearrangement between each charge state; the Δ term averages at -97 meV for As_{Ga}, while for Ga_{As} the average is -49 meV. The values of all GW-corrected electron addition energies and the DFT-calculated relaxation energies are shown in tables 4.2 and 4.3.

(i)

Defect	q	R_q	A
V_{Ga}	0	R_0	0.23
	+1	R_{+1}	0.07
Ga_{As}	0	R_0	0.41
	+1	R_{+1}	0.11
As_{Ga}	+1	R_{+1}	0.28

(ii)

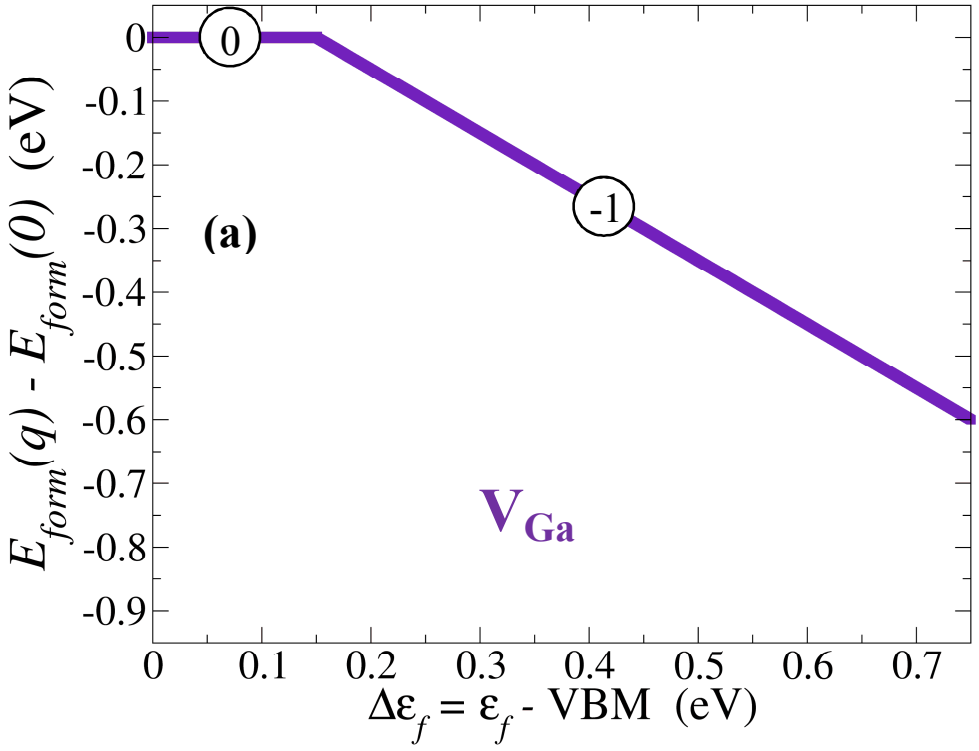
Defect	q	R_q	I
Ga_{As}	0	R_0	-0.08
As_{Ga}	0	R_0	-0.39
	+1	R_{+1}	-0.65

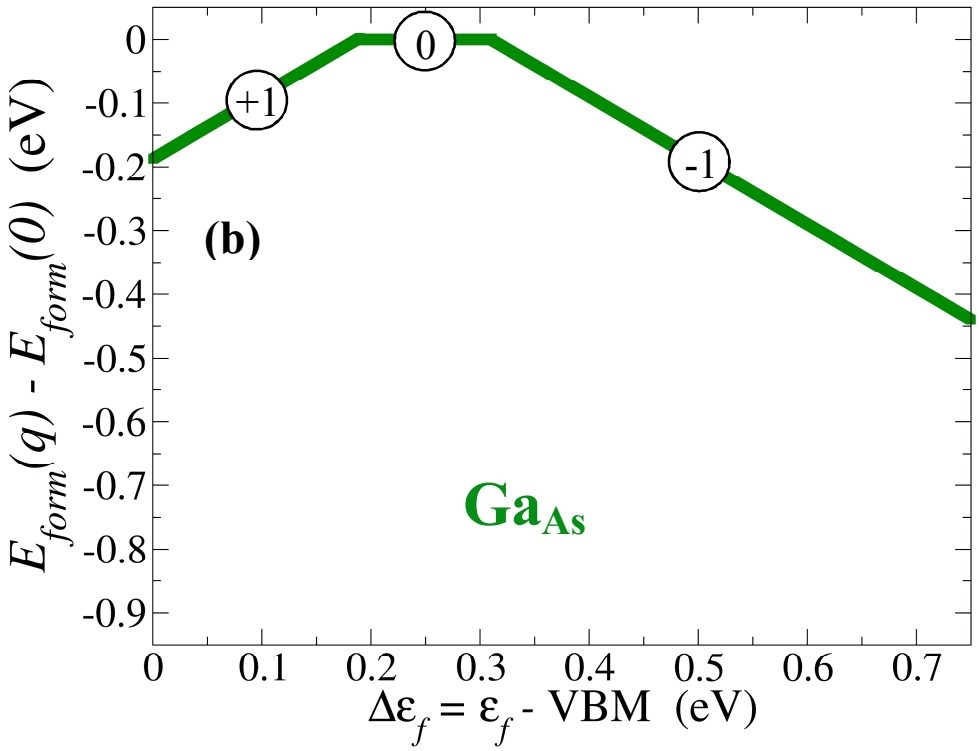
Table 4.2 - Table (i) shows the electron addition energies (A) obtained using the GW approximation as described in section 3.6 and used to calculate the formation energies of neutral and negative charge states. Table (ii) shows the electron removal energies (I) obtained using the GW approximation as described in section 3.6 and used to calculate the formation energies of positive charge states. Charge state is given by q, and R_q denotes the atomic configuration of the q^{th} charge state.

Defect	Δ		(eV)
V_{Ga}	$E(q_{-1}, R_{-1}) - E(q_{-1}, R_0)$	=	-0.020
	$E(q_0, R_0) - E(q_0, R_{+1})$	=	-0.010
Ga_{As}	$E(q_{-1}, R_{-1}) - E(q_{-1}, R_0)$	=	-0.045
	$E(q_0, R_0) - E(q_0, R_{+1})$	=	-0.050
	$E(q_{+1}, R_{+1}) - E(q_{+1}, R_0)$	=	-0.053
As_{Ga}	$E(q_0, R_0) - E(q_0, R_{+1})$	=	-0.096
	$E(q_{+1}, R_{+1}) - E(q_{+1}, R_0)$	=	-0.086
	$E(q_{+2}, R_{+2}) - E(q_{+2}, R_{+1})$	=	-0.108

Table 4.3 - Charge state relaxation energies calculated within DFT. The magnitude of Δ reflects the degree of structural relaxation. The values shown in this table are added to the electron addition (removal) energies of table 4.2 to obtain the formation energies of neutral and negative (positive) charge states. Note in all cases Δ consists of total energy terms involving the same charge state.

In order to investigate the calculated position of transition levels with respect to peaks in measured D_{it} for $In_{0.53}Ga_{0.47}As$/oxide interfaces [6], figure 4.7 presents the formation energies plotted against an energy scale set by the room temperature experimental bandgap 0.75 eV. The first graph represents the Ga vacancy, followed by the Ga_{As} antisite, and finally the As_{Ga} antisite. In figure 4.7 the formation energy of each charge state is shown relative to the formation energy of the neutral charge state. Charge transition levels correspond to values of $\Delta\varepsilon_f$ at which the slope of the lowest energy charge state changes. It appears that all defects studied here have at least one charge transition between 0.15 eV and 0.31 eV above the valence band edge. In particular, the As_{Ga} antisite has a +2/+1 transition 0.28 eV above the VBM. Within the range of theoretical accuracy, this is consistent with the midgap peak of the measured D_{it} distribution [6, 35, 38, 47].


0
-0.1
-0.2
-0.3
-0.4
-0.5
-0.6
-0.7
-0.8
-0.9
$E_{form}(q) - E_{form}(0)$ (eV)
(a)
0
-1
V_{Ga}
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
$\Delta\varepsilon_f = \varepsilon_f$ - VBM (eV)


0
-0.1
-0.2
-0.3
-0.4
-0.5
-0.6
-0.7
-0.8
-0.9
$E_{form}(q) - E_{form}(0)$ (eV)
(b)
+1
0
-1
Ga_{As}
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
$\Delta\varepsilon_f = \varepsilon_f$ - VBM (eV)

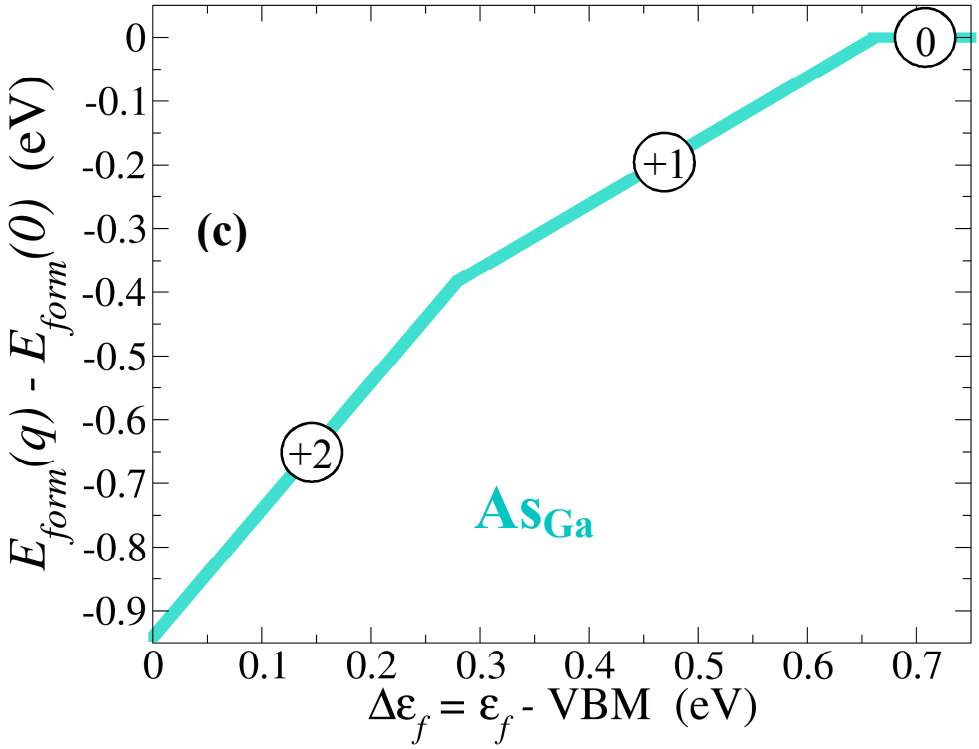


Figure 4.7 - V_{Ga} shown in (a), Ga_{As} in (b), and (c) represents As_{Ga}. For each defect, q-state formation energies relative to neutral (q = 0) formation energy are plotted against the position of the Fermi level relative to the valence band edge. Encircled numbers denote charge state.

Placing these values of $\varepsilon_{q/q'}$ on the graph reported by Komsa and Pasquarello [13], we can compare our results to the hybrid-DFT calculated quantities (figure 4.8). Komsa and Pasquarello employ the HSE hybrid functional to treat exchange and correlation effects and correct the DFT bandgap error [13, 131]; it should be mentioned that the same functional was used in section 4.2.1 to compare satellite valleys calculated using HSE and DFT+GW. Good agreement is obtained for transition levels near the valence and conduction band edges, although midgap levels are slightly lower compared to the hybrid-DFT results. Our bulk defect transition levels agree with hybrid-DFT results to within 0.1-0.2 eV.

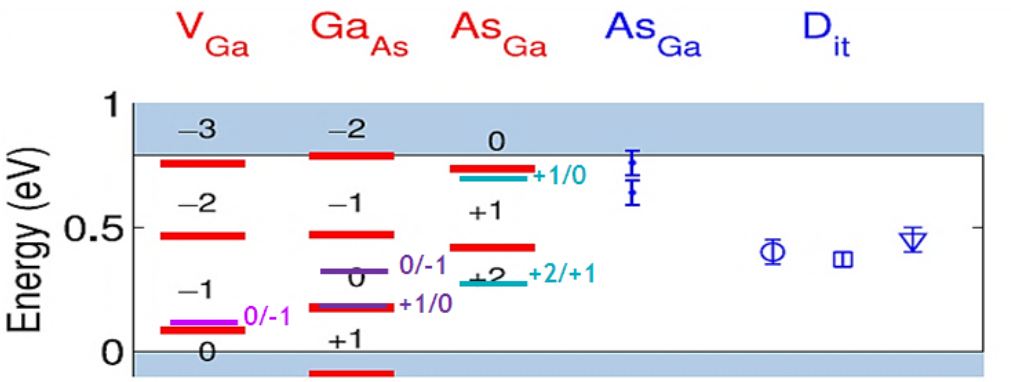


Figure 4.8 - Comparison of our DFT+*GW* calculated charge transition levels (pink, purple and turquoise lines for V_{Ga}, Ga_{As}, As_{Ga}, resp.) with transition levels calculated within hybrid DFT (reproduced from [13]). For the Ga_{As} transition levels, the defect center is surrounded by 2 In and 2 Ga atoms, *i.e.* $In_{0.5}Ga_{0.5}$ NN (see text).

We also investigated the effect of cation disorder on charge transition levels. For defects surrounded by cations, such as the Ga_{As} antisite, the local In and Ga content could have an effect on formation energies [134] and hence charge transition levels. To examine the influence of local cation disorder, the atoms surrounding the Ga_{As} antisite are replaced with either all Ga or all In, while the In_xGa_{1-x} content throughout the entire supercell is maintained at x = 0.53. When the defect is surrounded by four Ga nearest neighbors (denoted as $In_{0.0}Ga_{1.0}$ NN from here on) the 0/-1 charge transition level relative to the VBM $\varepsilon_{0/-1}(Ga_{As}) - E_v$ where E_v is the valence band maximum increases to 0.37 eV compared to 0.31 eV when there are 2 In and 2 Ga nearest neighbors ($In_{0.5}Ga_{0.5}$ NN). The $\varepsilon_{+1/0}(Ga_{As}) - E_v$ charge transition level moves much closer to the valence band, to 0.02 eV for $In_{0.0}Ga_{1.0}$ NN. These changes, relative to the case of $In_{0.5}Ga_{0.5}$ NN, occur mainly through a decrease by an average of 76 meV in the magnitude of the electron addition energies in the positive charge state, $A(q = +1, R_{q=0})$ and $A(q = +1, R_{q=+1})$, while the electron addition energy of the neutral charge state $A(q = 0, R_{q=0})$ decreases by 32 meV. All the relaxation energies maintain their values to within 10 meV compared to the values for $In_{0.5}Ga_{0.5}$ NN. When the Ga_{As} antisite is bonded to 4 In atoms ($In_{1.0}Ga_{0.0}$ NN) the $\varepsilon_{0/-1}(Ga_{As}) - E_v$ transition level again increases by ~50-60 meV compared to the $In_{0.5}Ga_{0.5}$ NN case, to 0.37 eV. The +1/0 charge transition level $\varepsilon_{+1/0}(Ga_{As}) -$

E_v = 0.16 eV for $In_{1.0}Ga_{0.0}$ NN which is 30 meV lower than the $In_{0.5}Ga_{0.5}$ NN case. While the 0/-1 charge transition level is very close for the two cases $In_{0.0}Ga_{1.0}$ NN and $In_{1.0}Ga_{0.0}$ NN, the latter exhibits a slightly increased value of $A(q = 0, R_{q=0})$ (by 11 meV) compared to $In_{0.5}Ga_{0.5}$ NN, while the value of $A(q = +1, R_{q=+1})$ decreases by 58 meV relative to $In_{0.5}Ga_{0.5}$ NN and $A(q = +1, R_{q=0})$ increases by 11 meV. Again, all relaxation energies for $In_{1.0}Ga_{0.0}$ NN change by less than 10 meV relative to $In_{0.5}Ga_{0.5}$ NN. The largest change in the charge transition levels of this defect as a function of local cation disorder occurs for the +1/0 transition level which is 0.02 eV above the VBM when bonding to 4 Ga atoms, compared to 0.19 eV above the VBM when bonding to 2 In and 2 Ga. Local cation disorder seems to have a greater effect for donor-like transitions of Ga_{As}, while the acceptor-like transitions change by less than 60 meV.

These findings could have consequences for the donor-like feature of the experimental D_{it} which extends into the valence band (see figure 4.9 on page 99). Local cation disorder can shift this donor-like transition state within a range of ~100-200 meV, yet always remain within the lower half of the bandgap (see figure 4.9). Therefore one could conclude that the D_{it} feature extending into the valence band may have broadening contributions from variable bonding arrangements of anion-situated defects due to the random cation alloy.

4.3.4 Conclusions

In this chapter, calculations of the pristine bulk InGaAs electronic structure were carried out using LDA-DFT supplemented by the GW approximation which corrects for the intrinsic limitations of the LDA and the KS-DFT formalism. Using a two-atom VCA model to test the DFT+GW approach, good qualitative agreement is obtained for the bandstructure when comparing DFT+GW to empirical pseudopotential calculations, and reasonable quantitative agreement is obtained between hybrid-DFT and DFT+GW, as compared to the position of the Γ, X and L satellite valleys. This provides a good preliminary comparison of the

efficacy of these methods for calculating charge transition levels, as satellite valleys play a significant role in the filling/emptying of states at the semiconductor surface as a function of gate voltage [130, 135]. Moving to a more sophisticated 64-atom model in order to explicitly account for the InGa random alloy sublattice, the small overestimation of the bandgap obtained using the VCA is largely eliminated and we find excellent agreement with the experimental low temperature bandgap.

In the next section, structural and electronic properties of point defects are presented. To the best of our knowledge, this is the first application of the GW approximation to point defects in bulk $In_{0.53}Ga_{0.47}As$. We find semi-quantitative agreement between our DFT+GW calculations and results obtained using hybrid functionals [13]. Good agreement is seen for the shallow charge transitions, while the deeper midgap transitions are lower with respect to the hybrid-DFT results by up to 0.16 eV. Thus, our results are largely consistent with the conclusion of ref. [13] in which the As antisite was assigned as the most likely physical origin of observed midgap D_{it} peak. However, a consistent comparison of our bulk results with a surface model, in particular a surface terminated with a high-k oxide will be used to explore this conclusion. Isolated As dimers have also been invoked [43] to explain observed interface state density profiles in GaAs; such a defect is not suitable for modeling in a bulk simulation cell, which further motivates a study involving a surface model.

All of the defects studied in this work exhibit charge transition energies in the vicinity of the experimental D_{it} feature which extends into the valence band. In addition, our calculations show that local alloying has a significant effect on defect levels. These results raise the question as to whether a range of different defects, including defects affected by bonding disorder due to the cation alloy, could be responsible for this part of the measured data. The charge transition levels calculated in this thesis are compared to the measured interface state density [6] in figure 4.9.

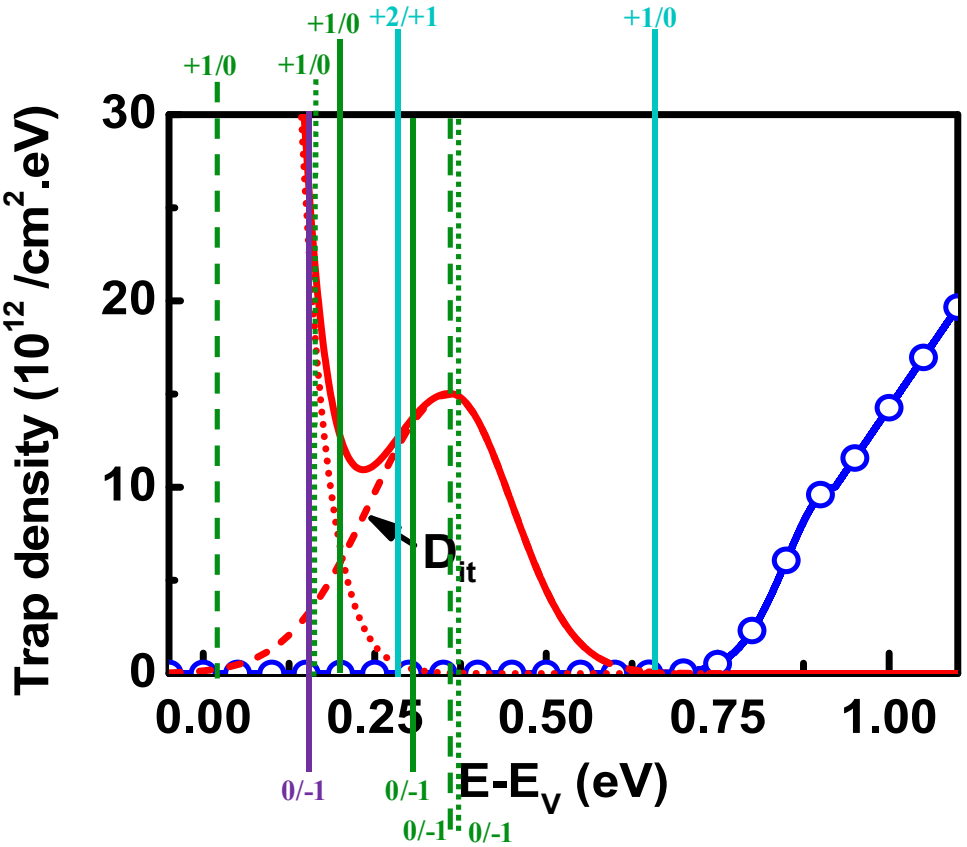


Figure 4.9 - Comparison of our DFT+*GW* calculated charge transition levels (green, purple, and turquoise lines for Ga_{As}, V_{Ga}, As_{Ga}, resp.) with experimentally inferred distribution of interface states. Experimental data and fits to data reproduced from [6]. For Ga_{As}, solid lines denote $In_{0.5}Ga_{0.5}$ NN, dashed lines denote $In_{0.0}Ga_{1.0}$ NN, and dotted lines denote $In_{1.0}Ga_{0.0}$ NN. All charge transition levels are given with respect to E_v.

The black curve is inferred from measurements [6] taken on an MOS stack containing an $In_{0.53}Ga_{0.47}As$-Al_2O_3 interface. While these measurements were unable to ascertain the nature of defects giving rise to the peaks, comparisons with other experimental works [35, 38] indicate they are independent of the deposited oxide and are thus likely associated with the substrate. What is still unclear however, is whether a single defect or multiple defects are responsible, and whether these defects are bulk-like or surface-like, *i.e.* far enough away from the oxide such that the local chemical environment can be considered as a bulk material, or close enough to the surface that the termination of bulk charge balance by bonding to the oxide has an effect on the transition levels. Moreover

for a given defect, what is the sensitivity of $\varepsilon_{q/q'}$ to proximity to the oxide interface? To investigate these issues we will present a computational study of a slab model consisting of (100) $In_{0.53}Ga_{0.47}As$ passivated with Al_2O_3 (chapter 5).

Chapter 5
InGaAs Surface Passivated with Al_2O_3

5.1 Introduction

The metal-oxide-semiconductor (MOS) capacitor consisting of an InGaAs semiconductor and a high-*k* oxide has been the subject of intense research in recent years [6, 30, 35, 47, 136]. The performance of MOSFETs is significantly affected by the presence of defects at the semiconductor-oxide interface, and the MOS capacitor is an ideal system for studying the sensitivity of MOSFETs to these interfacial defects [137]. These defects may consist of charge-trapping, recombination, as well as compensation centers. Thus, identifying the physical nature of such defects, both their electronic and atomic structure, is crucial to optimizing the electrical performance of devices and remains one of the goals of first-principles methods.

In this chapter, a model of the (100) oriented $In_{0.53}Ga_{0.47}As$ surface passivated with Al_2O_3 is presented and utilized to calculate the experimentally detectable charge transition levels of point defects. Comparing with the results of chapter 4, this study provides a means to compare and relate properties of defects in a bulk chemical environment with those either within a few bond lengths from or bonding directly to the gate dielectric. Such a study is highly noteworthy since in MOS stacks the band edges relative to the Fermi level of the semiconductor can change when moving from the bulk to the surface. The bonding arrangement of the semiconductor bulk is terminated at the surface, and the surface termination by necessity depends on a large number of conditions: surface preparation techniques, cleavage orientation, the oxide material, post-deposition processing, just to name the most significant influences. Hence, it is easy to appreciate that many important properties such as the distribution of defect states, their atomic structure, and of course the charge transition levels

can vary widely between a given defect located in the bulk or at the semiconductor/oxide interface. Assigning a particular charge transition level calculated in a bulk model to a measured D_{it} feature is thus incomplete without comparison to a semiconductor/oxide interface model.

5.2 Passivation Mechanisms

In contrast to Si or Ge, III-V compound semiconductors exhibit a polarity in their bonding arrangement; group V anionic species such as As exchange charge with group III cations such as Ga or In to achieve a charge balance within the bulk of the material. Cleaving along certain crystallographic planes such as (100) results in termination of the bulk structure with a monolayer of either cations or anions. Therefore these surfaces are polar in nature, and the excess negative or positive charge drives the creation of electrically active defects in order to return the Fermi level to within the bandgap [138]. In other words, for such semiconductors, polarity driven surface reconstruction occurs in the absence of a stable passivating material [9, 139]. Unlike Si which has 1 electron per sp^3 hybrid orbital, the cations formally have 3/4 and the anions 5/4 of an electron per orbital. Complex reconstructions occur at (100) III-V surfaces as a result of these different electron occupancies.

As discussed by Robertson and Lin, this surface reconstruction can be seen as an auto-compensation response to the electron-counting rule [9]. Using GaAs as an example, the electron-counting rule says that at (100) oriented surfaces all Ga dangling bonds must be completely empty, all As dangling bonds must be completely full, and defects will be created to establish a neutral surface, ensuring the Fermi level lies in the bandgap; *e.g.* if the surface "starts" with an excess of electrons, acceptor defects will be created to receive the excess electrons and bring the Fermi level out of the conduction band and into midgap. Robertson and Lin argue that this is the mechanism which drives the formation of gap states during the deposition and passivation processes. To achieve

insulating, abrupt interfaces with oxides, it is thus necessary to use an oxide with an electron count that corresponds to the monolayer that terminates the III-V surface.

Following ref. [9], we find that the trivalent oxide Al_2O_3 exhibits this property; the electron count of the O layer in Al_2O_3 is the equivalent to that of the As layer in $In_{0.53}Ga_{0.47}As$, and the Al layer in Al_2O_3 has the same electron count as the cation layer in $In_{0.53}Ga_{0.47}As$. The remaining O dangling bonds can then be passivated with H, and a neutral, insulating, atomically abrupt interface with InGaAs can be achieved. We calculate relaxed slabs of As-rich (100) $In_{0.53}Ga_{0.47}As$ passivated with the Al_2O_3 structure proposed by Robertson and Lin, an example of which is shown in figure 5.1 below (methods and numerical details are described in the next section). An As-rich model is selected for consistency with experimental conditions which are tuned to achieve As-rich (100) $In_{0.53}Ga_{0.47}As$ [6, 30, 140]. To emulate the continuation of bulk charge transfer, we terminate the bottom As-terminated slab face with pseudo-hydrogens consisting of noninteger nuclear charge of 0.75.

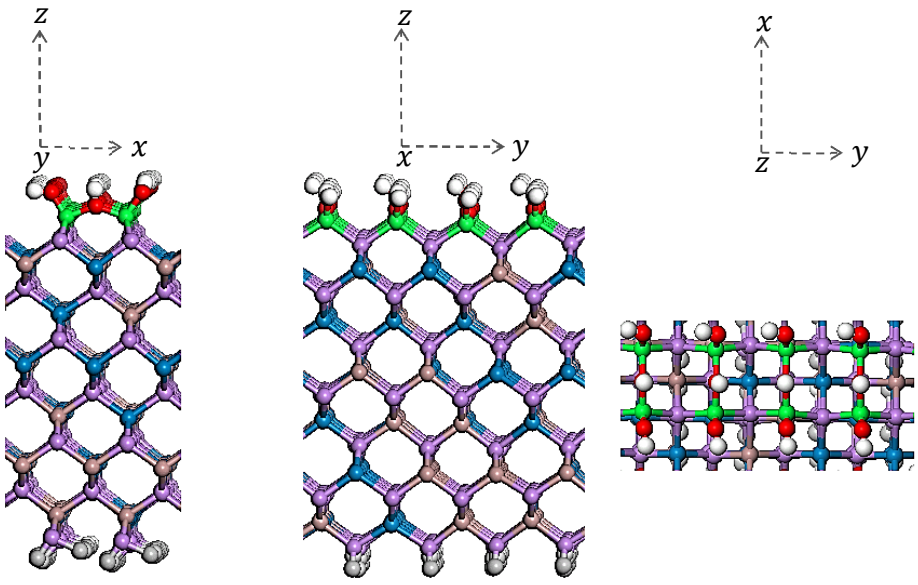


Figure 5.1 - Relaxed slabs of (100) InGaAs passivated with Al_2O_3, consisting of 152 atoms. x, y, and z supercell directions are shown at the top. White, green, and red spheres indicate H, Al, and O. Pink, blue and brown spheres indicate As, In, and Ga. Grey spheres on the bottom are for pseudo-hydrogens with valence = 0.75.

The graph below compares the local density of states of the Al_2O_3 layer with that calculated by Robertson and Lin [9] for (100)GaAs:Al_2O_3; despite the comparison between different semiconductors, as well as different calculation methods (ultrasoft PBE-GGA and plane wave cutoff energy ~28 Rydbergs used by Robertson and Lin [9], we use norm-conserving PZ-LDA and 60 Rydbergs), our model also exhibits no states within the bandgap, in qualitative agreement with ref. [9].

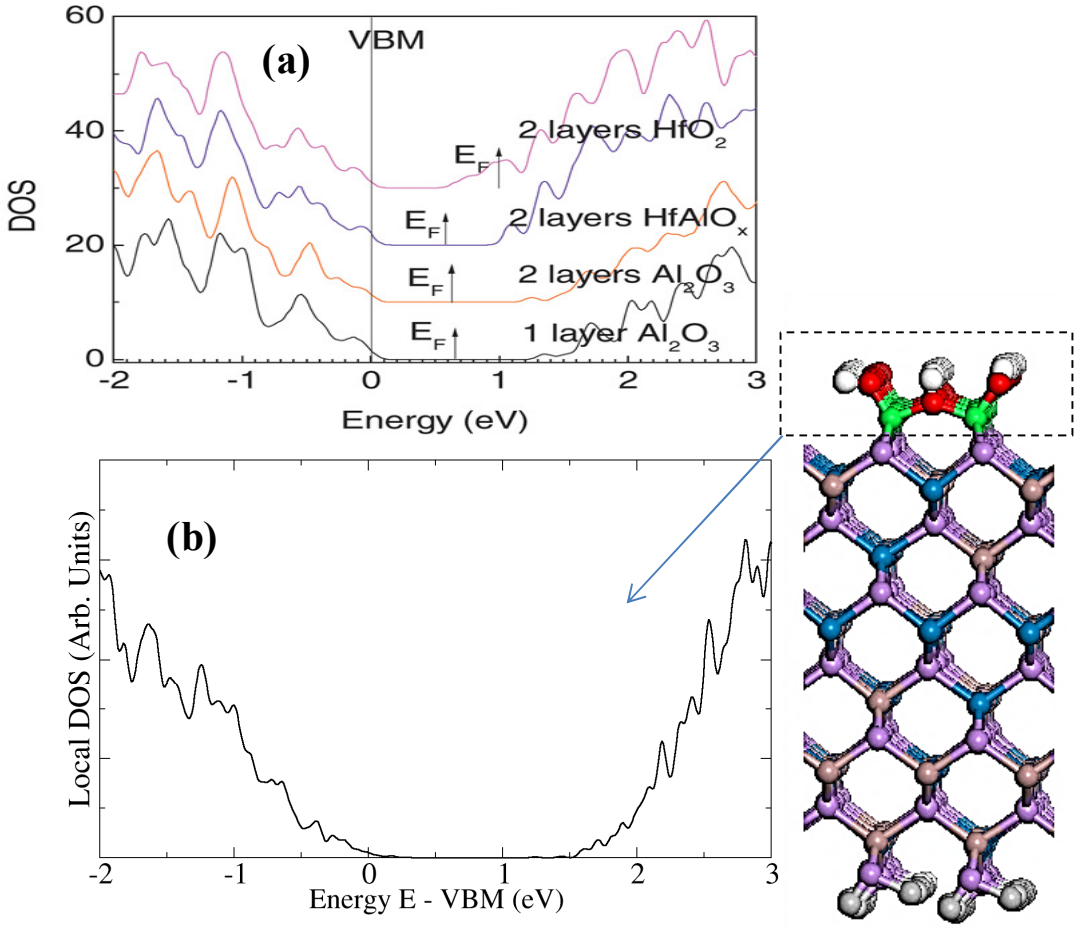


Figure 5.2 - DOS of one layer of Al_2O_3 bonding to (100) GaAs calculated by Robertson and Lin (bottom black curve of (a)) compared with our calculated DOS of one layer of Al_2O_3 bonding to (100)$In_{0.53}Ga_{0.47}As$ (b). Overall qualitative agreement is obtained. Top graph (a) reproduced from ref. [9].

5.3 Defects - As_2, Ga_{As} and As_{Ga}

This section describes the computational modeling of point defects incorporated into Al_2O_3 passivated (100) $In_{0.53}Ga_{0.47}As$. This work is an attempt to understand the atomic structure of defects giving rise to measured CV responses at InGaAs-high-*k* oxide interfaces. Point defects native to InGaAs were chosen due to the experimental indication that the measured responses are independent of the oxide [6, 35, 38]. The As dimer, As_2, provides a useful comparison with recent computational studies of InGaAs surfaces using hybrid-DFT [32]. Using the results of the previous chapter, the antisites allow for a comparison of transition levels between bulk-like and surface-like defects. Initially, defects were placed on the top monolayer and bonded directly to Al atoms at the surface. This resulted in charge transition levels resonating with the host bands. However, moving defects down from the interface by just a few monolayers results in a dramatic shift of the transition levels bringing them into the bandgap and much closer to the corresponding bulk transition levels.

The size of supercells adopted here for the structural optimization of defects allows for the use of Zunger's special quasi-random structures (SQS) [141, 142] to define the cation sublattice, leading to accurate defect properties such as atomic structure and energy differences. The 64 atom supercells used in the previous chapter are too small to adopt the special quasi-random structure while also maintaining a cation composition of 53% In and 47% Ga, thus a random configuration is chosen in that case. Fortunately, the 304 atom supercells used for the surface model allows for the adoption of SQS-8 while also maintaining the desired cation composition. Also, as described in chapter 3 the expression for charge state formation energies indicates that the error associated with electrostatic interaction of point defects may cancel; specifically, in the relaxation energy term Δ which consists of differences between simulation cell energies involving the same charge state. We extrapolate the energies of charged cells to infinite defect separation to show that these interactions are practically

negligible compared to elastic effects, which in turn are minimized in the supercells used to calculate Δ. As in the case of bulk defects, the GW approximation is utilized here to calculate the electron addition energy required to obtain the charge state formation energy.

5.3.1 Methods

For the structural relaxation of point defects, we employ 304 atom supercells (208 InGaAs, 40 Al_2O_3, 24 H and 32 pseudo-H) in which periodic images of defects are separated by 16 Å in the x and y directions. Periodic images of slabs are separated by approximately 18 Å of vacuum equivalent to almost 13 monolayers in the z direction, while 13 monolayers of InGaAs separate each slab face (figure 5.1, page 103). These distances minimize the interaction between periodic images of defects while maintaining a reasonable computational expense. Relaxation of the atomic geometries plus the supercell is performed before inserting point defects in order to reduce errors due to spurious elastic effects [5]. Following the full structural relaxation of the pristine cell, point defects are inserted and the atomic geometries re-optimized.

The approach of special quasi-random structures (SQS) developed by Wei, Zunger and co-workers [141] is utilized for the cation sublattice. The SQS method allows one to emulate as closely as possible the statistical correlations of an infinitely large random alloy of $In_{0.53}Ga_{0.47}As$ for the cation distribution using a finite supercell. To achieve this, Zunger *et al* envisage a binary substitutional alloy with N sites occupied by one of two atoms, atoms A or B. Defining a configuration as one of the 2^N distinct arrangements of this (sub)lattice, they calculate configurational averages of lattice properties to find those configurations whose properties best match those of the ensemble average of a random alloy distribution [141]. By stacking the (113) cation planes in the order of $2_A1_B2_A3_B$ (see figure 5.3, page 107 for clarification of notation) where A denotes In atoms and B denotes Ga, as shown in figure 5.3, the "SQS-8" configuration is achieved [141].

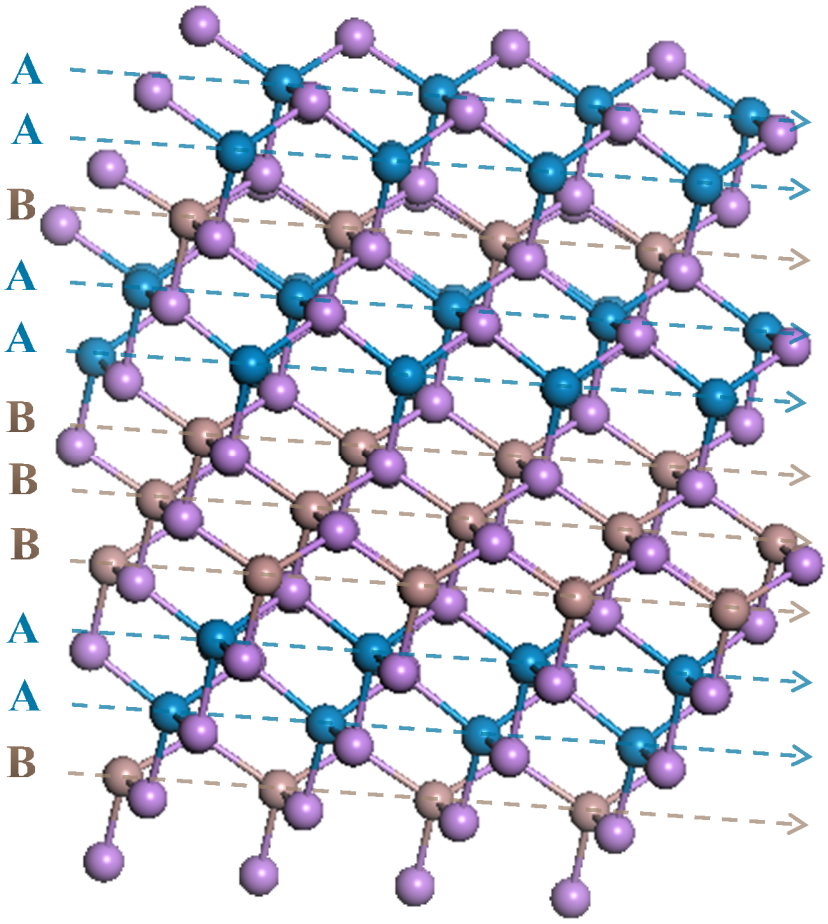


Figure 5.3 - InGaAs surface structure oriented to show the stacking arrangement of (113) cation planes. The latter are ordered as prescribed in the SQS-8 configuration [141]. A indicates In atoms and B indicates Ga atoms. Al_2O_3 and pseudo-hydrogen passivation are omitted for clarity.

Norm-conserving pseudopotentials, the PZ-LDA form of the exchange-correlation functional [89, 123], 60 Rydberg kinetic energy cutoff, and a 2x2x1 mesh for the $k_x \times k_y \times k_z$ grid of k-points (vacuum in the z-direction) are used for the optimization of atomic positions within DFT. The GW calculations involve 1772 unoccupied states and 5041 elements in the plasmon-pole approximated dielectric response matrix $\varepsilon_{\mathbf{G},\mathbf{G}'}(\mathbf{q}, E)$, and electron addition energies are calculated as the GW-corrected separation between highest occupied and lowest unoccupied states at the Γ-point. Due to the computational demand, 152 atom supercells (see figure 5.1) of half the width of the 304 atom cells were used to compute the GW-correction; smaller supercells for the electron addition energy

term have been used previously [22] when employing the GW-correction for both neutral and charged cells. The justification for the use of the smaller cell being the lack of electrostatic (Hartree) contributions to the self-energy correction Σ, *i.e.* the latter only involves exchange and correlation terms so the supercell size should not significantly affect Σ.

As was described in chapter 4 for the case of a bulk simulation cell, we also ensure the validity of the plasmon pole approximation (PPA) to the dielectric function for the case of the Al_2O_3 passivated (100) $In_{0.53}Ga_{0.47}As$ surface. This is done by plotting the dielectric response as a function of energy [50] as shown in figure 5.4. The single peak of the imaginary part of the inverse dielectric function at the plasmon energy reveals the applicability of the PPA to this system [50].

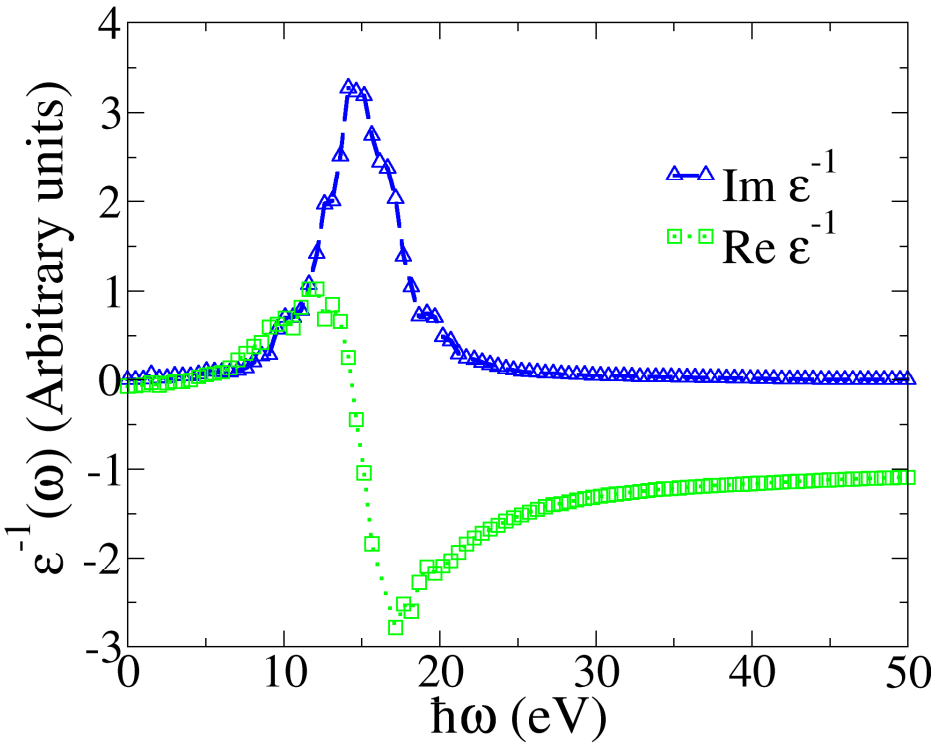


Figure 5.4 - Real and Imaginary parts at the Γ-point component ($G = G' = 0$) of the inverse dielectric function, obtained from the Al_2O_3 passivated (100) $In_{0.53}Ga_{0.47}As$ surface model, plotted as a function of energy.

5.3.2 Results

Before discussing the structural and charge transition properties of surface defects, we note that the charge state formation energies contain a term involving the energy difference between simulation cells that have the same charge state. Thus errors associated with electrostatic interactions may systematically cancel. This has been suggested in a review by Van de Walle and Neugebauer [5] in their discussion of the accuracy of charge transition levels. In order to investigate this suggestion, we take the Ga_{As} defect at the -1 charge state as an example, and in figure 5.5 plot the difference in supercell energies $\Delta = E(q_{-1}, R_{-1}) - E(q_{-1}, R_0)$ versus the inverse of the square root of the supercell surface area $S^{-1/2}$ where $S = x_{cell}y_{cell}$ and x_{cell} and y_{cell} are the dimensions of the supercell parallel to the surface plane. To this end four (100) $In_{0.53}Ga_{0.47}As$:Al_2O_3 supercells are constructed, consisting of 72, 114, 152, and 304 atoms corresponding to the four data points on each graph of figures 5.5, 5.6, and 5.7. The latter two are plots of the energy differences for the neutral cell relaxations $\Delta = E(q_0, R_0) - E(q_0, R_{-1})$ and $\Delta = E(q_0, R_0) - E(q_0, R_{+1})$, versus $S^{-1/2}$. In these cases the energy differences are between two uncharged cells so there is no spurious electrostatic component.

Comparison between figures 5.5, 5.6, and 5.7 reveals a qualitatively similar variation of the relaxation energy as a function of supercell size for all cases of charged and uncharged supercells. In fact, the neutral cell relaxation corresponding to a transition from the negative state (figure 5.6) yields the largest variation as a function of surface area. This indicates that the spurious Coulombic interactions between periodic defect images in charged supercells are negligible compared to elastic interactions. The latter are, within the computational constraints, minimized for the largest supercell due to the large separation (16 Å) between defect images. Given the accepted range of error of ~100 meV for these types of calculations [3, 5, 32, 44], in addition to the typical error bar (~50-100 meV) of the experiments with which we are comparing our

results [47, 143], the large computational expense associated with extrapolating the energies of all charge states of all defects as performed here is unjustified. Therefore, as suggested by the recent literature [5] we omit the electrostatic correction term from the formation energy assuming a cancellation between the electrostatic components of the two total energies of the equally charged cells. Due to the lack of a significant impact of both Coulombic and elastic interactions for the largest surface area (corresponding to a 304 atom supercell), this supercell is adopted for the calculation of the relaxation energy term of the formation energies (see section 3.6) of all surface defects. The partitioning of formation energies into an electron addition (GW) component and a geometry relaxation (DFT) component enables the use of like-charge cells for the latter term, resulting in a fortuitous cancellation of the charged defect-charged defect interactions within the periodic cells.

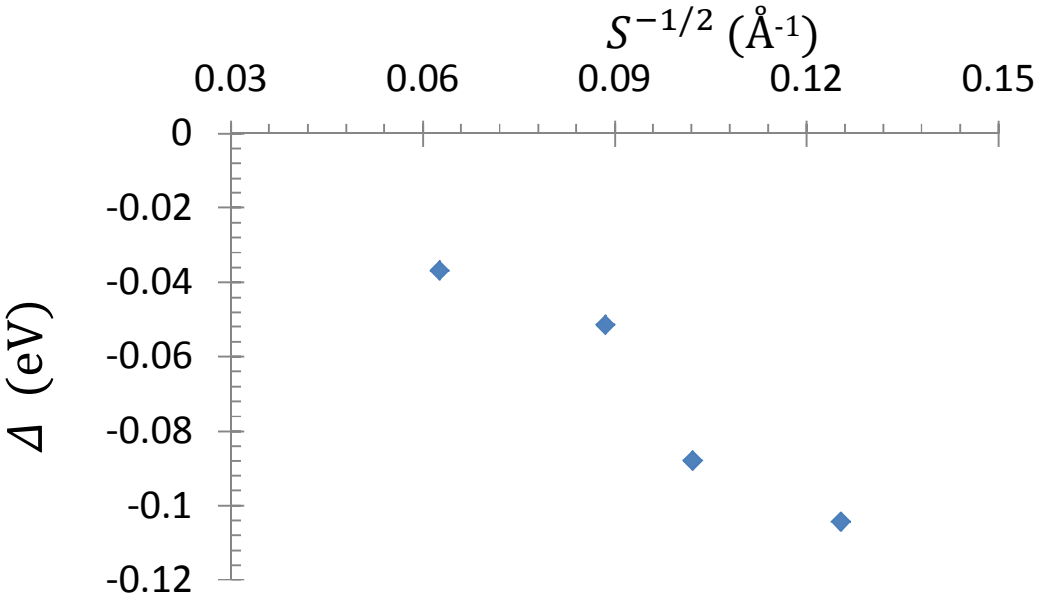


Figure 5.5 - Energy difference $\Delta = E(q_{-1}, R_{-1}) - E(q_{-1}, R_0)$ versus $S^{-1/2}$. Variation in this value from the smallest to largest supercell is 68 meV.

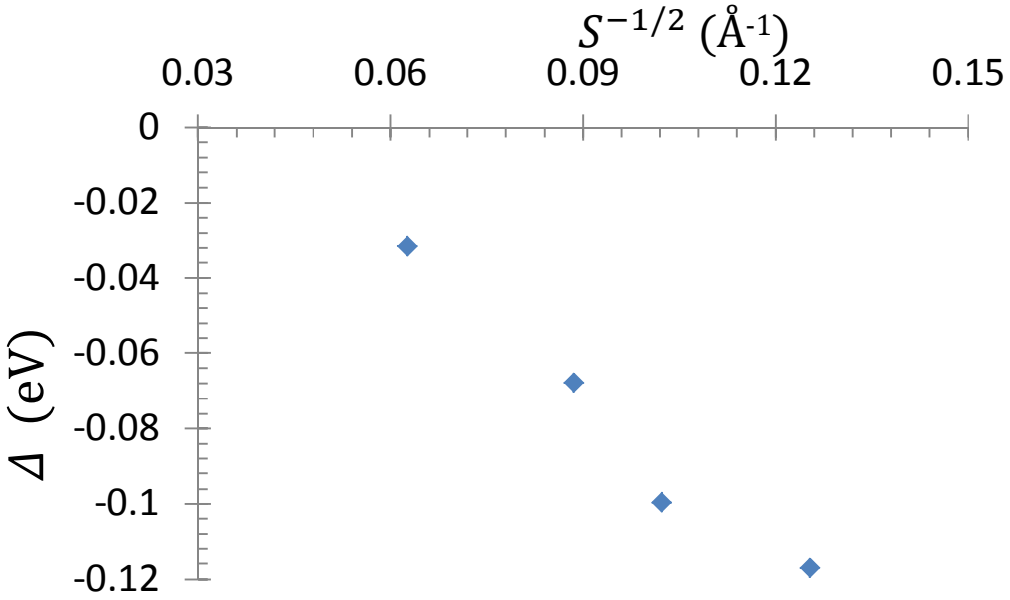


Figure 5.6 - Energy difference $\Delta = E(q_0, R_0) - E(q_0, R_{-1})$ versus $S^{-1/2}$. Variation in this value from the smallest to largest supercell is 85 meV.

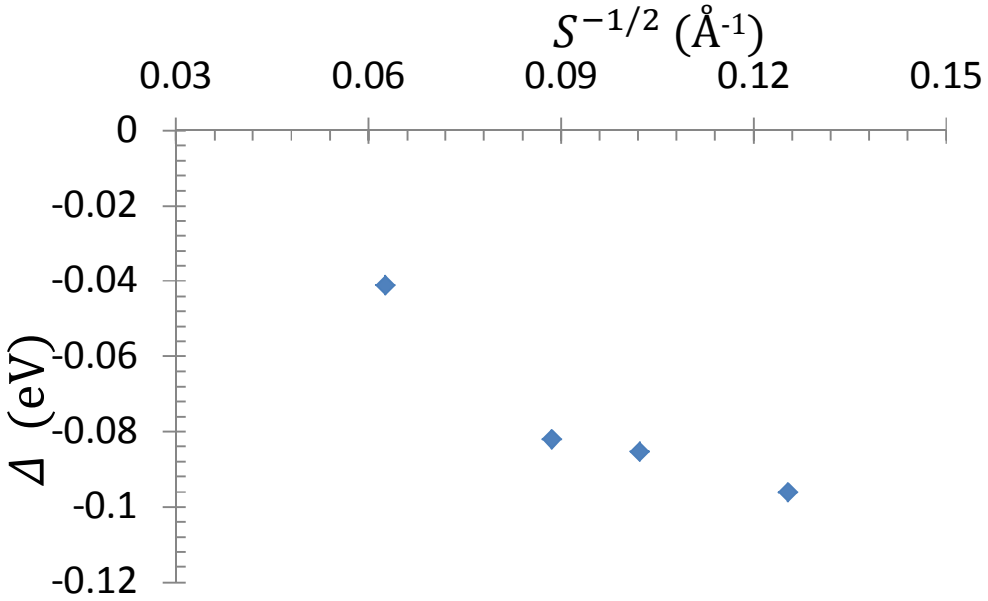


Figure 5.7 - Energy difference $\Delta = E(q_0, R_0) - E(q_0, R_{+1})$ versus $S^{-1/2}$. Variation in this value from the smallest to largest supercell is 55 meV.

5.3.2.1 Structural Properties

Focusing on the As dimer defect first, we create this defect using the model of Miceli and Pasquarello [32]. Displacing the oxide atoms in the (110) direction allows for the formation of two As-As and two Al-Al bonds. An O atom is inserted between both Al-Al bonds and one of the As-As bonds leaving a single As-As bond at the semiconductor-oxide interface (figure 5.8 (a). page 114). Relaxing this defect in the neutral charge state, $q = 0$, results in an As-As bond length of 2.56 Å. Charging to $q = -1$ increases the As-As bond length to 2.98 Å. This is in good agreement with the work of Miceli and Pasquarello in which an As-As bond length of 2.56±0.01 Å for $q = 0$ was reported, which increased to 2.95±0.03 Å after relaxing in the negative charge state [32]. Charging to $q = +1$ results in a small reduction to 2.54 Å in the As-As bond length compared to the $q = 0$ case. Thus a much larger energy difference is expected for the neutral to negative relaxation $\Delta(R_{-1}, R_0, q_{-1})$ than the neutral to positive relaxation $\Delta(R_{+1}, R_0, q_{+1})$.

Turning to the Ga_{As} defect, we find that in the neutral charge state the defect moves up towards the oxide relative to the unrelaxed substitutional site with an tendency to move away from Al and towards the nearest OH group (see figure 5.8 (b)); the relaxed Ga_{As}-OH and Ga_{As}-Al separations are 2.12 Å and 2.59 Å respectively, compared to the corresponding As-OH and As-Al distances which average to 3.42 Å and 2.42 Å, respectively. In analogy to the bulk Ga antisite, the surface Ga antisite moves towards the nearest Ga neighbor and away from the nearest In neighbor relative to the pristine As site. At variance with the bulk case however, the relaxed surface Ga_{As}-Ga distance is 0.07 Å less than the corresponding surface As-Ga bond length with a 0.08 Å reduction in the corresponding bond length for the bulk Ga_{As} case. The relaxed surface Ga_{As}-In distance increases by 0.16 Å relative to the corresponding surface As-In separation compared to a 0.04 Å increase of the same bond length in the bulk case. Charging to $q = +1$ results in a further contraction of the Ga_{As}-OH distance to 2.09 Å, while the Ga_{As}-Ga and Ga_{As}-In separations both increase by

an average of 0.02 Å (1%) relative to the neutral case. Adding an electron to the neutral state to bring q to -1 results in a Ga_{As}-OH distance of 2.13 Å, a slight increase relative to $q = 0$. The Ga_{As}-Ga and Ga_{As}-In separations both decrease by an average of 0.02 Å relative to $q = 0$; *i.e.* the $q = -1$ relaxation is almost the same magnitude but opposite direction compared to $q = +1$. These relaxations occur mainly through a movement of the defect center towards OH and away from the cations for $q = +1$, and away from OH and towards the cations for $q = -1$.

The surface As_{Ga} is also qualitatively similar to bulk As_{Ga} in terms of structural rearrangement (figure 5.8 (c)). The As_{Ga}-As bonds are on an average 0.12 Å greater than the pristine Ga-As bonds. Removing an electron and relaxing in the positive charge state results in 0.02 Å reduction in the As_{Ga}-As bond lengths relative to $q = 0$, and removing another electron to bring q to +2 yields a further 0.02 Å shortening of As_{Ga}-As bond lengths relative to $q = +1$. As these charged relaxations occur through a movement of the four surrounding anions and an accompanying local displacement of the oxide while the defect site remains more or less fixed, larger Δ terms are expected for the As antisite than for the Ga antisite.

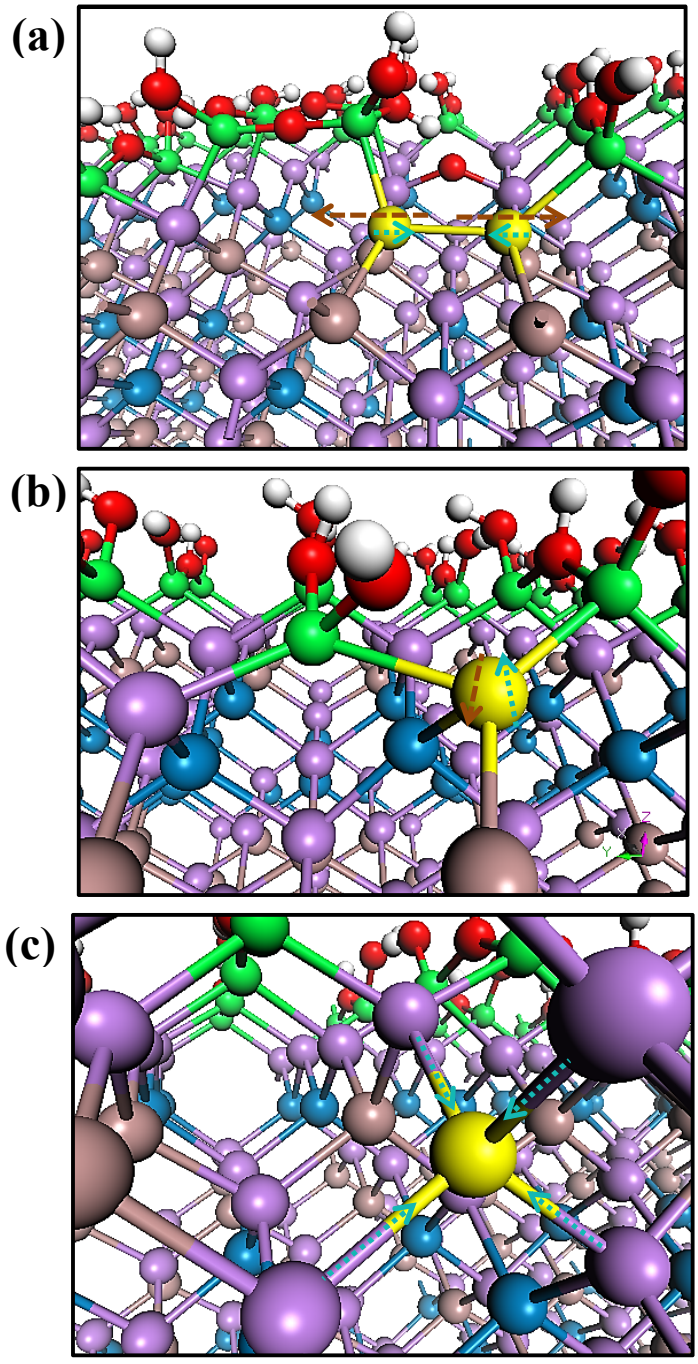


Figure 5.8 - Relaxed structures of neutral point defects in (100) InGaAs:Al_2O_3. Blue, brown and, pink spheres represent In, Ga, and As respectively. Red, white and green spheres represent O, H, and Al, respectively. For (a), the As atoms comprising the As dimer are highlighted in yellow. For (b) and (c), the Ga antisite and the As antisite, respectively, are highlighted in yellow. Arrows indicate the direction of movement of atoms for relaxation due to positive (turquoise dotted) and negative (orange dashed) charges.

5.3.2.2 Charge transition levels

In ref. [32] the $\varepsilon_{0/-1}$ charge transition level of an As dimer inserted into a (100) GaAs-Al_2O_3 interface in which up to 2 of the nearest neighbor Ga atoms are replaced by In was studied. The authors obtain a value of $\varepsilon_{0/-1}$ = 0.61 eV above the conduction band edge, which they find to be largely independent of the local In configuration [32]. Thus, they rule out the As_2 defect as a candidate for midgap D_{it}. We calculate the $\varepsilon_{0/-1}$ and $\varepsilon_{+1/0}$ transition levels of this defect in a similar fashion to the bulk charge transition levels of chapter 4. Using the GW approximation we obtain an electron addition energy of the neutral charge state $A(q_0, R_0) =$ 2.16 eV (table 5.1, page 120). Taking the energy difference between two negatively charged cells with atomic positions optimized to neutral and negative charge state configurations as the Δ term (table 5.2, page 121),

$$\Delta(R_{-1}, R_0, q_{-1}) = E(q_{-1}, R_{-1}) - E(q_{-1}, R_0)$$

and adding this to the electron addition energy of the neutral state $A(q_0, R_0)$, we find $E_{form}(q_{-1}) = A(q_0, R_0) + \Delta(R_{-1}, R_0, q_{-1}) =$ 1.87 eV ≡ formation energy of the negative charge state at zero gate bias. The neutral charge state is formed by adding an electron to the positive state and subsequently relaxing, yielding $E_{form}(q_0) = A(q_{+1}, R_{+1}) + \Delta(R_0, R_{+1}, q_0) =$ 0.38 eV. Taking the negative of the electron addition energy of the positive charge state in the atomic configuration of q = 0, *i.e.* $-A[q_{+1}, R_0]$, as the electron removal energy of the neutral charge state, *i.e.* $I[q_0, R_0]$ (table 5.1) and adding this to the relaxation energy between the neutral and positive charge state configurations with q fixed to $+1$ *i.e.* $\Delta(R_{+1},\ R_0,\ q_{+1})$ (table 5.2) yields $E_{form}(q_{+1}) =$ -0.34 eV. The +1/0 charge transition level is obtained by taking the difference between $E_{form}(q_0)$ and $E_{form}(q_{+1})$ at zero gate bias

$$\varepsilon_{+1/0} = \frac{E_{form}(q_0,\ \Delta\varepsilon_f = 0) - E_{form}(q_{+1}, \Delta\varepsilon_f = 0)}{q_{+1} - q_0}\ ,$$

yielding a positive to neutral charge transition level of $\varepsilon_{+1/0} - E_v = 0.71$ eV where E_v is the valence band maximum (figure 5.10 (a), page 122). Taking the difference in formation energies between the negative and neutral charge states at zero gate bias

$$\varepsilon_{0/-1} = \frac{E_{form}(q_{-1},\ \Delta\varepsilon_f = 0) - E_{form}(q_0, \Delta\varepsilon_f = 0)}{q_0 - q_{-1}}$$

yields $\varepsilon_{0/-1} - E_v$ = 1.49 eV or 0.74 eV above the conduction band minimum, in reasonable agreement with ref. [32].

The formation energies and charge transition levels of the other defects are calculated in a similar manner to that described above. For the Ga_{As} antisite bonding to Al_2O_3, the relaxed positive charge state exhibits a GW-corrected electron addition energy of 0.56 eV (table 5.1). Relaxing in the neutral charge state and adding the total energy difference to $A(q_{+1}, R_{+1})$ results in a formation energy of $E_{form}(q_0) = 0.52$ eV at zero gate bias. Comparing this with the formation energy of the negative charge state reveals a neutral to negative charge transition level of $\varepsilon_{0/-1} - E_v$ = 1.45 eV. Removing an electron from the neutral charge state, taking $-A(q_{+1}, R_0) = I(q_0, R_0)$ (table 5.1), relaxing in the positive state, and taking the difference between $E_{form}(q_0)$ and $E_{form}(q_{+1})$ at $\Delta\varepsilon_f$ = 0 results in $\varepsilon_{+1/0} - E_v$ = 1.06 eV (figure 5.10 (b)). Thus, as in the case of the As dimer which is also bonding to Al_2O_3, this defect exhibits charge transition levels residing within the conduction band of $In_{0.53}Ga_{0.47}As$ and does not display defect levels which could contribute to the experimental midgap D_{it} when bonding to the oxide.

Turning now to the As_{Ga} antisite reveals the same qualitative picture as the Ga_{As} antisite and the As dimer; defects in close proximity to the oxide, either bonding directly or up to one monolayer away from the oxide, do not exhibit

levels which correlate with the experimental distribution of midgap interface states. First, we note again that in qualitative agreement with [13] we find a +2/+1 charge transition level near midgap at $\varepsilon_{+2/+1} - E_v$ = 0.28 eV for the As_{Ga} antisite located in a bulk simulation cell (see chapter 4). Comparing to the bulk case, the charge transition levels are shifted into the semiconductor host bands when the defect is in close proximity to the oxide. Specifically, $\varepsilon_{+1/0} - E_v$ = 1.27 eV while $\varepsilon_{+2/+1} - E_v$ = -0.06 eV for the As_{Ga} antisite near the oxide (figure 5.10). It should be noted that this defect does not bond directly to Al_2O_3 in the As-terminated (100) $In_{0.53}Ga_{0.47}As$ surface (see figure 5.8 (c)). The shifts in charge transition level between bulk- and surface-modeled As_{Ga} are smaller than the charge transition level shifts for the bulk- and surface-modeled Ga_{As} antisite (compare figure 5.10 on pages 122-123 with figure 4.7 on pages 94-95). The latter is bonding directly to Al and exhibits charge transition level shifts of 0.9-1.0 eV between the bulk and surface models. The diagram below (diagram 5.1) shows the difference in charge transition levels for defects in bulk and surface models.

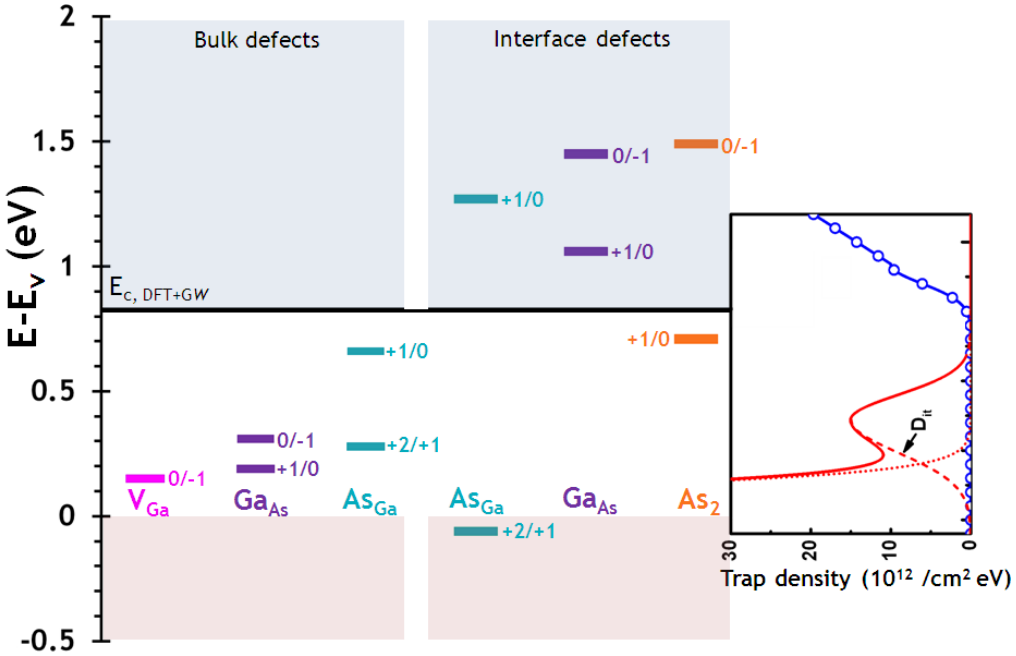


Diagram 5.1 - Comparison of charge transition levels calculated in bulk $In_{0.53}Ga_{0.47}As$ and in the Al_2O_3 passivated (100) $In_{0.53}Ga_{0.47}As$ surface. The bulk bandgap calculated within DFT+*GW* (see section 4.2) is also shown. The graph on the left represents the experimentally inferred D_{it} distribution extracted by Djara *et al* [6].

This behaviour prompts a quantitative investigation of the charge transition level as a function of distance from the oxide. The Ga_{As} antisite is chosen as the test case for this investigation, as this allows for an analysis of the effect of bonding directly to the oxide, compared to an entirely bulk-like environment. Starting from the relaxed pristine cell, an As atom is exchanged for a Ga atom in the next As layer "down" from the oxide (see figure 5.9).

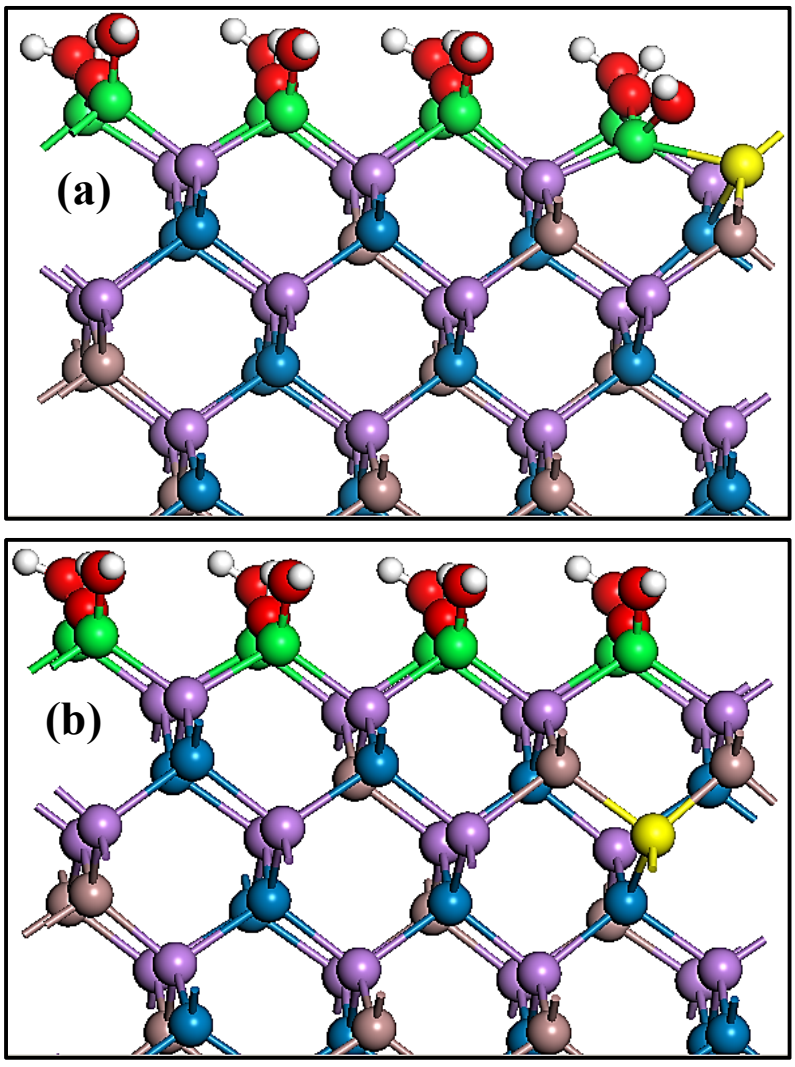


Figure 5.9 – (a) Ga_{As} antisite located on the atomic layer which binds to Al_2O_3. (b) Ga_{As} antisite located 1 As layer "down" from the oxide. For (a) the antisite is bonding to 2 Al, 1 In and 1 Ga, for (b) the antisite bonds to 2 Ga and 2 In (only 1 Ga_{As}-In bond shown here). Blue, brown and pink spheres represent In, Ga, and As respectively. Red, white and green spheres represent O, H, and Al, respectively. The Ga_{As} antisite defect is highlighted yellow in both (a) and (b).

Calculating the charge transition levels as described above, we note a significant reduction in the electron addition and removal energies (table 5.1) compared to values obtained when the antisite is bonding directly to Al_2O_3. The largest reduction of 1.73 eV occurs for the neutral charge state leading to a correspondingly large shift of the neutral to negative charge transition level, resulting in $\varepsilon_{0/-1} - E_v$ = 0.2 eV for Ga_{As} antisite formed in the 2nd As layer down from Al_2O_3 (figure 5.10, page 123). The latter is much closer to the corresponding charge transition level in the bulk case (0.31 eV, figure 4.7 (b) on page 94) compared to the value obtained when bonding to Al_2O_3. A large shift towards the valence band is also demonstrated by the positive to neutral charge transition level of the Ga_{As} antisite not bonding to the oxide compared to the case when the defect is bonding to Al_2O_3. We calculate a value of $\varepsilon_{+1/0} - E_v$ = 0.08 eV for this case (figure 5.10). While these values are closer to the corresponding values in the bulk case (0.19 eV), we consider them too close to the valence band to explain the midgap D_{it} peaks [6, 35, 38, 47] observed in experiment (see figure 5.12, page 126). Our calculations do indicate however that bonding to the oxide plays a significant role in the position of defect states, particularly those close to or inside the valence and conduction bands.

(i)

Defect	q	R_q	A
As_2	0	R_0	2.16
	+1	R_{+1}	0.39
Ga_{As} (a)	0	R_0	2.01
	+1	R_{+1}	0.56
As_{Ga}	+1	R_{+1}	0.89
Ga_{As} (b)	0	R_0	0.28
	+1	R_{+1}	0.07

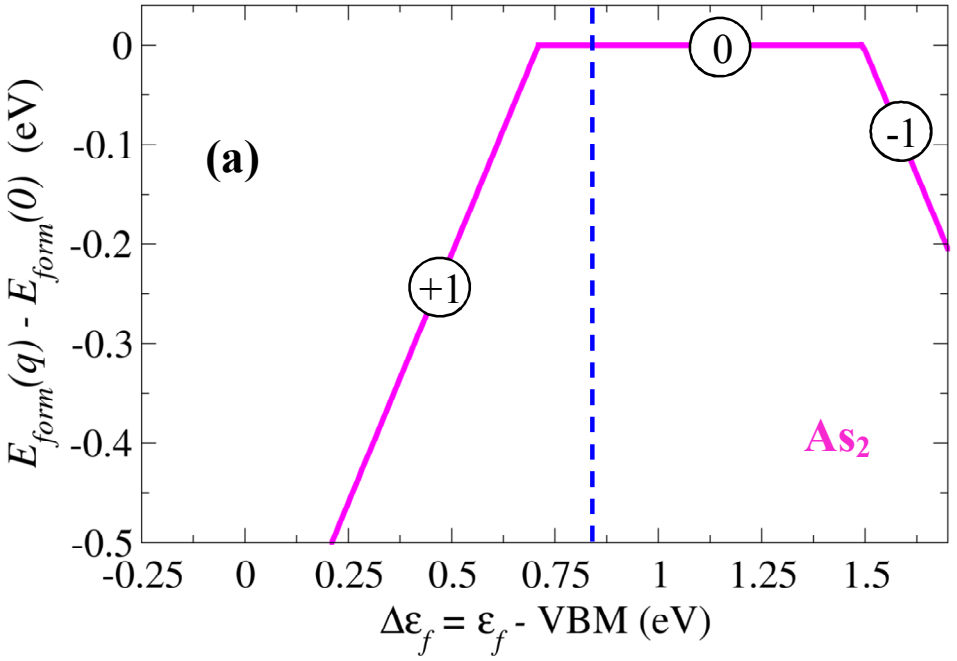
(ii)

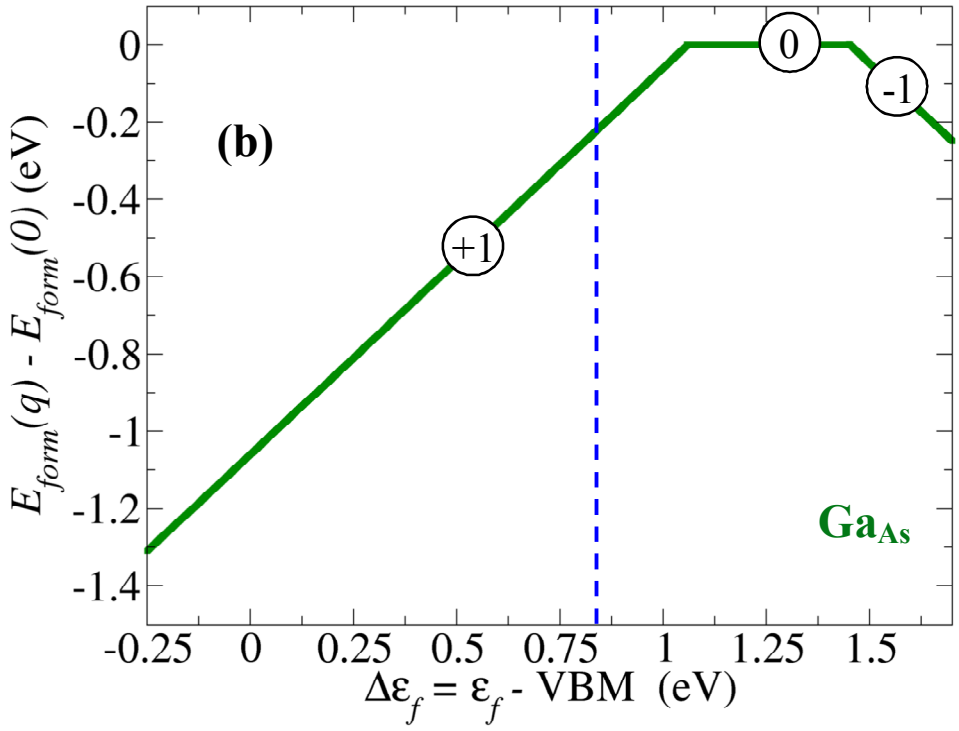
Defect	q	R_q	I
As_2	0	R_0	-0.32
Ga_{As} (a)	0	R_0	-0.52
As_{Ga}	0	R_0	-0.37
	+1	R_{+1}	-0.45
Ga_{As} (b)	0	R_0	-0.03

Table 5.1 - Table (i) shows the electron addition energies (A) obtained using the *GW* approximation as described in section 3.6 and used to calculate the formation energies of neutral and negative charge states. Table (ii) shows the electron removal energies (I) obtained using the *GW* approximation as described in section 3.6 and used to calculate the formation energies of positive charge states. Ga_{As} (a) refers to the antisite bonding to the oxide while Ga_{As} (b) denotes the antisite placed in the next As layer "down" from the oxide. Charge state is given by q, and R_q denotes the atomic configuration of the q^{th} charge state. The corresponding bulk values can be found in table 4.2, page 92.

Defect	Δ		(eV)
As_2	$E(q_{-1}, R_{-1}) - E(q_{-1}, R_0)$	=	-0.287
	$E(q_0, R_0) - E(q_0, R_{+1})$	=	-0.015
	$E(q_{+1}, R_{+1}) - E(q_{+1}, R_0)$	=	-0.018
Ga_{As} (a)	$E(q_{-1}, R_{-1}) - E(q_{-1}, R_0)$	=	-0.037
	$E(q_0, R_0) - E(q_0, R_{+1})$	=	-0.041
	$E(q_{+1}, R_{+1}) - E(q_{+1}, R_0)$	=	-0.025
As_{Ga}	$E(q_0, R_0) - E(q_0, R_{+1})$	=	-0.174
	$E(q_{+1}, R_{+1}) - E(q_{+1}, R_0)$	=	-0.181
	$E(q_{+2}, R_{+2}) - E(q_{+2}, R_{+1})$	=	-0.046
Ga_{As} (b)	$E(q_{-1}, R_{-1}) - E(q_{-1}, R_0)$	=	-0.066
	$E(q_0, R_0) - E(q_0, R_{+1})$	=	-0.058
	$E(q_{+1}, R_{+1}) - E(q_{+1}, R_0)$	=	-0.041

Table 5.2 - Charge state relaxation energies calculated within DFT. The magnitude of Δ reflects the degree of structural relaxation. The values shown in this table are added to the electron addition (removal) energies of table 5.1 to obtain the formation energies of neutral and negative (positive) charge states. Ga_{As} (a) refers to the antisite bonding to Al while Ga_{As} (b) denotes the antisite placed in the next As layer "down" from the oxide. Note in all cases Δ consists of total energy terms involving the same charge state. The corresponding bulk values can be found in table 4.3, page 93.


(a)
0
-0.1
-0.2
-0.3
-0.4
-0.5
-0.25
0
0.25
0.5
0.75
1
1.25
1.5
+1
0
-1
As2
Eform(q) - Eform(0) (eV)
Δεf = εf - VBM (eV)


(b)
0
-0.2
-0.4
-0.6
-0.8
-1
-1.2
-1.4
-0.25
0
0.25
0.5
0.75
1
1.25
1.5
+1
0
-1
GaAs
Eform(q) - Eform(0) (eV)
Δεf = εf - VBM (eV)

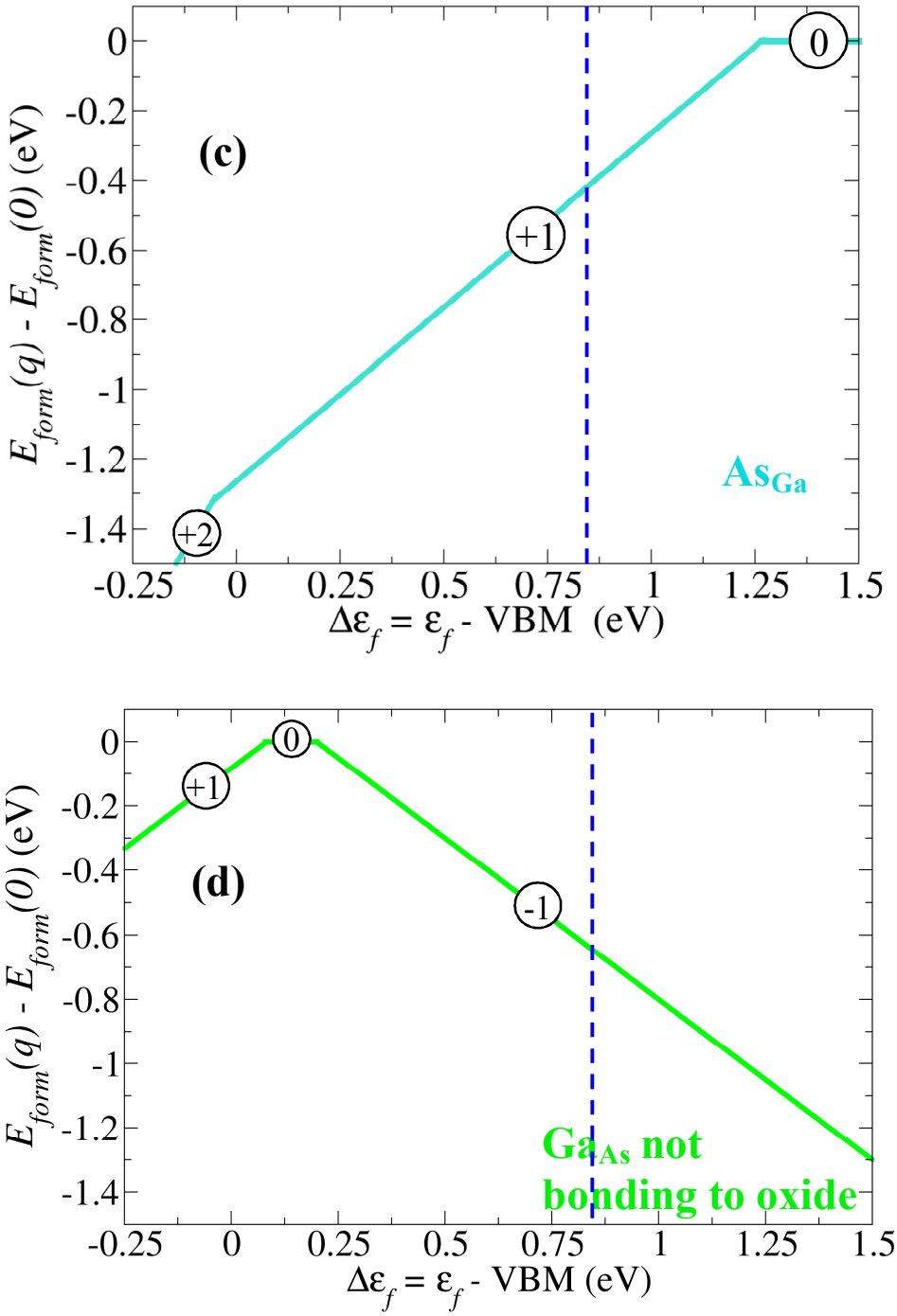


Figure 5.10 - Formation energy versus Fermi level of defects bonding to or near the InGaAs/oxide interface. As dimer (As_2) shown in (a), Ga_{As} in (b), As_{Ga} in (c), and (d) represents the Ga_{As} antisite located in the 2nd As layer down from the oxide *i.e.* not bonding to the oxide. For each defect, q-state formation energies relative to neutral (q = 0) formation energy are plotted against the position of the Fermi level relative to the VBM. Encircled numbers denote charge state. The blue dashed line indicates the position of the bulk CBM calculated within the DFT+*GW* approach (see section 4.2).

5.3.3 Conclusions

This study applies first principle electronic structure methods to identify the physical origin of experimentally inferred midgap states of electrically active defects in III-V materials, specifically of capacitance-voltage derived D_{it} profiles at $In_{0.53}Ga_{0.47}As$/high-k oxide interfaces. To the best of our knowledge, this is the first application of the GW approximation to point defects at (100) $In_{0.53}Ga_{0.47}As$ surfaces passivated with Al_2O_3. Our calculations show that no defects bonding directly to the oxide give rise to charge transition levels which correlate with the observed midgap peaks in the interface state distribution [6, 35, 38, 47]. Acceptor-like (neutral when the defect state is unoccupied and negatively charged when occupied) transitions in the conduction band are observed however (see figure 5.11, page 125), which correlates with recent computational studies [31, 32] and agrees with recent experiments indicating that the measured increase of D_{it} into the conduction band originates from acceptor-like defects [6].

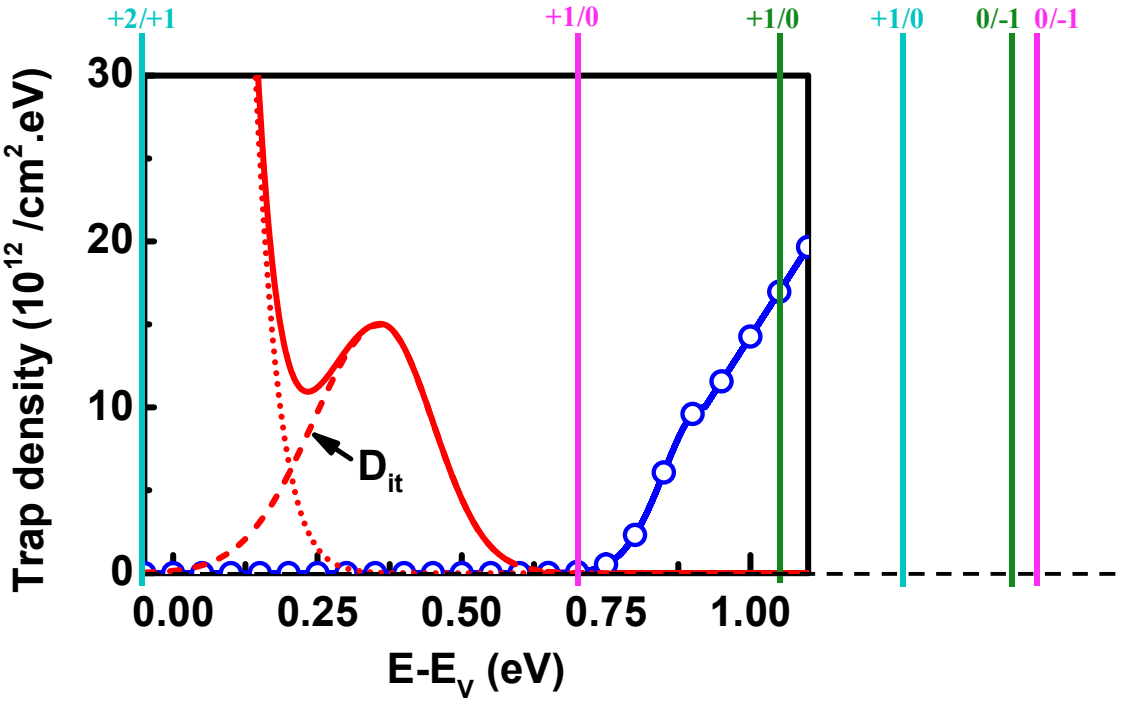


Figure 5.11 - DFT+*GW* calculated charge transition levels for InGaAs defects bonding directly to the oxide (except As_{Ga}) compared with experimentally inferred D_{it}. Experimental data and fits to these data reproduced from [6]. All charge transition levels are given with respect to E_v. Pink, green, and turquoise lines represent As_2, Ga_{As}, and As_{Ga}, resp.

While the defects we have studied in our model of an $In_{0.53}Ga_{0.47}As$ surface passivated with Al_2O_3 do not completely account for the donor-like midgap D_{it} peaks [6] observed in experiment, our calculations do indicate that bonding to the oxide plays a significant role in the position of defect levels. To show this, the Ga_{As} antisite was placed in the next As layer down from the oxide allowing for an investigation of the effect of proximity to Al_2O_3 on the charge transition levels. Our calculations show that the charge transition levels shift by a large degree as a function of distance from the oxide, bringing them in much closer agreement with the levels calculated in the bulk (see figure 5.12, page 126), as could be anticipated from the local character of the bonding.

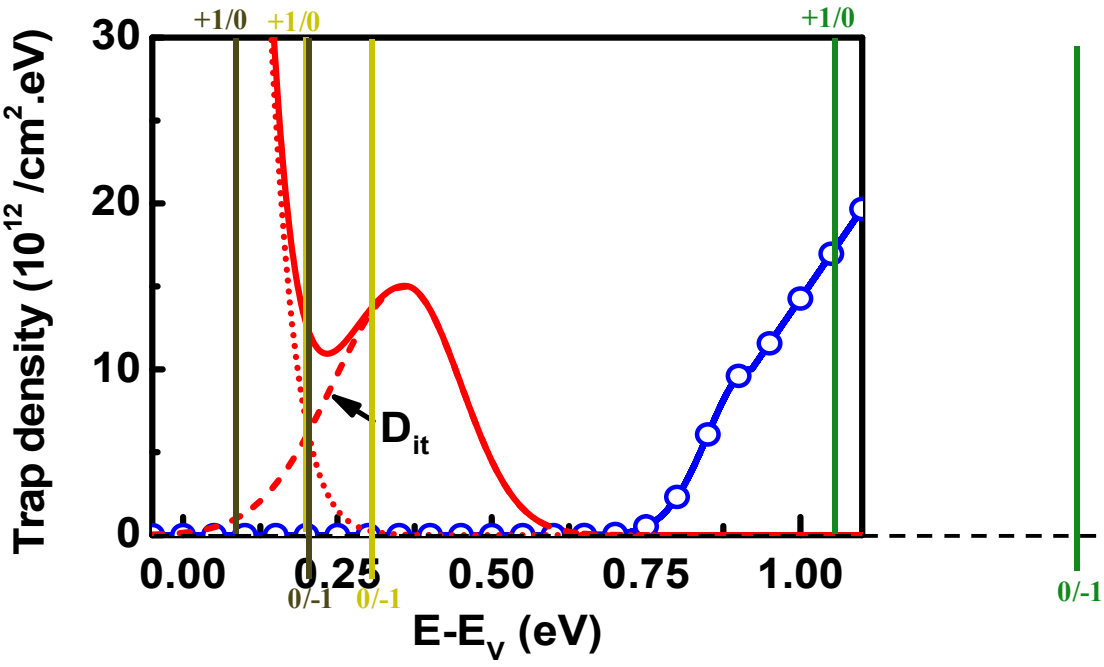


Figure 5.12 – DFT+*GW* calculated charge transition levels for the Ga_{As} defect bonding to 2 In and 2 Ga as a function of proximity to the oxide. Green lines indicate the charge transition levels of the Ga_{As} antisite bonding directly to the oxide. Yellow lines correspond to the bulk values presented in chapter 4. Brown lines represent Ga_{As} located in the next As layer down from the oxide. Experimentally inferred D_{it} shown for comparison and reproduced from [6]. All charge transition levels are given with respect to E_v.

Interestingly, a sharp change in the properties of native defects in III-V semiconductors as a function of depth from the surface towards bulk-like properties has been noticed in previous computational works [133]. In fact, this work shows that defect levels and formation energies of sub-surface anion vacancies approach their bulk values when located below the second atomic layer down from the (110) surface of GaP. In a separate study, other researchers [144] noticed a change of 1.92 eV in the total energy of As_{Ga} antisites in (110) GaAs, along with a 1.16 eV change in the defect energy levels as the defect is moved from the third layer below the surface to the surface layer. In ref. [144], it is argued that this abrupt change in the electronic structure occurs as a result of a change from bulk-like four-fold to three-fold coordination as the antisite is moved up to the surface layer. These works focus on defects at unpassivated surfaces, yet some similar behavior is found in our studies that include Al_2O_3

passivation, as well as differing in the surface orientation. We find that when 2 monolayers separate the defect from the oxide, charge transition levels of the Ga_{As} antisite can change by as much as 1.25 eV relative to the case of bonding to Al_2O_3, while the corresponding charge transition levels in bulk differ by ~0.1 eV relative to the case when the defect is located in the next As layer down from the oxide. Combining our findings with the results of [133, 144], strong indications are provided that point defects in III-V materials return to bulk-like characteristics within only a few atomic layers from the surface, despite having very different properties when located at the surface/interface layer. With this in mind, and considering that our calculated charge transition levels for defects as close as possible to the oxide are all close to or inside the semiconductor host bands, the following conclusion can be drawn: the observed midgap D_{it} originates from point defects *not* bonding directly to the high-*k* oxide, but rather from defects located inside the semiconductor a few atomic layers or more from the insulator. In agreement with [13], the As_{Ga} antisite in a bulk environment is a likely candidate for the midgap D_{it} peak (see figure 5.14, page 128). Considering the variation (~0.1-0.2 eV) in charge transition levels of the Ga_{As} antisite with local cation disorder, in addition to the much wider variation as a function of distance from the oxide, we postulate that such differences in the local environment of point defects contribute to the D_{it} features extending into the semiconductor bands (see figure 5.13, page 128). Local cation disorder of anion-situated bulk defects contribute to the D_{it} feature close to and extending into the valence band (see figure 5.13), while the peak in the conduction band arises from point defects bonding to or within two monolayers from the oxide.

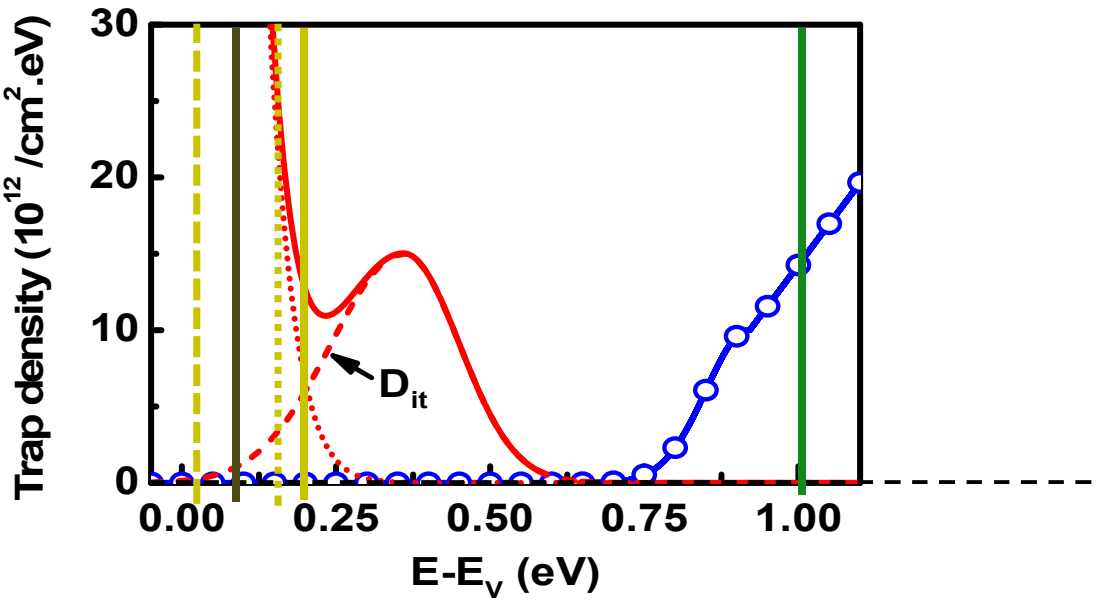


Figure 5.13 - DFT+*GW* calculated charge transition levels for the +1/0 charge transition of the Ga_{As} antisite. Green line indicates the charge transition levels of the Ga_{As} antisite bonding directly to the oxide. Solid, dashed, and dotted yellow lines correspond to the bulk values for the cases of bonding to 2Ga and 2 In, bonding to 4 Ga, and bonding to 4 In, respectively. Brown line represents Ga_{As} located 1 As layer down from the oxide. Graph of D_{it} shown for comparison and reproduced from [6]. Charge transition levels are given with respect to E_v.

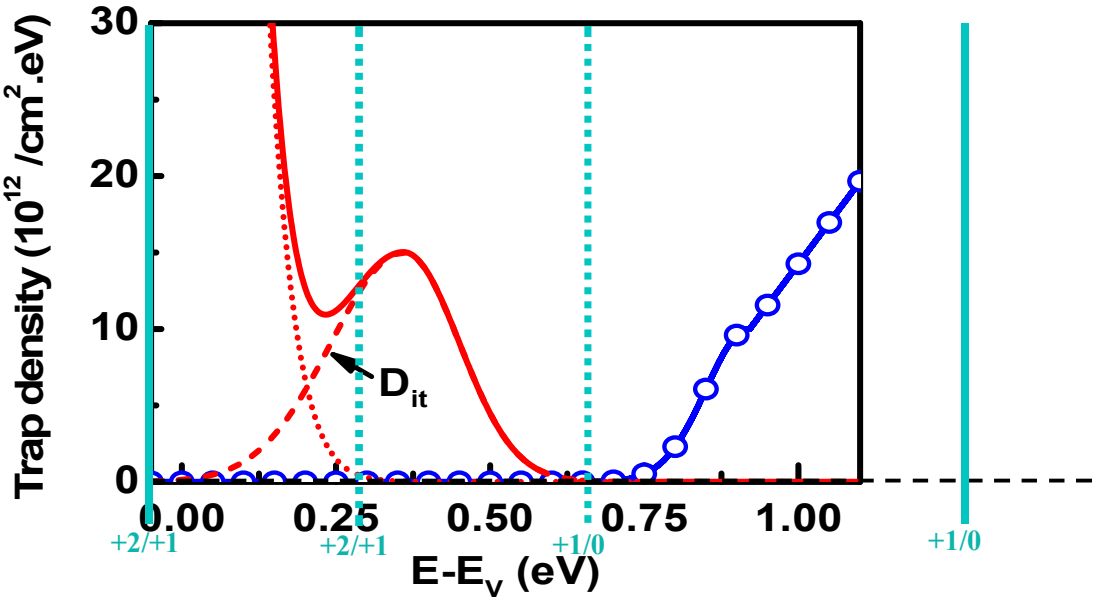


Figure 5.14 - DFT+*GW* calculated charge transition levels for the As_{Ga} antisite. Solid turquoise lines represent the defect close to the oxide. Dotted vertical lines correspond to the bulk values. Experimentally inferred D_{it} shown for comparison and reproduced from [6]. All transition levels are given with respect to E_v.

It is important to note that the red and blue colors of the experimental D_{it} curve (figures 5.11-5.14) correspond to amphoteric labels, where red indicates donor-like (+/0) and blue indicates acceptor-like (0/-) charge transitions. These labels were inferred from a fitting between theoretical and experimental CV curves [6], and correspond to *net* charge transitions of the participating defects; they are the net result of a large number of defect centers all changing their charge states. Hence, there is no inconsistency in the assignment of a single, calculated charge transition of a single, particular defect to a feature of the experimental D_{it} which has a different amphoteric label. For example, the midgap D_{it} peak labelled as +1/0 could easily involve a +2/+1 charge transition introducing the possibility that the observed midgap states could arise from the presence of As_{Ga} antisites.

Chapter 6
Divacancies in Carbon Nanotubes and their Influence on Electron Scattering

6.1 Introduction

First-principles calculations are applied to study the formation energies of various divacancy defects in armchair and zigzag carbon nanotubes of varying diameter, and the electron transport properties for the corresponding structures. We investigate the relative stabilities of the 585 and 555 777 defects as a function of CNT radius from first-principles calculations, extending previous studies to CNTs with much larger radii than previously investigated. Our explicit *ab-initio* calculations confirm that the lateral 585 divacancy is the most stable defect in small diameter tubes, with the 555777 divacancy becoming more stable in armchair tubes larger than (30, 30); images of the defect structures which clarify the defect nomenclature are shown in figures 6.1 and 6.2 (pages 133 and 134). Following the study of formation energies, the optimized structures are then used as inputs to a charge transport simulator [12, 116-118]. This allows us to calculate the transmission properties of charge carriers and obtain the scattering mean free path due to single defects. The scattering resulting from the introduction of the defects is calculated for a range of radii, chiralities and defect configurations. The strongest scattering is found for the 555777 divacancy configuration, which should be observable in electrical spectroscopy experiments. Using knowledge of the scattering contributions from individual defects, we apply an approximation (see section 3.8) that allows us to establish the localization length arising from multiple defects in much longer nanotubes and we are able to relate this to the resistivity of the nanotubes. We find that localization lengths increase with increasing diameter at fixed defect density, and can exceed 100 nm for typical defect densities.

6.2 Methods

6.2.1 Geometry and Electronic Structure

For total energy and electronic structure calculations, we use density functional theory (DFT) as implemented in the OpenMX code [76]. The method expresses Kohn–Sham molecular orbitals as a linear combination of numerical localized pseudo-atomic orbitals (NAOs)

$$\varphi_\mu = \sum_{i,\beta} c_{\mu,i,\beta}\, \vartheta_{i,\beta}(\boldsymbol{r} - \boldsymbol{R}_i) \qquad (6.1)$$

The NAO basis functions ϑ are obtained by solving the Schrödinger equation for an atom in a slightly modified environment that gives a finite cutoff radius. In equation (6.1) $c_{\mu,i,\beta}$ is the expansion coefficient, i is the site index, β is the orbital index - which includes the quantum numbers l and m - and μ is the Kohn–Sham molecular orbital index. Norm-conserving, fully relativistic pseudo-potentials and the corresponding basis sets used in all calculations are generated by the software package ADPACK [76, 95, 96]. In our calculations, the Ceperley-Alder form of the local density approximation (LDA) is used [84, 89]. A double-ζ NAO basis set with a cutoff radius of 1.85 Å is used to describe the carbon atoms. The basis set accuracy is numerically tuned via variational optimization [145] of the carbon basis set (s32p32), in which two optimized orbitals are generated from three primitive orbitals (for a full description of optimized NAO generation, see [95, 96]). A constant vacuum spacing of 13 Å in both directions perpendicular to the nanotube axis is maintained for all CNTs. Sampling in k-space is performed on a 4 × 1 × 1 *k-point* grid; 4 *k-points* are sampled in the direction corresponding to the CNT axis.

Defective CNTs are constructed by either removing a single carbon atom (monovacancy) or two carbon atoms (divacancies) from otherwise pristine structures. The latter are constructed using thirteen unit cells for $n \leq 10$, and

using seven unit cells for $n \geq 20$ in tubes of chirality (n, n) and (n, 0). This indicates a distance between periodic images of defects of over 55 Å and 30 Å for zigzag tubes of $n = 10$ and $n = 20$, respectively, and over 32 Å and 17 Å for armchair tubes of $n = 10$ and $n = 20$, respectively. For all structures, we observe on average a difference of 0.01 Å in the bond lengths situated midway between defects, relative to corresponding bonds in the pristine tubes. The atomic configurations of both pristine and defective tubes are determined by calculating the forces between atoms and updating atomic positions using a steepest decent optimization method until the maximum force converges to less than 0.005 eV/Å.

6.3 Results

6.3.1 Vacancy Formation

In this section, we present an analysis of the formation energies of the 585 and 555777 configurations [68]. The 585 divacancy consists of two pentagons on either side of an octagon, whereas the 555777 structure is formed by three pentagons and three heptagons. Divacancies are also classified as lateral or vertical, depending on the orientation relative to the tube axis of the pentagons formed as a result of the missing atoms. In figure 6.1, we indicate the various atomic arrangements for the largest studied CNTs for the relaxed atomic configurations. In figure 6.2 an example of the graphene supercell is shown. The results for graphene are included to compare with the nanotubes as the limiting case of zero curvature or infinite radius.

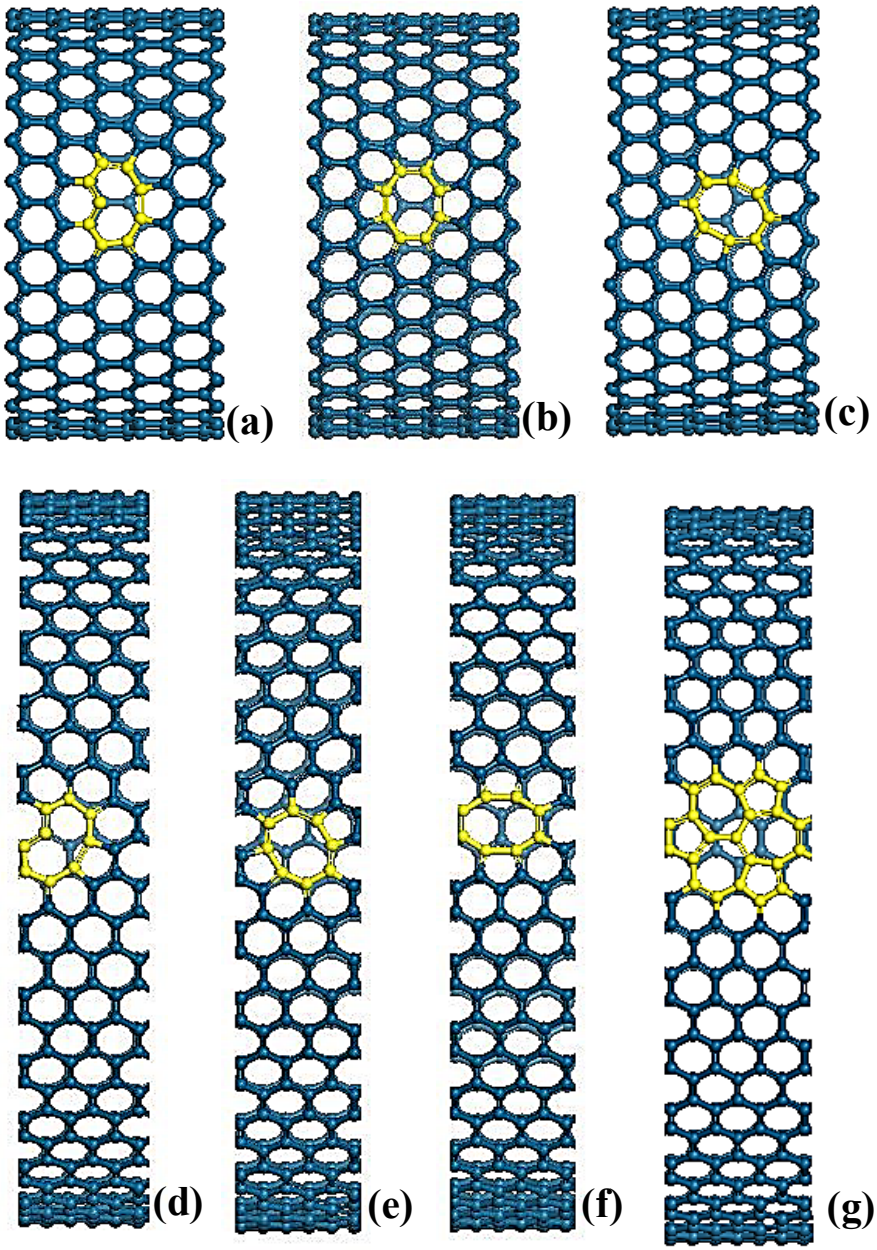


Figure 6.1 - Top three images show (30, 0) zigzag CNTs with: (a) monovacancy, (b) lateral and (c) vertical divacancies in the 585 configuration. The bottom four images show (30, 30) armchair CNTs with: (d) monovacancy, (e) lateral and (f) vertical divacancies in the 585 configuration, and (g) a divacancy in the 555777 configuration. All images show the central three unit cells extracted from the 13 (7) unit supercells for chiral vectors $n \leq 10$ (≥ 20).

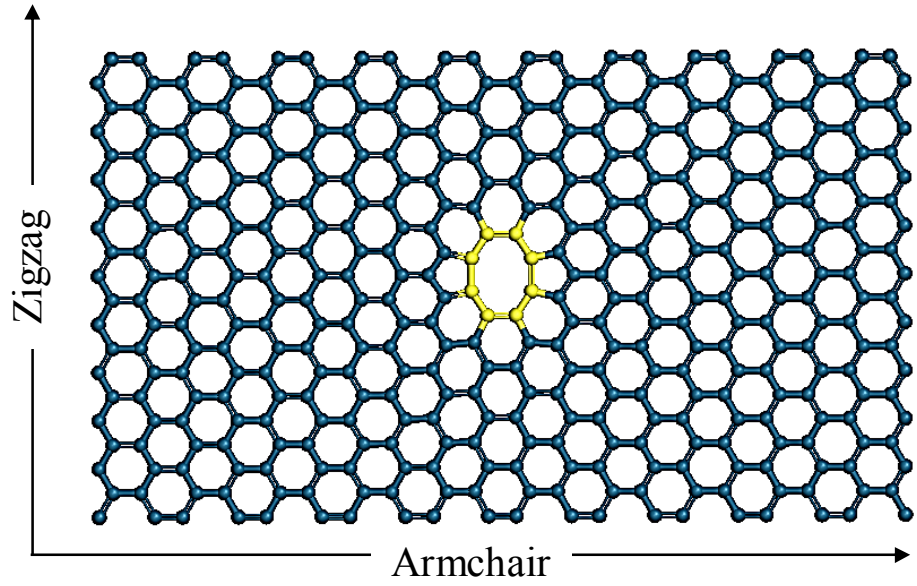


Figure 6.2 - An example of a graphene cell with a divacancy defect (lateral relative to the zigzag direction, vertical relative to the armchair direction).

The total energies of the defective tubes and the gas-phase carbon dimer were obtained from calculations performed on all structures after relaxing the geometries as prescribed in section 6.2.1. Spin polarization was included for the isolated carbon atom treated as a spin-unrestricted triplet state, while the carbon dimer and all CNTs were treated without spin polarization. This leads to a dissociation energy for C_2 of 6.8 eV, which slightly overestimates the experimental bond strength of the carbon dimer of 6.4 eV as is typical of DFT-LDA.

The defect formation energies (see section 3.7) and relative stabilities for armchair and zigzag CNTs are presented in tables 6.1 and 6.2, respectively. The formation of lateral divacancies is more stable than the formation of vertical divacancies: $\Delta E_{latdivac} < \Delta E_{vertdivac}$ for all CNTs. Our results also indicate that each divacancy is always more stable than 2 isolated monovacancies in a given CNT. This demonstrates qualitative agreement with calculations on tubes with diameters up to (10, 10) [12], which were performed using a self-consistent implementation of the Harris functional [146]. Changing chirality by looking at zigzag CNTs reveals the same qualitative picture for the defect formation: lateral divacancies are energetically preferred over vertical divacancies. With increasing diameter the difference in formation energies between the lateral and vertical

divacancies decreases for both armchair and zigzag CNTs, as may be anticipated due to a reduction in bond strain with decreasing curvature.

	(5, 5)	(7, 7)	(10, 10)	(20, 20)	(30, 30)	Graphene
ΔE_{mono}	-7.86	-7.81	-6.05	-4.98	-1.41	0.0
$\Delta E_{latdivac}$	-11.11	-8.63	-7.78	-7.23	-0.79	1.20
$\Delta E_{vertdivac}$	-7.14	-5.47	-5.41	-5.16	1.09	1.19
ΔE_{555777}	-4.53	-4.44	-4.13	-3.61	-2.46	0.86
$d\Delta$	3.97	3.16	2.38	2.06	1.88	0.01[a]

Table 6.1 - Defect formation energetics of the armchair CNTs, calculated using equation (3.63) (page 70) and then reported as the difference from the formation of a monovacancy in graphene. The last row gives $d\Delta$ *i.e.* the magnitude of difference in energy between lateral and vertical divacancies. All formation energies are in eV.

	(10, 0)	(20, 0)	(30, 0)	Graphene
ΔE_{mono}	-7.58	-6.82	-2.82	0.0
$\Delta E_{latdivac}$	-11.05	-9.64	-1.84	1.20
$\Delta E_{vertdivac}$	-6.36	-5.97	-0.27	1.19
ΔE_{555777}				0.86
$d\Delta$	4.69	3.67	1.56	0.01[a]

Table 6.2 - Defect formation energetics of the zigzag CNTs calculated using equation (3.63) and then reported as the difference from the formation energy of a monovacancy in graphene. For graphene, the defect name is given by its orientation relative to the tube axis of a CNT rolled up in the 'zigzag-axis' *e.g.* $\Delta E_{latdivac}$ = formation energy of the defect shown in figure 6.2. All formation energies are in eV.

[a] Denotes the difference in energy of the lateral and vertical divacancies in graphene due to finite simulation cell effects.

We now turn to the discussion of the formation energy of the 555777 divacancy. Amorim *et al* have previously estimated that 555777 divacancies are less stable than their 585 counterparts for (n, n) tubes with $n < 30$ and for (n, 0) tubes with $n < 117$ [68]. Their estimates of defect formation energies are not based on *ab-*

initio calculations, but on a parameterized molecular mechanics model. To test the conclusions reached by Amorim *et al*, we calculate formation energies of 555777 divacancies in armchair tubes. Studying the stability of this defect in (117, 0) zigzag tubes would not be possible given the computational resources at hand; hence, we consider this defect in armchair tubes only. As can be seen in table 6.1, the 555777 defect configuration becomes more stable than the 585 configurations for n = 30, in agreement with the molecular mechanics estimate in ref. [68].

6.3.2 Transport Properties

As stated in chapter 3, the transport simulator TiMeS [116] is used to calculate electronic transport through defective CNTs. A transport module [76] is also included in the OpenMX package. We find, however, that it is more memory intensive, thereby not allowing us to study large structures even within a single-ζ Hamiltonian representation. Nevertheless, we validated our transport calculations with TiMeS by comparing results for tubes of chirality (n, n) and (n, 0) for n = 10, 20 with results obtained using the transport module of OpenMX. The comparison is shown in figure 6.3. All results are in excellent agreement.

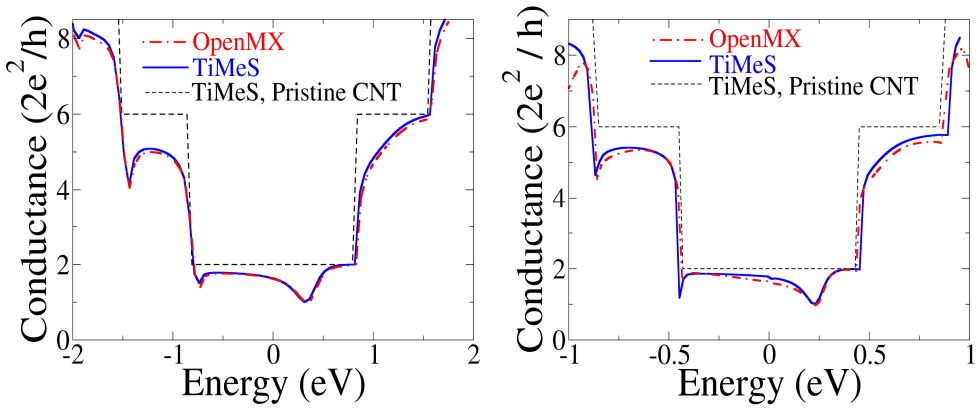


Figure 6.3 - This graph demonstrates the excellent agreement of our TiMeS transport calculations with results obtained from OpenMX for (10, 10) (left panel) and (20, 20) (right panel) CNTs with a lateral divacancy. The Fermi level is set to 0 eV.

Before discussing the conductance spectra for the largest defective CNTs, we show that the use of a double-ζ basis does not affect transport significantly, as may be expected for the simple sp^2 network of carbon nanotubes. We justify the choice of minimal basis for the calculation of transport properties by showing that for CNTs the variation of the conductance is in fact reasonably small when going from single-ζ to double-ζ basis sets. This is shown in figure 6.4 and allows us to use the less computationally demanding single-ζ basis. This result agrees with work from Ohfuchi *et al* who tested the validity of using large basis sets for geometry optimization and then switching to single-ζ basis sets for NEGF-based transport calculations of CNTs bridged between graphene electrodes [94].

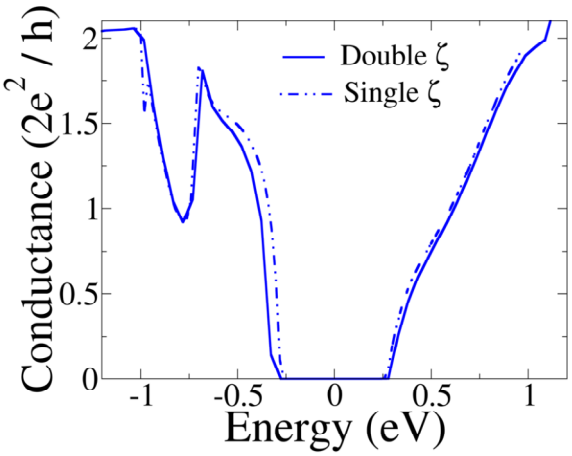


Figure 6.4 - Illustration of reasonably small discrepancy in transmission spectra between the two basis sets, using the (10, 0) CNT with a lateral divacancy. Notice the particularly good agreement in the conduction band. Both curves were obtained from TiMeS. The Fermi level is set to 0 eV.

Figure 6.5 shows the conductance spectra for the (20, 0), (20, 20), (30, 0), and (30, 30) tubes each with various types of defects. First, we note that a small bandgap (40 meV) appears in the transmission spectrum of the (30, 0) tube. A simple zone-folding analysis of the electronic structure (*i.e.* mapping of the graphene band structure to that of a tubular lattice by folding the appropriate Brillouin zone boundaries) would predict metallic behaviour for the (30, 0) tubes, as $(n - m)/3$ is an integer.

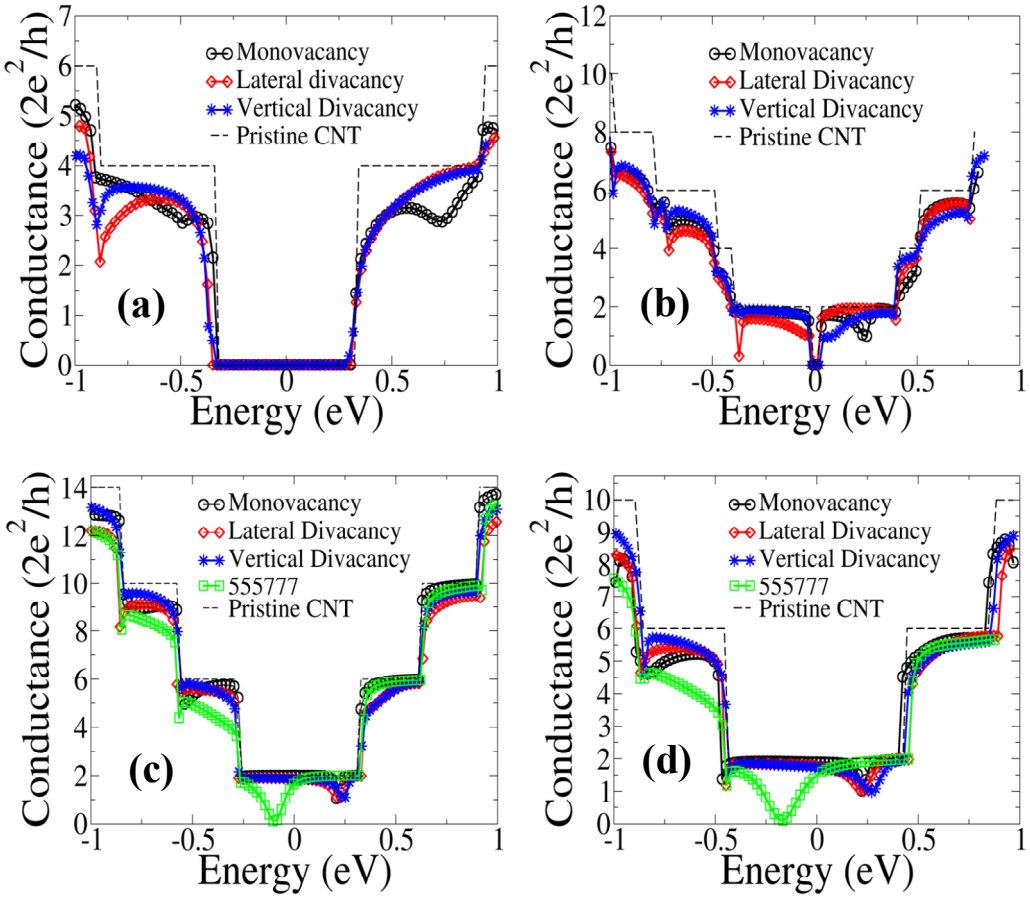


Figure 6.5 - Conductance spectra for zigzag tubes (graphs (a) and (b) for (20, 0) and (30, 0) respectively) and armchair tubes (graphs (c) and (d) for (20, 20) and (30, 30) respectively). Note the greater amount of scattering for the 555777 configuration of the divacancy defect. The Fermi energy is 0 eV.

The observed bandgap is likely due to curvature induced π-σ hybridization, which is not taken into account in the zone-folding scheme [147-151]. To investigate this effect, independent plane wave DFT calculations were performed using the Quantum Espresso package [99]. These calculations can be systematically converged in the plane wave basis. A kinetic energy cutoff of 100 Rydbergs and a 4x1x1 grid of *k*-points were used. This eliminates the question as to whether the atomic orbitals used in this work give rise to a spurious bandgap due to basis set incompleteness. These parameters were used to calculate the Γ-point gap of a pristine (30, 0) unit cell (equivalent to a lead principle layer (see section 3.5.2 page 58) used to calculate the transmission spectra of the (30, 0)

vacancies, see figure 6.5). A Γ-point gap of 46.4 meV was observed, compared to our OpenMX-calculated bandgap of 40.3 meV, in close agreement. Hence we conclude that the predicted bandgap is not a spurious effect arising from an incomplete basis set. The curvature induced bandgaps in (n/3, 0) tubes have been predicted to decrease with diameter and essentially vanish for tubes larger than (24, 0) [147-151]. The predictions are confirmed by experiment only up to (15, 0) tubes. Our calculations, on the other hand, suggest the finite bandgap persists for larger radii than previously anticipated.

Interestingly, the spectrum of the (30, 30) metallic CNT shows that significant resonant backscattering is generated by the divacancy in the 555777 configuration. In particular, the transmission drops to almost $0.1G_0$ ($G_0 = 2e^2/h$) at ~125 meV below the Fermi level. We argue that this is a quantum interference effect, implying the formation of localized electronic states at the 555777 complex which should be observable with transport spectroscopy. This is further evidenced by an analysis of the effect of the 555777 defect on the local density of states (DOS). The various contributions to the local DOS from each atom making up the 555777 divacancy is shown in figure 6.6. The atoms corresponding to the 555777 defect in the (30, 30) CNT are labelled in the insets of figure 6.6. The local DOS of selected atoms ('S' and 'U') and the average DOS over the defect region are shown in figure 6.6 (b). The local DOS of a carbon atom far from the defect is also plotted for comparison and is labelled as 'pristine'. These graphs show that the atoms associated with the 555777 defect make a significant contribution to the DOS near the Fermi level. Combining these data with the resonant scattering observed in the transmission spectrum in the same energy range, strong evidence is provided for the formation of localized electronic states at this defect.

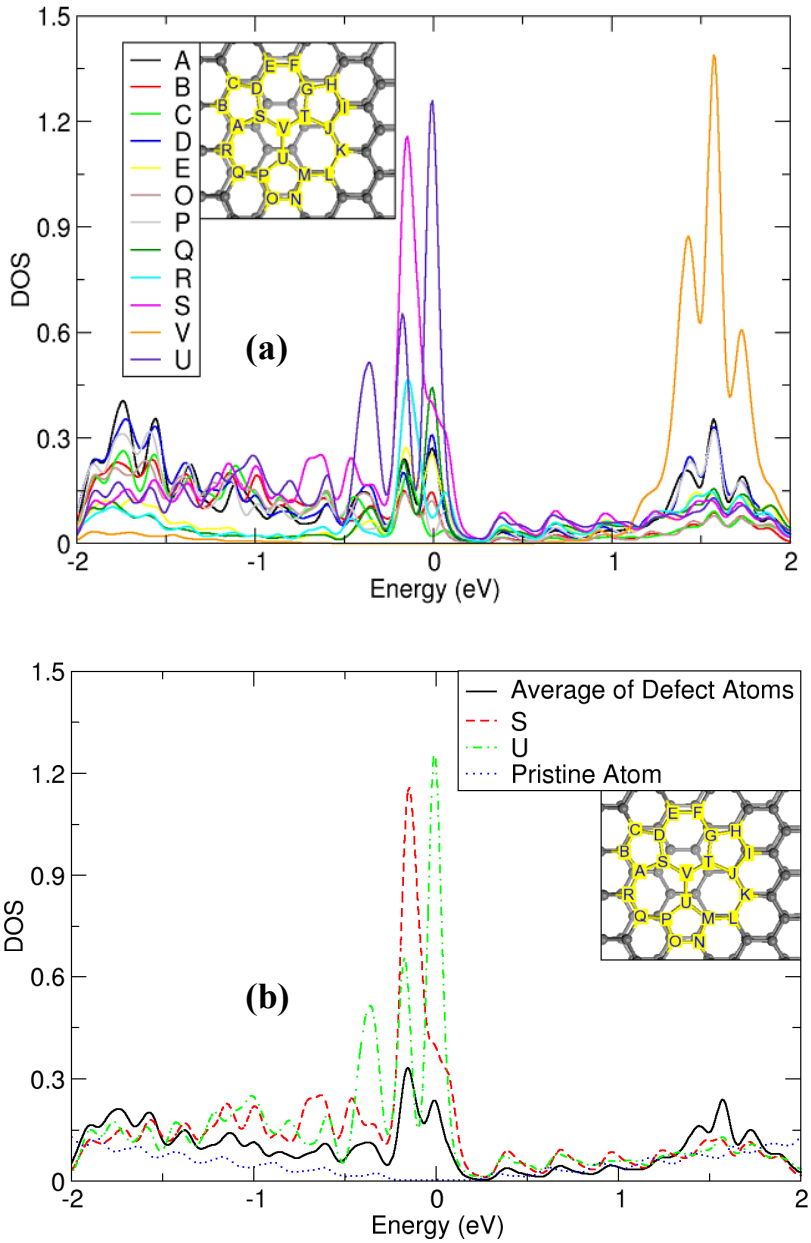


Figure 6.6 - (a) LDOS due to each atom of the 555777 defect. Notice the significant contribution of the "S" and "U" atoms. (b) Comparison of the LDOS of defect atoms (as labelled in the inset) with atoms from a pristine section of a (30, 30) CNT. The Fermi level is set to zero. LDOS units are $(eV)^{-1}$.

Our observations confirm previous suggestions for small diameter tubes [68]. First, the variations in transmission spectra allow one to distinguish between the two types of divacancy defects using transport measurements as a function of diameter. Second, such large amounts of scattering would likely hinder

conductance (or at least would need to be taken into consideration) in applications involving large diameter CNTs, in which 555777 defect geometries become more stable than the 585 configuration. For example, wide CNTs are often considered as candidate interconnects due to the vanishing bandgaps with increasing diameter [152]. Hence our findings for the stability of the 555777 in large diameter CNTs coupled with a large scattering cross section indicates their importance for understanding electron transmission in these applications. Below, we elaborate further on the characteristic transport lengths due to scattering of charge carriers at vacancy defects in large diameter CNTs.

The mean free path λ is estimated using the independent scattering approximation (ISA) described in chapter 3. The ISA relates λ from a concentration of defects to the calculated scattering properties of independent impurities. As explained in section 3.8, this approximation leads to a linear dependence of λ with respect to the mean distance between defects l_d, namely, $\lambda = (R_c/R_s)l_d$ where R_c and R_s are the CNT contact resistance and the resistance of the isolated defect, respectively (see section 3.8.2). The defect concentration is expressed here through l_d. In addition, the localization length in quasi-one-dimensional systems can in the absence of magnetic fields be written as $\xi = (N_{ch} + 1)\lambda = (N_{ch} + 1)(R_c/R_s)l_d$ where N_{ch} is the number of conducting channels [24, 111].

Interestingly, a linear relationship between λ (or ξ) and l_d was also observed earlier using a more elaborate approach that accounts for interference effects from scattering between defects with a rigorous scattering formalism [7, 10, 11]. Our starting point is to compare the previously obtained slopes of the localization length versus mean defect distance to our calculated $(N_{ch} + 1)(R_c/R_s)$ estimate. This is shown in table 6.3.

Chirality	This work. $(N_{ch}+1)(R_c/R_s) = \xi/l_d$ for lateral divacancy.	Previous Reports. ξ/l_d for lateral divacancy
(5, 5)	3.40	5.1 [11]
		5.2 [11]
(7, 7)	6.78	6.4 [11]
(10, 10)	9.61	4.1 [7]
		8.4 [11]

Table 6.3 - Comparison between previous work and our results on the relation $\xi = (constant) \times l_d$, where ξ, l_d and N_{ch} are localization length, distance between defects, and the number of channels at the Fermi level, respectively. R_c and R_s are the contact and scattering resistances defined in section 3.8.2.

The agreement justifies the ISA since interference effects between defects are expected to diminish in larger CNTs with similar defect concentrations. This allows us to extend the analysis to armchair and zigzag CNTs of much larger diameter and for the various types of vacancy defects. Our results are summarized in table 6.4.

Chirality	Defect	$(N_{ch}+1)(R_c/R_s)$
(10, 0)	Monovacancy	10.44
	Vertical divacancy	6.79
	Lateral divacancy	5.54
(20, 0)	Monovacancy	19.88
	Vertical divacancy	19.10
	Lateral divacancy	18.69
(30, 0)	Monovacancy	75.97
	Vertical divacancy	24.96
	Lateral divacancy	61.49
(5, 5)	Monovacancy	9.03
	Vertical divacancy	3.11
	Lateral divacancy	3.40
	555777	1.88
(7, 7)	Monovacancy	24.44
	Vertical divacancy	4.27
	Lateral divacancy	6.78
	555777	3.41
(10, 10)	Monovacancy	55.42
	Vertical divacancy	14.61
	Lateral divacancy	9.61
	555777	8.10
(20, 20)	Monovacancy	388.70
	Vertical divacancy	34.09
	Lateral divacancy	42.10
	555777	5.07
(30, 30)	Monovacancy	9995
	Vertical divacancy	70.30
	Lateral divacancy	59.68
	555777	22.90

Table 6.4 - The 3rd column shows, for each defect, the proportionality constant between localization length and the distance between defects l_d in the relation $\xi = (constant) \times l_d$.

A graphical representation of selected data follows in figure 6.7. There are a few observations to make.

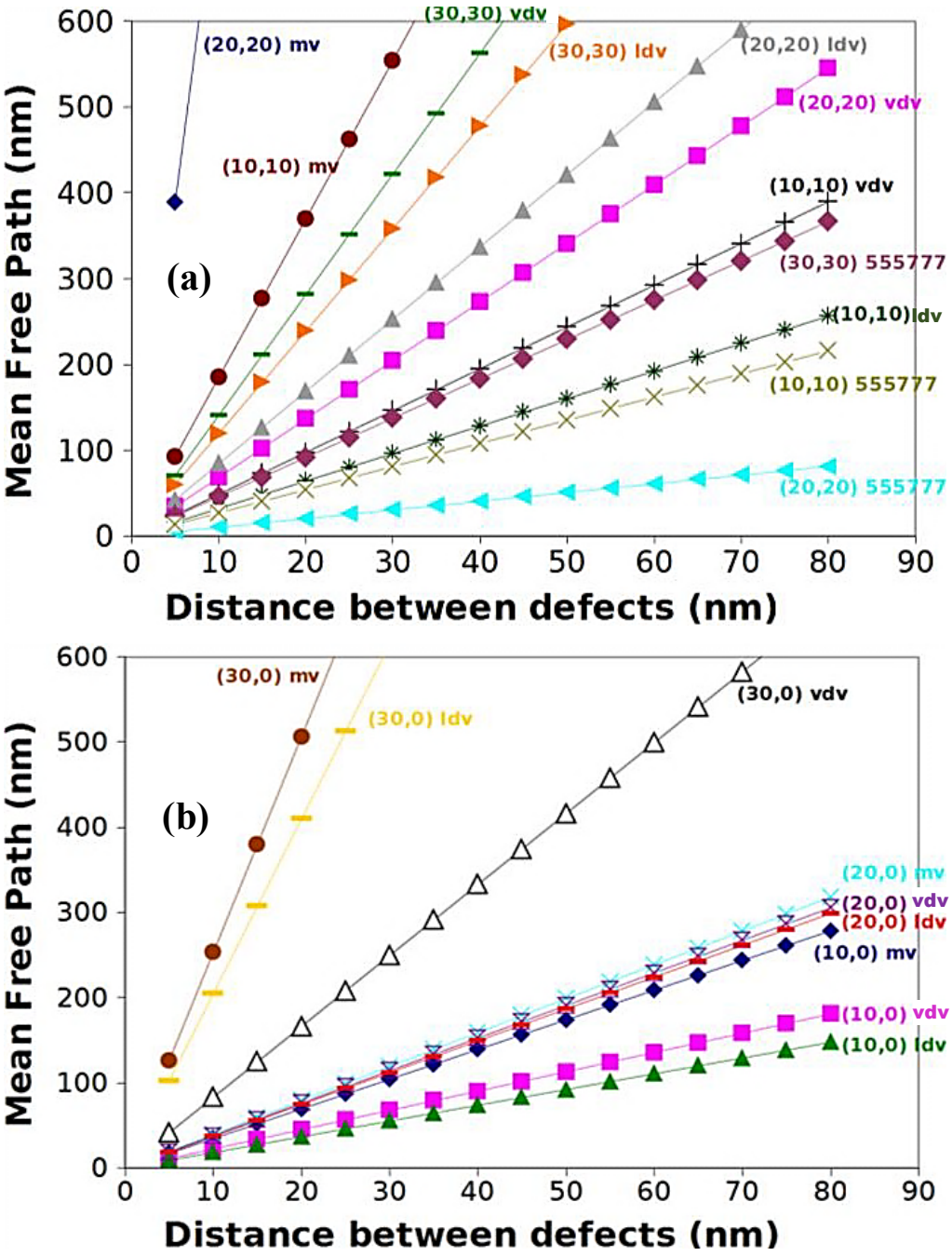


Figure 6.7 - Mean free path λ versus distance between defects l_d for (a) armchair and (b) zigzag carbon nanotubes.

First, as the tube size increases the mean free path extends significantly for a

fixed defect concentration. This is attributed to the diminishing R_s, which is proportional to the ratio of the scattering cross section of the defect to the tube diameter. This can be seen in figure 6.8.

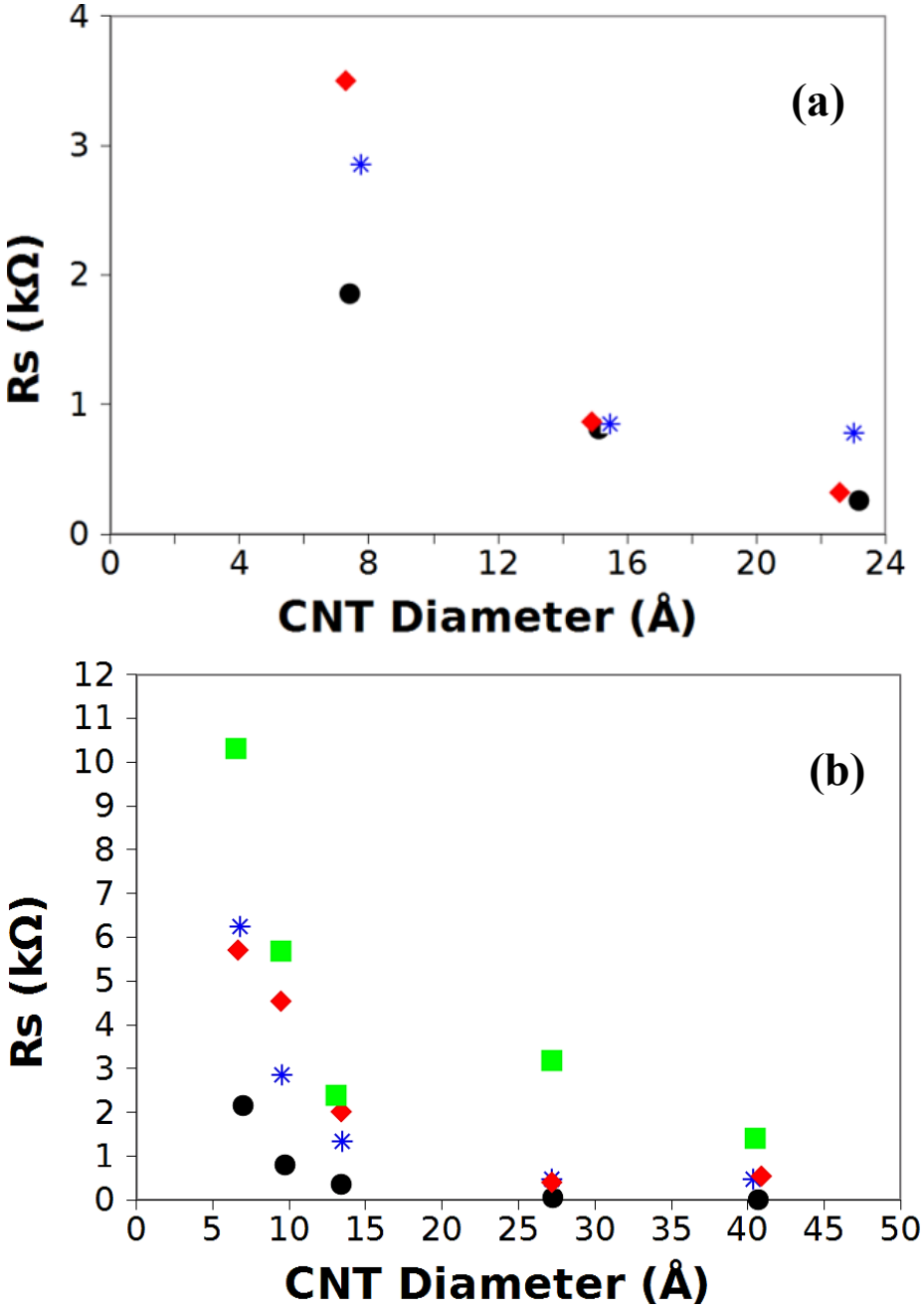


Figure 6.8 - The scattering resistance of the various isolated defects for (a) zigzag and (b) armchair tubes decreases with increasing CNT diameter. Black circles represent monovacancies, blue stars represent vertical divacancies, red diamonds represent lateral divacancies, and green squares represent the 555777 defect.

Second, monovacancies and divacancies introduce different amounts of scattering depending on the chirality. Monovacancies have a much lower effect than divacancies on the conductance of metallic armchair tubes. This confirms previous results showing that monovacancies lead to very little scattering at the Fermi level [12]. In contrast, both monovacancies and divacancies have comparable R_s when $n > 10$ in $(n, 0)$ zigzag CNTs. Also, the latter show very similar scattering between lateral and vertical divacancies, unlike the armchair tubes. Nevertheless, in all cases the differences between the various divacancy configurations decline as the diameter increases. This should be expected in analogy to the convergence of the formation energies for larger tube sizes.

6.3 Conclusions

Our main findings for defective CNTs are as follows: (i) extending previous results [68], we confirm from explicit *ab-initio* calculations that the 555777 divacancy becomes more stable in armchair tubes larger than (30, 30); (ii) evaluating electron transmission as a function of diameter and chirality for a range of defects and defect configurations, we find strong scattering from the 555777 divacancy configuration, which should be observable in electrical spectroscopy experiments; (iii) establishing an approximation that relates scattering contributions from independent defects to treat otherwise computationally intractable structures, we extract the localization length in large diameter CNTs.

The validation of an independent scattering approximation was achieved by comparing our calculated transport properties for small diameter tubes with single defects against the work of Rubio, Flores, and Biel *et al* in which experimental data was combined with a rigorous scattering formalism [7, 10, 11]. This validation allows for a study of Anderson localization characteristics in larger CNTs. Specifically, the localization length increases with tube diameter and, for typical defect densities, may be in excess of 100 nm. Nevertheless,

strong scattering from the 555777 configuration is observed. As mentioned, this can be used in electrical spectroscopy experiments to differentiate between defect geometries. Our predictions regarding the larger diameter structures have important implications for technological applications, particularly interconnects, in which wider CNTs are often considered.

Chapter 7
Summary and Conclusions

In chapter 4, a study of the bulk $In_{0.53}Ga_{0.47}As$ electronic structure was carried out using the DFT+GW approach which corrects for the intrinsic limitations of LDA and the KS-DFT formalism. Good agreement with the low temperature experimental bandgap was found. Using this bulk simulation cell, point defects were inserted and their atomic structure and charge transition levels were calculated. The calculation of charge transition levels involves electron addition and removal energies requiring systems with different numbers of electrons. Thus, accurate assessment of charge transition levels necessitates the avoidance of the limitations of DFT, and to this end the GW approximation to the electron self-energy was also employed for these defective simulation cells. Comparing with recent hybrid DFT studies of charge transition levels of point defects in bulk InGaAs [13], overall qualitative agreement is obtained with the largest discrepancy of 0.16 eV exhibited by deep transition levels, while good agreement is seen for shallow levels. Despite the discrepancy at midgap, our bulk calculations are consistent with the conclusion of ref. [13] in which the As antisite was assigned as the most likely physical origin of observed midgap D_{it} peak. However, we explore this conclusion further using a consistent comparison with an oxide-passivated surface model. Before focusing on the surface model, we note that our bulk simulations show that for anion-situated defects, a significant variation of charge transition levels as a function of local cation disorder is observed. Such variations, which are significant yet remain in the lower half of the bandgap, may at least partially account for the D_{it} feature which rises from just below midgap and extends into the valence band, a persistent feature reported in many experimental studies of $In_{0.53}Ga_{0.47}As$/high-k oxide interfaces [6, 37, 48].

It is easy to appreciate that the effect of bonding to an oxide can significantly alter both the electronic and atomic structure of point defects

compared to a bulk environment. Thus, following the study of point defects in bulk $In_{0.53}Ga_{0.47}As$, a consistent comparison with a high-k oxide passivated surface is necessary to investigate the premise that defects in an entirely bulk environment are responsible for the midgap peaks observed in the experimental D_{it}. Such an investigation not only allows for a comparison between "bulk" and "surface" defects (specifically the antisites, in this work), but also allows for a study of defects which are not amenable to a bulk simulation cell, namely the As dimer. To this end, a model of the (100) $In_{0.53}Ga_{0.47}As$ surface passivated with Al_2O_3 was constructed. Our calculations show that all defects bonding to the oxide exhibit charge transition levels which resonate with the semiconductor host bands, most of which are deep into the conduction band.

The large variation between bulk and surface defect properties prompts a quantitative investigation of the defect levels as a function of distance from the oxide. Since the Ga_{As} antisite is amenable to both bulk and surface simulations, and also bonds directly to the oxide in our surface model, this defect was chosen for this investigation. It is found that when the Ga_{As} antisite is located in the second As layer down from the oxide, the defect levels shift dramatically to within ~0.1 eV of their bulk values. Combining our studies of defective bulk and surface models, we conclude that the experimental distribution of interface states can be explained by a combination of the following effects:

(i) Defects bonding directly to the oxide contribute to the conduction band D_{it} feature. As_{Ga} antisites located 1 monolayer down (or in the first In/Ga layer) from the oxide contribute states to the D_{it} feature that extends into the valence band.

(ii) Anion-situated defects bonding to a locally variable cation alloy contribute to the D_{it} feature that rises sharply towards the valence band; the corresponding charge transition levels can change by ~0.1-0.2 eV as a function of the local cation configuration, yet remain in the lower half of the $In_{0.53}Ga_{0.47}As$ bandgap near the valence band edge.

(iii) The position of the +2/+1 charge transition level of the As_{Ga} antisite in an entirely bulk environment is consistent, within the experimental and computational error, with the position of the midgap D_{it} peak, in agreement with previous theoretical investigations [13].

Point defects were also studied in semiconducting and metallic CNTs. The remarkable electronic properties of CNTs make them one of the most important emerging research materials for both front-end and back-end devices [57]. However, the challenges associated with their large scale manufacturing and integration precludes their widespread commercial use in CMOS technology in the near future. Despite this fact, they are good candidates for studying the physics of low-dimensional systems, and the information gleaned from such studies will undoubtedly prove beneficial if the VLSI of CNTs becomes commercially viable. Hence, we study point defects in CNTs using explicit ab-initio methods; localized orbital based DFT for the electronic structure, in conjunction with Green's function based methods to evaluate the electronic transport.

Our calculations confirm the greater stability of the 555777 divacancy over the 585 configuration in armchair CNTs for diameters larger than 40 Å, which extends previous results based on parameterized molecular mechanics [68]. Following the analysis of defect energetics, transport through defective CNTs was assessed as a function of diameter and chirality; we find evidence for strong resonant backscattering from the 555777 defect. This should be observable in electrical transport spectroscopy, allowing for the experimental differentiation of defect geometry. Keeping in mind that large diameter CNTs are often considered for interconnect applications due to the inverse relationship between CNT bandgap and diameter, in conjunction with our evidence that 555777 defects become more stable with diameter, we note the deleterious effect this defect structure may have on potential interconnect applications of CNTs, especially as the backscattering occurs within ~0.1 eV of the Fermi level. Utilizing the independent scattering approximation (which was validated by

comparing to the more rigorous scattering formalism of [7, 10, 11]), we considered defects at typical experimental densities, allowing for an estimate of the mean free path and localization length within the Anderson localization regime of transport for large (up to ~40 Å diameter) CNTs. We find that, for a given defect structure, the localization length rises monotonically with diameter in both armchair and zigzag CNTs and may be greater than 100 nm for typical defect densities.

There are a few points to bear in mind regarding the limitations of the approaches undertaken in this work. For the $In_{0.53}Ga_{0.47}As$ defects (chapters 4 and 5), a study of the formation energies of the defect centers in which the chemical potentials of the added/removed species are taken into account (using equation 2.3 on page 18, for example) was not carried out. Such (temperature- and pressure-dependent) chemical potentials characterize the energy required to transfer atoms to/from their respective reservoirs, and so can provide an estimate of the formation energies of defect centers as a function of experimental growth conditions [13]. For example, under As-rich growth conditions, one uses a large chemical potential for As atoms [5], which in turn would lower the As_{Ga} antisite formation energy (see equation 2.3), and the concentration (c) of this defect would then be expected to increase as [5]

$$c \propto e^{-(E_{form}/k_B T)} \tag{7.1}$$

assuming thermodynamic equilibrium (k_B is Boltzmann's constant and T is temperature). If included in this work, such a study would provide a comparison of the concentrations of each defect in both bulk and high-*k* oxide passivated surface chemical environments. This would yield information on which defect(s) is (are) most likely to form for a given set of growth conditions, and how these dependencies on growth conditions change between bulk and surface models. Komsa and Pasquarello carried out such a study for bulk defects [13], and under As-rich growth conditions the As_{Ga} antisite (which exhibits a midgap charge transition level) would likely be present in large concentrations. A corresponding

study of the formation energetics of the defects studied in this thesis near an interface with a high-k oxide, and whether or not these energetics change significantly for different oxides, would yield a stronger comparison of our approach against hybrid DFT [13] (as opposed to relying on a comparison between charge transition levels only). Such simulations of the relative stabilities of defects as a function of surface preparation techniques (cation- or anion-rich growth conditions, and nature of the oxide) would also provide a more rigorous characterization of the candidates giving rise to experimentally observed D_{it} [6, 37, 48].

In relation to the limitations concerning the modeling of defects in CNTs (chapter 6), our study would benefit from larger simulation cells. In this work, the validity of the independent scattering approximation (ISA) is assumed based on a previous application involving impurities in Si nanowires [119], and a comparison against a rigorous approach in which multiple defects are described in a single simulation cell [7, 10, 11]. However, our study would benefit from a complementary analysis of multiple defects performed using our choice of methods (described in chapter 3). While being very computationally expensive, longer simulation cells for the scattering region (see section 3.5.2, page 58, and section 3.8.1, page 72) would allow for the incorporation of more than one defect, which in turn would provide an explicit study of the scattering properties as a function of distance between defects *i.e.* defect density (as opposed to the implicit study of these properties in which the defect density is accounted for within an assumed linear relationship between the mean free path of a single defect and a variable distance between defects). This would give a stronger test of the applicability of the ISA to the systems we study in this thesis, and in particular where the ISA breaks down; how close point defects within a CNT have to be before interference effects between scattered electron waves have a significant impact, and in addition how this relationship changes with decreasing curvature as the CNT diameter is increased.

In this thesis, first principles electronic structure methods were applied to the study of point defects in technologically relevant post-Si device materials.

The materials considered here, $In_{0.53}Ga_{0.47}As$ and CNTs, represent both short and long-term solutions to the problem of continued Moore's law scaling beyond what is achievable using conventional Si-based MOSFETs integrated with Cu interconnects. The practical significance of point defects in nanoscale devices prompts accurate, quantitative studies of their atomic and electronic structure within the framework of state-of-the-art computational methods for atomistic simulations. We stress that such studies are imperative for the realization of post-Si Moore's Law scaling. To this end, a promising direction for future studies would be, in addition to addressing the limitations mentioned above, the characterization of defects in other III-V materials within a consistent first-principles framework, i.e. the DFT+GW methodology, to investigate the possibility of universal trends in defect properties, such as charge transition level positions and hence Fermi-level pinning, across these materials. Also, utilization of *ab-initio* transport simulators with at least 2-dimensional periodic boundary conditions would allow for the study of electronic transport throughout the III-V/high-k interface. Inclusion of a gating field in such a study would enable an *ab-initio* evaluation of the electrical properties of highly realistic models of III-V FETs in the presence of point defects, allowing for rigorous comparison between theory and experiment. For CNTs, multi-dimensional transport simulations would allow for the study of CNT arrays, which would be highly beneficial as a first principles assessment of realistic CNT interconnects.

References

1. Derbyshire, K., *ITRS 2010: A More-than-Moore Roadmap?* Solid State Technology, 2011. **54**(3): p. 7-7.
2. Drabold, D.A. and S.K. Estreicher, *Theory Of Defects In Semiconductors*. Topics in Applied Physics. Vol. 104. 2007: Springer. 295 p.
3. Freysoldt, C., et al., *First-principles calculations for point defects in solids.* Reviews of Modern Physics, 2014. **86**(1): p. 253-305.
4. Lany, S. and A. Zunger, *Accurate prediction of defect properties in density functional supercell calculations.* Modelling and Simulation in Materials Science and Engineering, 2009. **17**(8): p. 084002 (14).
5. Van de Walle, C.G. and J. Neugebauer, *First-principles calculations for defects and impurities: Applications to III-nitrides.* Journal of Applied Physics, 2004. **95**(8): p. 3851-3879.
6. Djara, V., et al., *Electrically active interface defects in the In0.53Ga0.47As MOS system.* Microelectronic Engineering, 2013. **109**: p. 182-188.
7. Gomez-Navarro, C., et al., *Tuning the conductance of single-walled carbon nanotubes by ion irradiation in the Anderson localization regime.* Nature Materials, 2005. **4**(7): p. 534-539.
8. Zanolli, Z. and J.C. Charlier, *Defective carbon nanotubes for single-molecule sensing.* Physical Review B, 2009. **80**(15): p. 155447 (6).
9. Robertson, J. and L. Lin, *Bonding principles of passivation mechanism at III-V-oxide interfaces.* Applied Physics Letters, 2011. **99**(22).
10. Biel, B., et al., *Anderson localization in carbon nanotubes: Defect density and temperature effects.* Physical Review Letters, 2005. **95**(26): p. 266801 (4).
11. Flores, F., et al., *Anderson localization regime in carbon nanotubes: size dependent properties.* Journal of Physics-Condensed Matter, 2008. **20**(30): p. 304211 (7).
12. Biel, B., et al., *Ab initio study of transport properties in defected carbon nanotubes: an O(N) approach.* Journal of Physics-Condensed Matter, 2008. **20**(29): p. 294214 (8).
13. Komsa, H.P. and A. Pasquarello, *Comparison of vacancy and antisite defects in GaAs and InGaAs through hybrid functionals.* Journal of Physics-Condensed Matter, 2012. **24**(4).
14. http://en.wikipedia.org/wiki/Moore's_law.
15. ITRS, International Technology Roadmap for Semiconductors [Online] http://www.itrs.net/reports.html.
16. Kaeslin, H., *Digital Integrated Circuit Design : From VLSI Architectures To CMOS Fabrication*2008, Cambridge: Cambridge University Press. xxii, 845 p.

17. Ebert, P., *Atomic structure of point defects in compound semiconductor surfaces.* Current Opinion in Solid State & Materials Science, 2001. **5**(2-3): p. 211-250.
18. Ebert, P., *Nano-scale properties of defects in compound semiconductor surfaces.* Surface Science Reports, 1999. **33**(4-8): p. 121-303.
19. Ebert, P., *Defects in III-V semiconductor surfaces.* Applied Physics a-Materials Science & Processing, 2002. **75**(1): p. 101-112.
20. Shockley, W. and W.T. Read, *Statistics of the Recombinations of Holes and Electrons.* Physical Review, 1952. **87**(5): p. 835-842.
21. Sze, S.M., *Physics Of Semiconductor Devices*1969, New York,: Wiley-Interscience. xiv, 812 p.
22. Hedstrom, M., et al., *Quasiparticle corrections to the electronic properties of anion vacancies at GaAs(110) and InP(110).* Physical Review Letters, 2006. **97**(22): p. 266401 (4).
23. Lany, S. and A. Zunger, *Assessment of correction methods for the band-gap problem and for finite-size effects in supercell defect calculations: Case studies for ZnO and GaAs.* Physical Review B, 2008. **78**(23): p. 235104 (25).
24. Datta, S., *Electronic Transport In Mesoscopic Systems*. Cambridge studies in semiconductor physics and microelectronic engineering 1995, Cambridge ; New York: Cambridge University Press. xv, 377 p.
25. Dennard, R.H., et al., *Design of ion-implanted MOSFETs with very small physical dimensions.* IEEE Journal of Solid-State Circuits, 1974. **Sc 9**(5): p. 256-268.
26. del Alamo, J.A., *Nanometre-scale electronics with III-V compound semiconductors.* Nature, 2011. **479**(7373): p. 317-323.
27. Frank, D.J., *Power-constrained CMOS scaling limits.* IBM Journal of Research and Development, 2002. **46**(2-3): p. 235-244.
28. Spicer, W.E., et al., *Unified defect model and beyond.* Journal of Vacuum Science & Technology, 1980. **17**(5): p. 1019-1027.
29. Ye, P.D., et al., *Inversion-mode In(x)Ga(1-x)As MOSFETs (x=0.53,0.65,0.75) with atomic-layer-deposited high-k dielectrics.* Silicon Nitride, Silicon Dioxide, and Emerging Dielectrics 10, 2009. **19**(2): p. 605-614.
30. Hurley, P.K., et al., *Structural and Electrical Properties of HfO2/n-InxGa1-xAs structures (x: 0, 0.15, 0.3 and 0.53).* Physics and Technology of High-K Gate Dielectrics 7, 2009. **25**(6): p. 113-127.
31. Lin, L. and J. Robertson, *Defect states at III-V semiconductor oxide interfaces.* Applied Physics Letters, 2011. **98**(8): p. 082903 (3).
32. Miceli, G. and A. Pasquarello, *Defect energy levels of the As-As dimer at InGaAs/oxide interfaces: A first principles study.* Microelectronic Engineering, 2013. **109**: p. 60-63.
33. Gu, J.J.J., et al., *Size-Dependent-Transport Study of In0.53Ga0.47As Gate-All-Around Nanowire MOSFETs: Impact of Quantum Confinement and Volume Inversion.* IEEE Electron Device Letters, 2012. **33**(7): p. 967-969.

34. Datta, S., et al., *Ultrahigh-speed 0.5 V supply voltage In0.7Ga0.3As quantum-well transistors on silicon substrate.* IEEE Electron Device Letters, 2007. **28**(8): p. 685-687.
35. Zhang, P.F., et al., *The structural and electrical properties of the SrTa2O6/In0.53Ga0.47As/InP system.* Microelectronic Engineering, 2011. **88**(7): p. 1054-1057.
36. Brammertz, G., et al., *Electrical properties of III-V/oxide interfaces.* Graphene and Emerging Materials for Post-Cmos Applications, 2009. **19**(5): p. 375-386.
37. Hwang, Y., et al., *Analysis of trap state densities at HfO2/In0.53Ga0.47As interfaces.* Applied Physics Letters, 2010. **96**(10): p. 102910 (3).
38. Ali, A., et al., *Small-Signal Response of Inversion Layers in High-Mobility In(0.53)Ga(0.4)7As MOSFETs Made With Thin High-kappa Dielectrics.* Ieee Transactions on Electron Devices, 2010. **57**(4): p. 742-748.
39. LaBella, V.P., et al., *Atomic structure of the GaAs(001)-(2 x 4) surface resolved using scanning tunneling microscopy and first-principles theory.* Physical Review Letters, 1999. **83**(15): p. 2989-2992.
40. Komsa, H.P. and A. Pasquarello, *Dangling bond charge transition levels in AlAs, GaAs, and InAs.* Applied Physics Letters, 2010. **97**(19): p. 191901 (3).
41. Komsa, H.P. and A. Pasquarello, *Identification of defect levels at InxGa1-xAs/oxide interfaces through hybrid functionals.* Microelectronic Engineering, 2011. **88**(7): p. 1436-1439.
42. Komsa, H.P. and A. Pasquarello, *Intrinsic defects in GaAs and InGaAs through hybrid functional calculations.* Physica B-Condensed Matter, 2012. **407**(15): p. 2833-2837.
43. Lin, L. and J. Robertson, *Defect states at III-V semiconductor oxide interfaces.* Applied Physics Letters, 2011. **98**(8): p. 082903.
44. Rinke, P., et al., *Defect formation energies without the band-gap problem: Combining density-functional theory and the GW approach for the silicon self-Interstitial.* Physical Review Letters, 2009. **102**(2): p. 026402 (4).
45. Guo, Y.Z. and J. Robertson, *Chemical trends and passivation of defects at Al2O3:GaAs/InAs/InP/GaSb interfaces.* Microelectronic Engineering, 2013. **109**: p. 274-277.
46. Djara, V., et al., *Impact of forming gas annealing on the performance of surface-channel In0.53Ga0.47As MOSFETs with an ALD Al2O3 gate dielectric.* IEEE Transactions on Electron Devices, 2012. **59**(4): p. 1084-1090.
47. O'Connor, E., et al., *A systematic study of (NH4)(2)S passivation (22%, 10%, 5%, or 1%) on the interface properties of the Al2O3/In0.53Ga0.47As/InP system for n-type and p-type In0.53Ga0.47As epitaxial layers.* Journal of Applied Physics, 2011. **109**(2): p. 024101 (10).

48. Brammertz, G., et al., *Electrical properties of III-V/oxide interfaces.* ECS Transactions, 2009. **19**(5): p. 375-386.
49. Broqvist, P., A. Alkauskas, and A. Pasquarello, *A hybrid functional scheme for defect levels and band alignments at semiconductor-oxide interfaces.* Physica Status Solidi a-Applications and Materials Science, 2010. **207**(2): p. 270-276.
50. Hybertsen, M.S. and S.G. Louie, *Electron correlation in semiconductors and insulators - band-gaps and quasi-particle energies.* Physical Review B, 1986. **34**(8): p. 5390-5413.
51. Aryasetiawan, F. and O. Gunnarsson, *The GW method.* Reports on Progress in Physics, 1998. **61**(3): p. 237-312.
52. Hybertsen, M.S. and S.G. Louie, *Theory of quasiparticle surface-states in semiconductor surfaces.* Physical Review B, 1988. **38**(6): p. 4033-4044.
53. Bushmaker, A.W., et al., *Memristive behavior observed in a defected single-walled carbon nanotube.* IEEE Transactions on Nanotechnology, 2011. **10**(3): p. 582-586.
54. Ganzhorn, M., et al., *Strong spin-phonon coupling between a single-molecule magnet and a carbon nanotube nanoelectromechanical system.* Nature Nanotechnology, 2013. **8**(3): p. 165-169.
55. Collins, P.G., H. Bando, and A. Zettl, *Nanoscale electronic devices on carbon nanotubes.* Nanotechnology, 1998. **9**(3): p. 153-157.
56. Narlikar, A.V. and Y.Y. Fu, *The Oxford Handbook Of Nanoscience And Technology. Volume III of III, Applications* 2010, Oxford: Oxford University Press. 936 p.
57. Awano, Y., et al., *Carbon nanotubes for VLSI: interconnect and transistor applications.* Proceedings of the IEEE, 2010. **98**(12): p. 2015-2031.
58. Saito, R., G. Dresselhaus, and M.S. Dresselhaus, *Physical Properties Of Carbon Nanotubes* 1998, London: Imperial College Press. xii, 259 p.
59. Fiori, G., G. Iannaccone, and G. Klimeck, *Performance of carbon nanotube field effect transistors with doped source and drain extensions and arbitrary geometry.* IEEE International Electron Devices Meeting 2005, Technical Digest, 2005: p. 529-532.
60. Ceric, H. and S. Selberherr, *Electromigration in submicron interconnect features of integrated circuits.* Materials Science & Engineering R-Reports, 2011. **71**(5-6): p. 53-86.
61. Srivastava, N., et al., *On the applicability of single-walled carbon nanotubes as VLSI interconnects.* IEEE Transactions on Nanotechnology, 2009. **8**(4): p. 542-559.
62. Chai, Y., et al., *Electromigration studies of Cu/carbon nanotube composite interconnects using Blech structure.* Ieee Electron Device Letters, 2008. **29**(9): p. 1001-1003.
63. Telling, R.H., et al., *Wigner defects bridge the graphite gap.* Nature Materials, 2003. **2**(5): p. 333-337.

64. Charlier, J.C., *Defects in carbon nanotubes.* Accounts of Chemical Research, 2002. **35**(12): p. 1063-1069.
65. Ebert, P., et al., *Nanoscale fluctuations in the distribution of dopant atoms: Dopant-induced dots and roughness of electronic interfaces.* Journal of Vacuum Science & Technology B, 2004. **22**(4): p. 2018-2025.
66. Marquardt, C.W., et al., *Reversible metal-insulator transitions in metallic single-walled carbon nanotubes.* Nano Letters, 2008. **8**(9): p. 2767-2772.
67. Krasheninnikov, A.V., K. Nordlund, and J. Keinonen, *Production of defects in supported carbon nanotubes under ion irradiation.* Physical Review B, 2002. **65**(16): p. 165423 (8).
68. Amorim, R.G., et al., *Divacancies in graphene and carbon nanotubes.* Nano Letters, 2007. **7**(8): p. 2459-2462.
69. Krupke, R., et al., *Separation of metallic from semiconducting single-walled carbon nanotubes.* Science, 2003. **301**(5631): p. 344-347.
70. Arnold, M.S., et al., *Sorting carbon nanotubes by electronic structure using density differentiation.* Nature Nanotechnology, 2006. **1**(1): p. 60-65.
71. Tu, X.M., et al., *DNA sequence motifs for structure-specific recognition and separation of carbon nanotubes.* Nature, 2009. **460**(7252): p. 250-253.
72. Shimauchi, H., et al., *Suppression of hysteresis in carbon nanotube field-effect transistors: Effect of contamination induced by device fabrication process.* Japanese Journal of Applied Physics 2006. **45**(6B): p. 5501-5503.
73. Fagas, G. and J.C. Greer, *Ballistic conductance in oxidized Si nanowires.* Nano Letters, 2009. **9**(5): p. 1856-1860.
74. Giannozzi, P., et al., *Quantum ESPRESSO: A modular and open-source software project for quantum simulations of materials.* Journal of Physics-Condensed Matter, 2009. **21**(39): p. 395502 (19).
75. Marini, A., et al., *YAMBO: An ab initio tool for excited state calculations.* Computer Physics Communications, 2009. **180**(8): p. 1392-1403.
76. Ozaki, T., K. Nishio, and H. Kino, *Efficient implementation of the nonequilibrium Green function method for electronic transport calculations*. Physical Review B, 2010. **81**(3): p. 035116 (19).
77. Oppenheimer, R. and M. Born, *Zur Quantentheorie der Molekeln.* Annalen Der Physik, 1927. **84**(20): p. 457-484.
78. Martin, R.M., *Electronic Structure : Basic Theory And Practical Methods* 2004, Cambridge ; New York: Cambridge University Press. xxiii, 624 p.
79. Kohn, W., *Nobel Lecture: Electronic structure of matter-wave functions and density functionals.* Reviews of Modern Physics, 1999. **71**(5): p. 1253-1266.
80. James, H.M. and A.S. Coolidge, *The ground state of the hydrogen molecule.* Journal of Chemical Physics, 1933. **1**(12): p. 825-833.

81. Thomas, L.H., *The calculation of atomic fields*. Mathematical Proceedings of the Cambridge Philosophical Society, 1927. **23**(05): p. 542-548.
82. Fermi, E., *Un Metodo Statistico per la Determinazione di alcune Prioprietà dell'Atomo*. Rend. Accad. Naz. Lincei, 1927. **6**: p. 602-607.
83. Kohn, W. and L.J. Sham, *Self-consistent equations including exchange and correlation effects*. Physical Review, 1965. **140**(4A): p. 1133-1138.
84. Ceperley, D.M. and B.J. Alder, *Ground-state of the electron-gas by a stochastic method*. Physical Review Letters, 1980. **45**(7): p. 566-569.
85. Gunnarsson, O., M. Jonson, and B.I. Lundqvist, *Descriptions of exchange and correlation effects in inhomogeneous electron-systems*. Physical Review B, 1979. **20**(8): p. 3136-3164.
86. Feynman, R.P., *Forces and stresses in molecules*, 1939, Massachusetts Institute of Technology. p. 340-343.
87. Johnson, D.D., *Modified broyden method for accelerating convergence in self-consistent calculations*. Physical Review B, 1988. **38**(18): p. 12807-12813.
88. Kresse, G. and J. Furthmuller, *Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set*. Physical Review B, 1996. **54**(16): p. 11169-11186.
89. Perdew, J.P. and A. Zunger, *Self-interaction correction to density-functional approximations for many-electron systems*. Physical Review B, 1981. **23**(10): p. 5048-5079.
90. Godby, R.W., M. Schluter, and L.J. Sham, *Accurate exchange-correlation potential for silicon and its discontinuity on addition of an electron*. Physical Review Letters, 1986. **56**(22): p. 2415-2418.
91. Perdew, J.P. and M. Levy, *Physical content of the exact kohn-sham orbital energies - band-gaps and derivative discontinuities*. Physical Review Letters, 1983. **51**(20): p. 1884-1887.
92. Payne, M.C., et al., *Iterative minimization techniques for ab initio total-energy calculations - molecular-dynamics and conjugate gradients*. Reviews of Modern Physics, 1992. **64**(4): p. 1045-1097.
93. Ozaki, T., *Efficient low-order scaling method for large-scale electronic structure calculations with localized basis functions*. Physical Review B, 2010. **82**(7): p. 075131 (17).
94. Ohfuchi, M., T. Ozaki, and C. Kaneta, *Large-scale electronic transport calculations of finite-length carbon nanotubes bridged between graphene electrodes with lithium-intercalated contact*. Applied Physics Express, 2011. **4**(9): p. 095101 (3).
95. Ozaki, T., *Variationally optimized atomic orbitals for large-scale electronic structures*. Physical Review B, 2003. **67**(15): p. 155108 (5).
96. Ozaki, T. and H. Kino, *Numerical atomic basis orbitals from H to Kr*. Physical Review B, 2004. **69**(19): p. 195113 (19).
97. Harrison, W.A., *Pseudopotentials in the theory of metals*. Frontiers in physics1966, New York,: W.A. Benjamin. xv, 336 p.
98. http://en.wikipedia.org/wiki/Pseudopotential.

99. Giannozzi, P., et al., *QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials.* Journal of Physics-Condensed Matter, 2009. **21**(39).
100. Lehmann, D., Zahn D.R.T., *Photoemission Spectroscopy - Organic/inorganic interfaces: From band dispersion to chemical reaction* Demand and Characteristics of a New VUVis-beamline at BESSY II, 30.05.2011, Berlin, Germany, 2011.
101. Onida, G., L. Reining, and A. Rubio, *Electronic excitations: density-functional versus many-body Green's-function approaches.* Reviews of Modern Physics, 2002. **74**(2): p. 601-659.
102. Bruus, H. and K. Flensberg, *Many-Body Quantum Theory In Condensed Matter Physics : An Introduction*. Oxford graduate texts 2004, Oxford ; New York: Oxford University Press. xix, 435 p.
103. Aulbur, W.G., L. Jonsson, and J.W. Wilkins, *Quasiparticle calculations in solids.* Solid State Physics: Advances in Research and Applications, Vol. 54, 2000. **54**: p. 1-218.
104. Hedin, L., et al., *Effects of electron-electron interaction on one-electron properties of atoms and solids.* Arkiv for Fysik, 1969. **39**(2): p. 97-112.
105. Hybertsen, M.S. and S.G. Louie, *First-principles theory of quasiparticles - calculation of band-gaps in semiconductors and insulators.* Physical Review Letters, 1985. **55**(13): p. 1418-1421.
106. Wendin, G.r., *Breakdown of the one-electron pictures in photoelectron spectra*. Structure and Bonding 1981, Berlin; New York: Springer-Verlag. vii, 30 p.
107. Marsh, A.C. and J.C. Inkson, *Scattering matrix theory of transport in heterostructures.* Semiconductor Science and Technology, 1986. **1**(4): p. 285-290.
108. Rocha, A.R., et al., *Spin and molecular electronics in atomically generated orbital landscapes.* Physical Review B, 2006. **73**(8).
109. Imry, Y., *Introduction To Mesoscopic Physics*. 2nd ed 2009, Oxford: Oxford University Press. xiii, 236 p.
110. Beenakker, C.W.J. and H. Vanhouten, *Quantum transport in semiconductor nanostructures.* Solid State Physics-Advances in Research and Applications, 1991. **44**: p. 1-228.
111. Beenakker, C.W.J., *Random-matrix theory of quantum transport.* Reviews of Modern Physics, 1997. **69**(3): p. 731-808.
112. Makov, G. and M.C. Payne, *Periodic boundary-conditions in ab-initio calculations.* Physical Review B, 1995. **51**(7): p. 4014-4022.
113. Zunger, A., *Practical doping principles.* Applied Physics Letters, 2003. **83**(1): p. 57-59.
114. Marini, A., G. Onida, and R. Del Sole, *Quasiparticle electronic structure of copper in the GW approximation.* Physical Review Letters, 2002. **88**(1): p. 016403 (4).
115. Lykke, K.R. and P. Wurz, *Direct detection of neutral products from photodissociated-C60.* Journal of Physical Chemistry, 1992. **96**(8): p. 3191-3193.

116. Sharma, D., et al., *Transport properties and electrical device characteristics with the TiMeS computational platform: Application in silicon nanowires*. Journal of Applied Physics, 2013. **113**(20): p. 203708 (8).
117. Gutierrez, R., et al., *Theory of an all-carbon molecular switch.* Physical Review B, 2002. **65**(11): p. 113410 (4).
118. Gutierrez, R., et al., *Conductance of a molecular junction mediated by unconventional metal-induced gap states.* Europhysics Letters, 2003. **62**(1): p. 90-96.
119. Markussen, T., et al., *Scaling theory put into practice: First-principles modeling of transport in doped silicon nanowires*. Physical Review Letters, 2007. **99**(7): p. 076803 (4).
120. Dargam, T.G., R.B. Capaz, and B. Koiller, *Disorder and size effects in the envelope-function approximation.* Physical Review B, 1997. **56**(15): p. 9625-9629.
121. Degironcoli, S., P. Giannozzi, and S. Baroni, *Structure and thermodynamics of sixge1-X alloys from ab initio Monte-Carlo simulations.* Physical Review Letters, 1991. **66**(16): p. 2116-2119.
122. Bellaiche, L. and D. Vanderbilt, *Virtual crystal approximation revisited: Application to dielectric and piezoelectric properties of perovskites.* Physical Review B, 2000. **61**(12): p. 7877-7882.
123. Bachelet, G.B., D.R. Hamann, and M. Schluter, *Pseudopotentials that work - from H to Pu*. Physical Review B, 1982. **26**(8): p. 4199-4228.
124. Madelung, O., *Semiconductors: Data handbook*. 3rd ed2004, Berlin ; New York: Springer. 691 p.
125. Vurgaftman, I., J.R. Meyer, and L.R. Ram-Mohan, *Band parameters for III-V compound semiconductors and their alloys*. Journal of Applied Physics, 2001. **89**(11): p. 5815-5875.
126. Pickett, W.E., H. Krakauer, and P.B. Allen, *Smooth Fourier Interpolation of Periodic-Functions.* Physical Review B, 1988. **38**(4): p. 2721-2726.
127. Yang, L., et al., *Quasiparticle energies and band gaps in graphene nanoribbons*. Physical Review Letters, 2007. **99**(18): p. 186801 (4).
128. Liang, L.B., E.C. Girao, and V. Meunier, *Quasiparticle band gaps of graphene nanowiggles and their magnetism on Au(111)*. Physical Review B, 2013. **88**(3): p. 035420 (4).
129. Fischetti, M., private communication, 2014.
130. O'Regan, T.P., et al., *Modeling the capacitance-voltage response of In0.53Ga0.47As metal-oxide-semiconductor structures: Charge quantization and nonparabolic corrections.* Applied Physics Letters, 2010. **96**(21): p. 213514 (3).
131. Heyd, J., G.E. Scuseria, and M. Ernzerhof, *Hybrid functionals based on a screened Coulomb potential.* Journal of Chemical Physics, 2003. **118**(18): p. 8207-8215.

132. Schulz, S., et al., *Composition-dependent band gap and band-edge bowing in AlInN: A combined theoretical and experimental study*. Applied Physics Express, 2013. **6**(12): p. 121001 (5).
133. Schwarz, G., et al., *Electronic and structural properties of vacancies on and below the GaP(110) surface*. Physical Review B, 1998. **58**(3): p. 1392-1400.
134. Murphy, S.T., et al., *Phase stability and the arsenic vacancy defect in InxGa1-xAs*. Physical Review B, 2011. **84**(18): p. 184108 (7).
135. Saraswat, K.C., et al., *High mobility materials and novel device structures for high performance nanoscale MOSFETs*. 2006 International Electron Devices Meeting, Vols 1 and 2, 2006: p. 395-398.
136. Cherkaoui, K., et al., *Investigation of High-k/InxGa1-xAs Interfaces*. Dielectrics for Nanosystems 4: Materials Science, Processing, Reliability, and Manufacturing, 2010. **28**(2): p. 181-190.
137. Cherkaoui, K., et al., *Can metal/Al2O3/In0.53Ga0.47As/InP MOSCAP properties translate to metal/Al2O3/In0.53Ga0.47As/InP MOSFET characteristics?* ECS Transactions, 2012. **45**(3): p. 79-88.
138. Zhang, S.B., S.H. Wei, and A. Zunger, *Microscopic origin of the phenomenological equilibrium "doping limit rule" in n-type III-V semiconductors*. Physical Review Letters, 2000. **84**(6): p. 1232-1235.
139. Lin, L. and J. Robertson, *Passivation of interfacial defects at III-V oxide interfaces*. Journal of Vacuum Science & Technology B, 2012. **30**(4): p. 04E101 (14).
140. Melitz, W., et al., *Atomic imaging of atomic layer deposition oxide nucleation with trimethylaluminum on As-rich InGaAs(001) 2 x 4 vs Ga/In-rich InGaAs(001) 4 x 2*. Journal of Chemical Physics, 2012. **136**(15).
141. Wei, S.H., et al., *Electronic properties of random alloys - special quasirandom structures*. Physical Review B, 1990. **42**(15): p. 9622-9649.
142. Zunger, A., et al., *Special quasirandom structures*. Physical Review Letters, 1990. **65**(3): p. 353-356.
143. Lin, H.C., et al., *Electrical study of sulfur passivated In0.53Ga0.47As MOS capacitor and transistor with ALD Al2O3 as gate insulator*. Microelectronic Engineering, 2009. **86**(7-9): p. 1554-1557.
144. Capaz, R.B., K. Cho, and J.D. Joannopoulos, *Signatures of bulk and surface arsenic antisite defects in GaAs(100)*. Physical Review Letters, 1995. **75**(9): p. 1811-1814.
145. Sharma, D., H.H. Arefi, and G. Fagas, *Atomic basis sets for first-principles studies of Si nanowires*. Computational and Theoretical Chemistry, 2012. **991**: p. 32-39.
146. Demkov, A.A., et al., *Electronic-structure approach for complex silicas*. Physical Review B, 1995. **52**(3): p. 1618-1630.
147. Blase, X., et al., *hybridization effects and metallicity in small radius carbon nanotubes*. Physical Review Letters, 1994. **72**(12): p. 1878-1881.

148. Zolyomi, V. and J. Kurti, *First-principles calculations for the electronic band structures of small diameter single-wall carbon nanotubes.* Physical Review B, 2004. **70**(8): p. 085403 (8).
149. Gulseren, O., T. Yildirim, and S. Ciraci, *Systematic ab initio study of curvature effects in carbon nanotubes.* Physical Review B, 2002. **65**(15).
150. Lu, X. and Z.F. Chen, *Curved Pi-conjugation, aromaticity, and the related chemistry of small fullerenes (< C-60) and single-walled carbon nanotubes.* Chemical Reviews, 2005. **105**(10): p. 3643-3696.
151. Matsuda, Y., J. Tahir-Kheli, and W.A. Goddard, *Definitive band gaps for single-wall carbon nanotubes.* Journal of Physical Chemistry Letters, 2010. **1**(19): p. 2946-2950.
152. Chai, Y., Z.Y. Xiao, and P.C.H. Chan, *Horizontally aligned carbon nanotube bundles for interconnect application: diameter-dependent contact resistance and mean free path.* Nanotechnology, 2010. **21**(23): p. 235705 (5).