Physics of Existing and Novel Transparent Conducting Oxide Semiconductors

by

Jack E. N. Swallow

Contents

Acknowledgments

A great many people have contributed to helping me finish this body of work, without whom this would not have been possible. Here I would like to give my thanks to the people who have provided the most notable contributions.

I would firstly like to express my utmost gratitude to my supervisor Prof. Tim Veal. Tim welcomed me to his group at a time when I felt a lack of direction, and he provided me with fantastic opportunities and imparted a great deal of knowledge and scientific enthusiasm, without which I would not be the scientist I have become. I must also thank him for his guidance, encouragement, and friendship he has given me over my time here in Liverpool, and can only apologise for the constant nagging and unsavoury jokes which I subjected him to on a daily basis. My thanks are extended to Prof. Vin Dhanak, who allowed me use of his lab, where I developed a keen interest in surface science and XPS.

I have been tremendously lucky that so much of my work has been performed on a collaborative basis. For this, I need to thank the people who performed experiments, calculations, and generally assisted in contributing to "the story". A massive thank you is owed to David Scanlon (Big Dave) and Ben Williamson from UCL, who performed DFT calculations for many of our TCO materials. Without them this work would not have been possible. In a similar capacity, I would like to thank Joel Varley at Laurence Livermore National Laboratories, for his contribution to all things Ga_2O_3 related. I am grateful to Louis Piper at Binghamton University for XAS measurements and assistance. During my PhD I have spent many hours locked in the control cabin of I09 at Diamond Light Source (and loved every second of it!) For this I thank the principal beamline scientist, Tien-Lin Lee, who entrusted us with his creation, and was always helpful and informative. I would also like to thank Pardeep Thakur for his assistance in these experiments. I am very grateful to Anna Regoutz at UCL who put up with me on many trips to Diamond, and allowed me to collect such good data. My sincere gratitude goes to Paul Warren, Alex Abbott, and Mark Farnworth at NSG who supported me throughout my PhD.

I would like to thank Max Birkett for helping me develop my coding, and for imparting his knowledge and wisdom in so many areas of physics. I would also like to thank Tom Whittles and David Hesp for guiding me into the world of all things UHV related. I would like to thank Tom Baines, hopefully living with me wasn't too traumatic. I would also like to give my appreciation to everybody in the Stephenson Institute who has been involved in this work, and those who were responsible for making the SIRE

such an enjoyable place to work, for which particular thanks must be given to Thomas Featherstone (Flaps), Matt Smiles, James Gibbon, Tom Shalvey (Shelves), Huw Shiel, Nicole Fleck, and Phil Murgatroyd (pea head). There are of course so many others, I couldn't possibly name them all here.

Extra thanks are reserved for my family. As a wise woman once said, "families are a complicated thing", and ours probably more so than most. However, whether its been through nature or nurture I would not be the man I am today without you all. In so many cases, your love and support have allowed me to peruse this. I would like to thank my Dad, Grandad, sisters and brother, who have provided varied and extensive support during my time at Liverpool, and indeed all other times.

Finally, a very special thanks is reserved for the three people who have had the biggest influence on my life. To my Nan, for giving me strength and support, stability and guidance, and of course a large amount of wit. To my Mum, for her generosity and for instilling in me the work ethic I have required to finally finish this thesis. To Leanne, I cannot hope to put into words my appreciation for all you have done for me. For this and so much more, I dedicate this thesis to you three, as a small token of what you mean to me.

Declaration

I declare that this thesis contains an account of my research carried out in the Department of Physics at the University of Liverpool between October 2015 and September 2019 under the supervision of Prof. T. D. Veal. The research reported here has not been submitted, either wholly or in part, in this or any other academic institution for admission to a higher degree.

The SIMS standards used in chapters 4 and 5 were ion implanted by N. Peng (at Surrey ion beam centre), and the associated SIMS data were obtained by M. Farnworth (at Pilkington Technology Centre). XAS measurements shown in chapter 5 were taken by Z. W. Lebens-Higgins in the group of L. Piper (at beamline 23-ID of the National Synchrotron Light Source II) . DFT calculations in chapters 4 and 5 were performed and analysed by B. A. D. Williamson in the group of D. O. Scanlon (at University College London), and in chapter 6 by J. B. Varley (at Lawrence Livermore National Laboratory). IMO and ITO films were deposited, and subsequent SEM images taken by, S. Sathasivam in the group of I. Parkin (at University College London). Calculations solving the Poisson equation shown in chapter 6 were performed by T. D. Veal. Apart from DFT calculations, all interpretation, analytical and theoretical work pertaining to these data were performed by the author.

Jack E. N. Swallow
September 2019

Several articles based on this research have been published:

- **J. E. N. Swallow**, B. A. D. Williamson, T. J. Whittles, M. Birkett, T. J. Featherstone, N. H. Peng, A. Abbott, M. Farnworth, K. J. Cheetham, P. Warren, D. O. Scanlon, V. R. Dhanak, and T. D. Veal, Self-Compensation in Transparent Conducting F-Doped SnO_2. Adv. Funct. Mater. 28, 1701900 (2018).

- **J. E. N. Swallow**, B. A. D. Williamson, S. Sathasivam, M. Birkett, T. J. Featherstone, P. A. E. Murgatroyd, H. J. Edwards, Z. W. Lebens-Higgins, D. A. Duncan, M. Farnworth, P. Warren, N. Peng, T-L. Lee, L. F. J. Piper, A. Regoutz, C. J. Carmalt, I. P. Parkin, V. R. Dhanak, D. O. Scanlon and T. D. Veal, Resonant doping for high mobility transparent conductors: the case of Mo-doped In_2O_3. Mater. Horiz. 7, 236–243 (2020).

- **J. E. N. Swallow**, J. B. Varley, L. A. H. Jones, J. T. Gibbon, L. F. J. Piper, V. R. Dhanak, and T. D. Veal, Transition from electron accumulation to depletion at β-Ga_2O_3 surfaces: The role of hydrogen and the charge neutrality level. APL Materials 7, 022528 (2019).

The author also contributed to the following other publications:

- M. J. Wahila, G. Paez, C. N. Singh, A. Regoutz, S. Sallis, M. J. Zuba, J. Rana, M. B. Tellekamp, J. E. Boschker, T. Markurt, **J. E. N. Swallow**, L. A. H. Jones, T. D. Veal, W. Yang, T-L. Lee, F. Rodolakis, J. T. Sadowski, D. Prendergast, W-C. Lee, W. A. Doolittle, and L. F. J. Piper, Evidence of a second-order Peierls-driven metal-insulator transition in crystalline NbO_2. Phys. Rev. Materials 3, 074602 (2019).

- B. A. D. Williamson, T. J. Featherstone, S. Sathasivam, **J. E. N. Swallow**, H. Shiel, L. A. H. Jones, M. Smiles, A. Regoutz, T-L. Lee, X. Xia, C. Blackman, P. Thakur, C. J. Carmalt, I. P. Parkin, T. D. Veal, D. O. Scanlon, Resonant Ta Doping for Enhanced Mobility in Transparent Conducting SnO_2. Chem. Mater. 5, 1964–1973 (2020)

Abstract

This thesis explores some of the fundamental properties of a rather special class of materials, the transparent conducting oxides. These 'semiconductors' combine the usually mutually exclusive properties of optical transparency with high electronic conductivity. Conventionally, TCOs are formed from metal-oxide structures doped with an element from the right hand column in the periodic table of either the anion (e.g. F in SnO_2) or cation (Sn in In_2O_3).

However, TCOs that use unconventional dopants (i.e. dopant elements not from the subsequent column in the periodic table) which display much improved optoelectronic properties are quite frequently reported in the literature. Most often, the host materials are doped with transition metal elements which can display unusual electronic configurations and many common oxidation states, making their properties as dopants hard to predict and understand.

In this thesis, the properties of selected TCOs are presented in three case studies. The first study is on a conventionally doped and commercially available TCO, $F:SnO_2$, in which the limitation of the electronic performance of the material is attributed to extrinsic defects associated with the dopant element. The second is a comparison between an In_2O_3 system doped with the transition metal Mo, and its conventionally doped and commercially relevant counterpart $Sn:In_2O_3$. In this study, the system with the novel dopant ($Mo:In_2O_3$) is shown to display much improved properties over the conventionally doped system. This is explained in the context of the band structure modifications due to choice of dopant. Finally, the surface properties of a relatively novel oxide semiconductor system, β-Ga_2O_3, are investigated. β-Ga_2O_3 has gained a lot of attention in the literature recently, and is a rare case of a III-VI oxide whose properties are not well established. The natural surface electronic state of this material is determined and the effect that surface H adsorption has on the surface space charge is investigated.

1

Introduction

1.1 Overview

Wide band gap materials by their very nature tend to be insulating, with many also displaying a high degree of optical transparency in the visible range, take for example silica glass. On the other hand, conducting materials such as bulk metals are optically opaque to visible light, instead displaying a high degree of reflectivity. There do however, exist a few exceptions to this general rule, one such exception being the class of materials known as the transparent conducting oxides (TCOs). The fortuitous combination of high transparency and electrical conductivity means TCOs have generated a great deal of scientific interest since their discovery in 1902, where F. Streintz created CdO as a pressed powder[1]. The first thin film was deposited by K. Bädeker in 1907[2], since which thin film TCOs have found applications in a multitude of modern tech-

nologies. These include: acting as transparent contacts in solar cells[3,4] allowing light through to generate carriers in the active layer whilst also enabling current collection; the infrared reflective layer in low emissivity (energy efficient) windows[5], where the high free carrier concentration aids in reflecting infrared light to minimise thermal losses whilst a wide band gap allows visible light to pass; in conductance-type gas sensors[6,7], where atmospheric gases can interact with a sensor material through processes such as surface adsorption, and cause a measurable resistivity variation; as the transparent electrode in touch screen panels[8], enabling a user interface while allowing light through; and in electrochromic windows[9], acting as an electro-active layer, allowing an applied charge to modulate the properties of the electrochromic material, to name but a few.

Since their inception over 100 years ago, it is clear TCOs have become commonplace in consumer electronic devices. However, these materials have the potential to go beyond simply acting as transparent contacts. It is now commonly accepted that silicon-based electronics are quickly approaching their physical and electronic limits[10], and new approaches are sorely needed. TCOs may well provide some breakthroughs required in this area[11−17]. One such stride towards integrating TCOs into more far reaching fields than simply invisible electronic contacts was made by Nomura *et al.*[18] who fabricated an all-oxide transparent field-effect transistor, an impressive step towards fully transparent electronics. It is also worth mentioning that despite commercially produced TCOs typically being manufactured via scalable deposition methods such as chemical vapour deposition (CVD) or magnetron sputtering which tend to create large defect densities in the films, great progress has been made in controlling the crystal growth of oxide semiconductors and oxide heterostructures. This was demonstrated by Tsukazaki *et al.*[19,20] who observed both integer and fractional quantum Hall effect in all oxide heterostructures, a phenomena typically only seen in high-quality

2

semiconductor quantum wells. The extremely high mobilities seen by Tsukazaki *et al.* in their heterostructures (up to 180 000 cm²V⁻¹s⁻¹) maybe be another way of opening new avenues to high performance devices where traditional semiconductors have up to now reined supreme.

CURRENT CLIMATE FOR TRANSPARENT CONDUCTING OXIDES

Currently TCOs are marketed primarily as transparent contacts, and so this thesis focuses on the materials used as such. Surprisingly, despite being a multi-billion pound industry, the TCO market is dominated by only a few materials. This is due to a combination of limiting factors which include the stringent properties expected of a new TCO to give adequate performance and suitability for commercial devices, in addition to some level of industrial reluctance to adopt new technologies. Whilst there is little that can be done by the scientific community about the second point the second point, it is possible to consider the first. Before even thinking about the physical properties of a TCO, market forces have effectively screened the materials that can be used. The largest market for TCOs is in consumer electronics and devices and so they are required to be non-toxic. This rules out materials such as CdO, which displays excellent opto-electronic properties, but is hazardous to health. They also need to be low cost in order to be picked up in the consumer market, and so more abundant elements are preferable. Figure 1.1[21] shows the abundance of elements up to Z=83 available in the earth's crust (top) and in the sea (bottom). This does not account for the difficulty in extracting these materials, the demand on these materials and the availability of them to us. These things being very hard to quantify, considerations are often only given to elemental abundances as a simple guide to availability. On top of this, a TCO needs to be cheap to deposit as large quantities of TCOs are required in the modern market for applications ranging from phone screens, where relatively small areas of coverage are
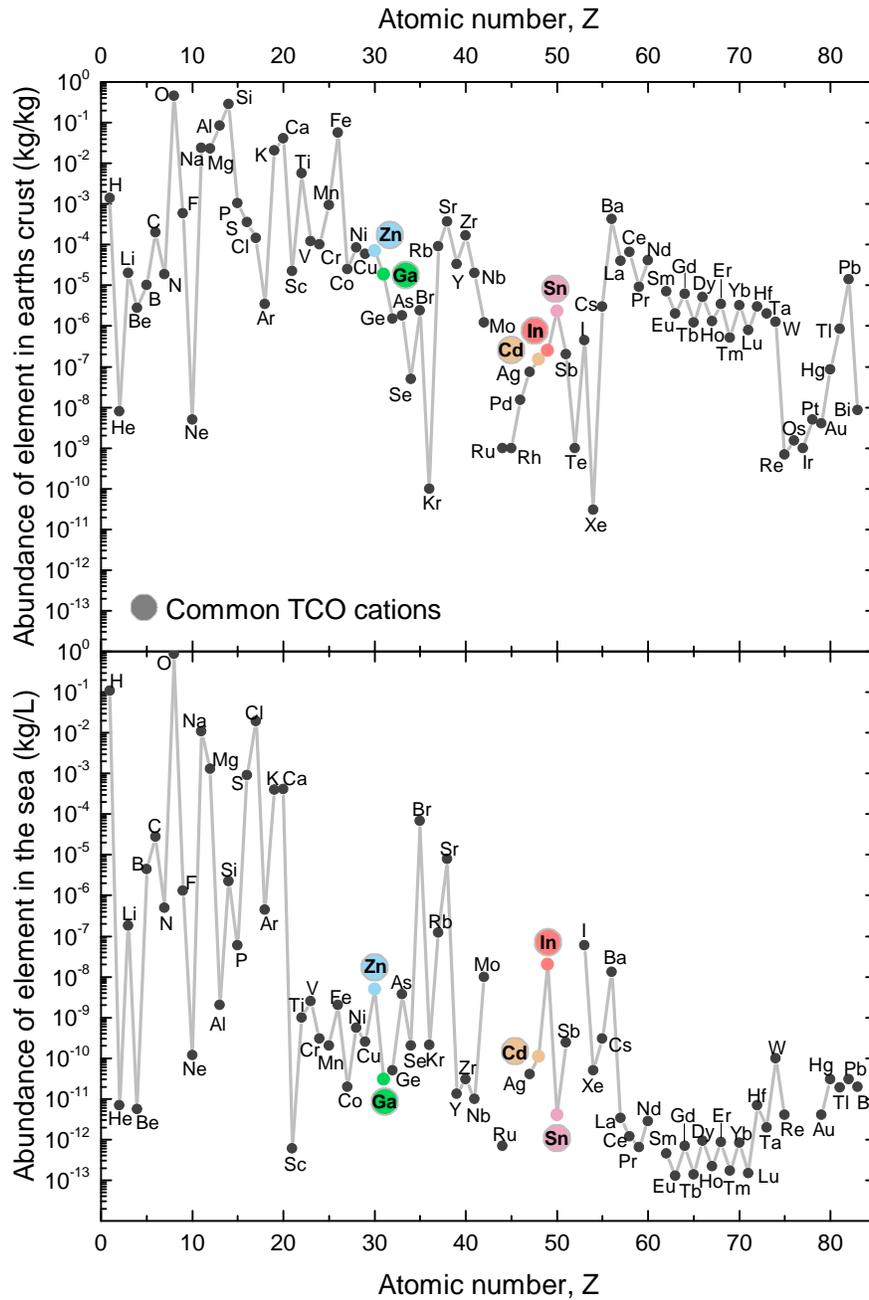
Figure 1.1: Abundance of elements available in the world crust (top) and in the sea (bottom).[21]

required for one handset but many millions of handsets are made per year, to large scale solar panels and solar farms which may require a single TCO to be deposited on the order of square meters.
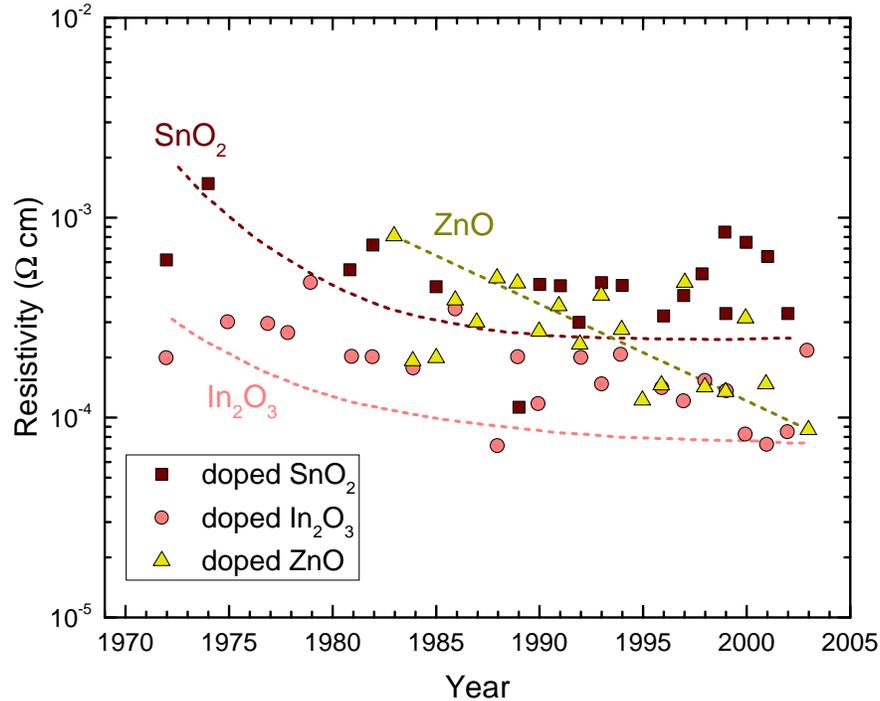
Figure 1.2: The historical development of resistivity for doped SnO, In_2O_3 and ZnO thin films from 1972-2005. Figure adapted from T. Minami. [22]

The market informs us on the practical information relevant to the consumer. The actual physical properties required of a TCO are arguably more easily comprehensible by comparison. A TCO requires a large optical band gap, $E_g^{opt} > 3eV$, large enough to allow for high optical transparency (>80%) but not too large that the material is forced to be insulating. The transmission window is further limited by the relatively diminished intensity of UV light compared to the visible in the solar spectrum. TCOs also need to display good electrical properties, for example a resistivity on the order of $\sim 10^{-4}\Omega cm$. This is usually achieved through intentional inclusion of chemical impurities (doping).

With all of this in mind, practically only a few TCOs that are non-toxic and relatively inexpensive to produce are of industrial importance. The 'major players' are highlighted in figure 1.2 (replicated from a figure by T. Minami [22]), which shows the

historical development in resistivity of impurity doped In_2O_3, SnO_2, and ZnO. Whilst the values are slightly outdated, this figure nicely illustrates how the TCO market has been split up for the past few decades. Sn-doped In_2O_3 (ITO) is the dominant TCO material, accounting for over 60% of the market[23]. Most often deposited via magnetron sputtering, ITO boasts the best conductivities out of the commercial TCOs. However, elemental indium is scarce (see figure 1.1) and displays high price volatility meaning a replacement for this material is sorely sought after. F-doped SnO_2 (FTO) has worse opto-electonic properties by comparison, but is usually used as a cheaper alternative for large scale requirements, such as PV or low emissivity glass. Cost effectiveness is made even more prominent for FTO as it is most often deposited via chemical vapour methods which do not require high vacuum, and it can be deposited on-line immediately after the float glass is made which takes advantage of the heat from the glass furnace. ZnO has seen a growing research interest in recent years because Zn is far more abundant than In and even Sn. This work has promoted a large increase in its performance as a TCO (usually doped with Al), making ZnO a very promising material for the future of TCOs. Difficulties still persist with depositing doped-ZnO in a controlled manner, due in part to Zn being more chemically active in an oxidizing atmosphere than both Sn and In[22]. These issues need to be addressed for ZnO to be considered as a replacement for ITO.

1.2 Fundamentals of Transparent Conducting Oxides

The previous section specified the physical conditions that a transparent conductor needs to fulfil to be successful. They are required to have a large optical gap $E_g^{opt} > 3\text{eV}$, and to display good resistivity on the order of $\sim 10^{-4}\Omega\text{cm}$. Now this section will specify how this is achieved in these materials. The following sections will have general information relevant to essentially all semiconductors, as well as specific cases

related to TCOs. This should give much of the background information and funda-
mental physics required to describe these materials. Hopefully, it will become clear
how all of the properties discussed below are inextricably linked, and are essential in
building an understanding of these materials.

1.2.1 OPTICAL PROPERTIES

The optical and electronic properties of materials are intimately linked. The important
parameters determined experimentally for TCOs are the transmission and reflection
(and therefore the absorption) of light in the ultraviolet (10 nm < λ < 400 nm, 124 eV >
E > 3.3 eV), visible (400 nm < λ < 740 nm, 3.3 eV > E > 1.7 eV), and infra-red (700 nm <
λ < 10^6 nm, 1.7 eV > E > 10^{-3} eV) spectral regions. Figure 1.3 displays a representation
of these properties for a typical TCO material.
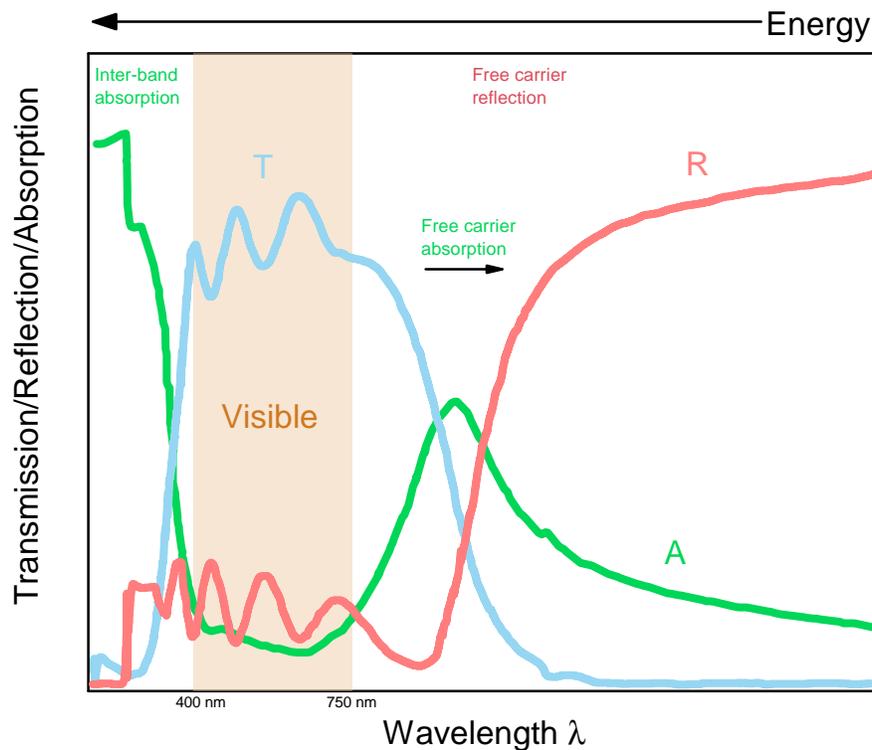
Figure 1.3: Representation of the transmission (T), reflection (R) and absorption (A) spectra of a typical TCO. Figure adapted from Pasquarelli *et al.*[24]

In the three spectra in figure 1.3, there are three characteristic regions of interest for a TCO. Firstly, a TCO is required to be transparent over the visible spectrum (highlighted as an shaded box). Typically a transmission of around 80% over a wavelength range of at least 400-750 nm is expected. This is achieved through the material having a large band gap, low plasma energy allowing for low reflection and absorption over this region. Secondly, at low wavelengths (high energy) light is absorbed causing inter-band transitions. This is shown in figure 1.3 where a dramatic increase in absorption can be seen and the transmission drops off heavily. As previously suggested, typical TCOs have a band gap greater than 3 eV and so the absorption onset occurs at wavelengths < 410 nm (>3 eV). Finally, at higher wavelength typically around > 1550 nm (lower energy < 0.8 eV) another decrease in optical transmission is seen. This is coupled with a large increase in the reflection signal.

The increase in reflection is related to the free carrier response to the electric field of the incoming light in the degenerately doped TCO, where the free carriers are treated as a free electron gas or plasma whose oscillatory motion with respect to the positive ions in the fixed lattice is naturally counteracted by Coulombic forces. The charge density oscillates at a natural frequency called the plasma frequency (ω_p). At frequencies of light higher than the plasma frequency of the material ($\omega_p < \omega$) the electric field oscillates too quickly for the electrons to respond and so light can propagate through the material and the material is transparent. If the frequency of light matches the plasma frequency ($\omega_p = \omega$) the electron gas oscillation couples strongly to the excitation, oscillating in phase with the electric field component of the electromagnetic wave. The imaginary component of the dielectric function has a maximum and the light is absorbed by the free carriers and the energy is dissipated as heat due to strong damping. Finally, if the frequency of light is much below the plasma frequency ($\omega_p > \omega$) the electron gas easily follows the oscillations of the excitation. Because the oscillation

amplitude is small, energetic losses due to damping are essentially negligible. The electron gas is completely out of phase with the light's electric field which cannot penetrate into the material and so the light is reflected.

INTERBAND TRANSITIONS

Figure 1.4 shows some of the interband transition mechanisms that are displayed in most common semiconductors. For a non-degenerate direct band gap semiconductor, an interband excitation can occur between the two band extrema at wavevectors equal to zero ($k = 0$). This is true also for an indirect band gap semiconductor where the band extrema are offset in reciprocal space, but additional energy (e.g. a phonon) is required to allow for the promotion of an electron to the conduction band. Under heavy doping and in the case of degeneracy (as is the case for transparent conductors) some of the states above the conduction band minimum (CBM) are already occupied. The lowest energy allowed transition is therefore away from the zone centre, and so do not occur between the band extrema (Burstein-Moss band filling). In this case the fundamental band gap E_g is no longer the same as the optical band gap E_g^{opt}.
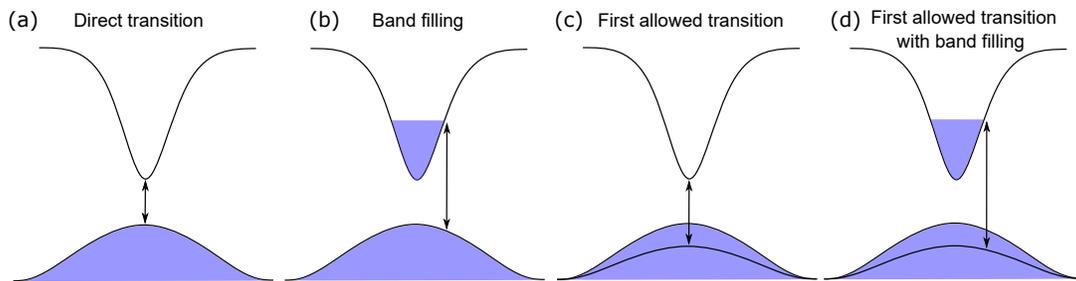
Figure 1.4: Schematic representations of different interband transition processes. (a) a direct transition from the valence band maximum (VBM) to conduction band minimum (CBM), here E_g is equivalent to E_G^{opt}. (b) the effect of band filling (Burstein-Moss shift) where the states at the CBM are occupied and so the transition is forced away from the zone centre. (c) the fist allowed transition applies to semiconductors with symmetry forbidden transitions between the VBM to CBM states. (d) the first allowed transition in a degenerate semiconductor displaying band filling with symmetry forbidden transitions.

Finally, materials displaying certain crystallographic symmetry often have what

is called a 'forbidden' electric dipole transition between the band extrema. More correctly these transitions can be described as forbidden by selection rules. An important example of a TCO structure exhibiting this sort of transition is In_2O_3's body-centred cubic (bcc) bixbyite structures[25]. In In_2O_3, the VBM consisting of mixed O 2p and In 4d orbitals (T_g symmetry) does not have an allowed transition to the CBM which is primarily formed from a mixture of In 5s and O 2s states (A_g symmetry). This is because these two states share the same parity, and electric dipole transitions are only allowed between states of opposing parity for these materials. Photon absorption through this transition is extremely weak, and very hard to see experimentally. At $\sim$0.8 eV below the VBM at the Γ-point the wave function character becomes more p like (T_u symmetry), where strong optical transitions begin to occur[25,26]. This phenomena is discussed more in Appendix A.

BURSTEIN-MOSS BAND FILLING AND BAND GAP RENORMALIZATION

As discussed above the high density of n-type dopant incorporation in TCOs usually leads to a high degree of degeneracy, where the Fermi level is pushed above the CBM and so normally unoccupied states become filled. This results in a widening of the optical band gap, the effect is known as the Burstein-Moss shift[27,28]. The Mott criterion (the critical carrier density for the metal-insulator transition[29]) can be estimated by $a_d n_c^{1/3} \sim 0.25$ where the donor states effective radius a_d is related to the Bohr radius, $a_d = a_B \frac{\varepsilon(0)}{m^*/m_0}$, where $\varepsilon(0)$ is the static dielectric constant ($\varepsilon(0) = 8.9$ for In_2O_3[5]). If the effective mass of In_2O_3 is taken to be $m^* = 0.22 m_0$[30,31] then the Mott criterion is $n_c \sim 1.6 \times 10^{18} cm^{-3}$, much lower than the typical free carrier densities of doped In_2O_3 ranging from $10^{20} cm^{-3}$ to over $10^{21} cm^{-3}$.

In modelling the Burstein-Moss band filling, it is assumed that a direct band-to-band transition takes place (phonon contributions are ignored), the lowest unoccupied
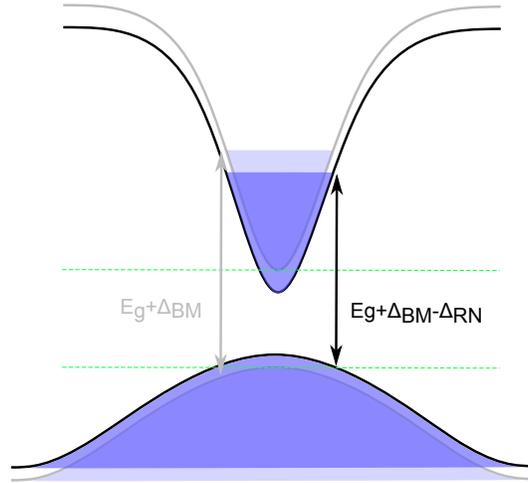
Figure 1.5: schematic representation of the CBM and VBM in a heavily doped semiconductor, where the effects of Burstein-Moss band filling Δ_{BM} and band gap renomalisation Δ_{RN} are included.

state is situated at E_F (temperature effects are ignored), and that the valence band is non-dispersive (a valid assumption in oxide semiconductors due to the non-dispersive nature of the O 2p states at the VBM). The magnitude of the Burstein-Moss shift under free electron theory in a parabolic band approximation is described by

$$\Delta_{BM} = \frac{\hbar^2}{2\,m_{vc}^*} \left(3\pi^2 n\right)^{2/3} \tag{1.1}$$

where n is the free electron concentration, and the reduced effective mass is given as $\frac{1}{m_{vc}^*} = \frac{1}{m_v^*} + \frac{1}{m_c^*}$, taking the VBM and CBM effective masses into account. The observed carrier dependence of the band gap is often much smaller than that expected from equation 1.1[32–34]. A band gap renormalization term was subsequently introduced to account for this discrepancy[33,35]. The magnitude of this renormalization Δ_{RN} is related to many body effects: electron exchange, the attractive Coulomb interactions and impurity scattering (i.e. interactions between electrons with themselves and with the dopant ions redistribute states close to the band edges, pushing the VBM upwards and the CBM down). This effect can be seen demonstrated schematically in figure 1.5. Prac-

tically, many people have adopted the methodology of Berggren and Sernelius[33,35-37] which was originally introduced to account for band gap renormalization in degenerately n-type doped Si and Ge:

$$\Delta_{RN} = \hbar\Sigma_v - \hbar\Sigma_c.$$

(1.2)

The shift in the band gap relative to the band edges is given by equation 1.2, where $\Sigma_{v,c}$ represents the self-energies associated with the many body interactions in the valence and conduction bands. These can be further split into contributions from electron-electron and electron-ion interactions in the valence and conduction band. A useful limiting approximation for these values is that at zero frequency (again ignoring thermal effects and phonons), sometimes referred to as the static approximation. It should be noted that although the concept of band gap renormalization is widely used in the literature, the origin of the phenomena has not been well explored and is not well understood. This is in part due to the difficulty in deconvolving the weight of the contributions Δ_{BM} and Δ_{RN} to the measured absorption onset[38]. Many attempts at providing a quantitative description of the renormalization term have been presented[32,35,36,39-41] but none have provided the necessary generality to be fully adopted.

ALTERNATIVE BAND GAP MODELLING APPROACHES

While closed-form equations such as those mentioned above are desirable to describe and model band filling and renormalization effects due to their ease of implementation, often *ab initio* approaches are sought to capture more of the physics related to these phenomena. These are most often performed through density functional theory (DFT) approaches, which are discussed in more detail in section 3.2. Here, a comparison between the results from such approaches are compared to the closed-form models, highlighting the usefulness and customizability of the DFT method for this style of

problem.

Recently, Walsh *et al.*[38] carried out density functional calculations using the PBE functional on undoped and doped In_2O_3 systems. They perform a correction to the ground-state effective mass to account for exaggerated coupling between the valence and conduction bands at the zone centre. It was found that conduction band non-parabolicity had a large impact on the extent that renormalization affects the system at high carrier densities. Furthermore, they show doping can further enhance the amount of renormalization due to band hybridization.
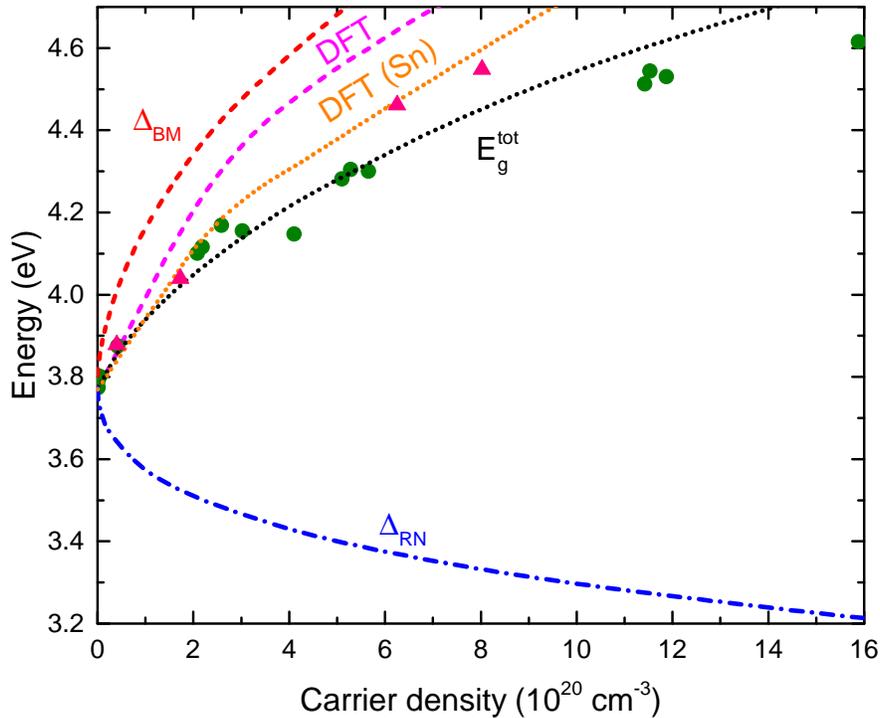
Figure 1.6: The effects of Burstein-Moss shift and band gap renormalization on E_g^{opt} for doped In_2O_3. This figure has been adapted from that of Feneberg *et al.*[42] and Walsh *et al.*[38]. Data was taken from Feneberg *et al.* via spectroscopic ellipsometry measurements (green circles)[42] and Hamberg *et al.* through absorption measurements[33].

Figure 1.6 highlights two different approaches to modelling the evolution of the optical gap with carrier density. Feneberg *et al.* adopted the analytical expressions of Berggren and Sernelius[35] described above, accounting for conduction band non-parabolicity[43,44] in their equation for the Δ_{BM} (red curve) and Δ_{RN} (blue curve) to pro-

duce the total optical gap $E_g^{opt} = E_g + \Delta_{BM} - \Delta_{RN}$ (black curve). This approach is in good agreement with their data at low carrier densities, only beginning to overestimate the extent of the Bursteen-Moss shift at very high densities $n > 10^{21}$ cm^{-3}. Walsh *et al.* who predicted the band structure change through DFT calculations also accounts for conduction band non-parabolicity in a similar way[45]. Only accounting for non-parabolicity in their calculations produced the pink curve labelled DFT in figure 1.6. Whilst the shape of the curve reflects that of the one by Feneberg *et al.*, the calculated gaps are overestimated. This difference may be due to the free parameter C used by Fenerberg *et al.* in his approximation for the conduction band non-parabolicty allowing for a larger Δ_{RN} term ($\sim$50% greater), or possibly due to the simplistic approach used to replicate doping in the DFT calculation underestimating the effect on the bands. More interestingly, the deviation taken by Walsh *et al.* from the standard approach comes at high doping densities, where they attempt to account for the dependence of the dopant chosen, something that the Berggren and Sernelius model neglects. When including Sn as a dopant in their calculations (swapping an In atom for a Sn atom in their unit cell), they show a great reduction in their calculated optical gaps (orange curve). This is in reasonable agreement with Sn-doped experimental absorption data[33]. Ignoring the fact that the two experimental data sets do not agree (a number of experimental uncertainties could possibly account for this, or this could stem from the different ways the optical gaps are extracted form the raw data), both fits seem in reasonable agreement at lower doping densities (n<3 $\times$ 10^{20}cm^{-3}). Walsh *et al.* also displays a model taking Ge as the dopant and shows the Δ_{RN} term is much greater for Ge-doping which they attribute to dopant s-levels hybridizing with the CBM, greatly affecting the band curvature at the CBM (Ge having more of a broadening effect than Sn). This dopant-dependent approach to optical modelling allows for a more complete model, and should allow for a higher degree of accuracy in the model

assuming the effects of each dopant can be accurately quantified.

Other than the band gap, the absorption and reflection of light by the free carriers in a material places constraints on the limit to both the transparency and conductivity (or carrier density) of doped systems. It is therefore important to have some appreciation of the physics of free carriers in heavily doped semiconductors. Fortunately, it is often possible to use simple oscillator models to describe the effects of free carriers in these systems. This section follows on from the content in Appendix B, describing the reflectivity of a free electron gas. The model used for a free electron gas can be modified slightly to describe the reflectivity of a doped semiconductor to a fairly high degree of accuracy. This is done by accounting for the fact that the free electrons are moving in the conduction band which means each has an associated effective mass m^* instead of the free electron mass, and there is a contribution to the polarizability from bound electrons.

Taking this into account, the frequency dependent dielectric function can be written as

$$\varepsilon(\omega) = \varepsilon(\infty) - \frac{ne^2}{m^*\varepsilon_0}\frac{1}{\omega^2 + i\gamma_p\omega} \qquad (1.3)$$

where n is the density of free electrons, ω the frequency of incoming light, γ_p the damping coefficient of the plasmons, and $\varepsilon(\infty)$ is the dielectric constant at the high frequency limit. This equation is often written as

$$\varepsilon(\omega) = \varepsilon(\infty)\left(1 - \frac{\omega_p^2}{\left(\omega^2 + i\gamma_p\omega\right)}\right) \qquad (1.4)$$

where the plasma frequency is given as

$$\omega_p^2 = \frac{ne^2}{\varepsilon(\infty)\varepsilon_0 m^*}. \tag{1.5}$$
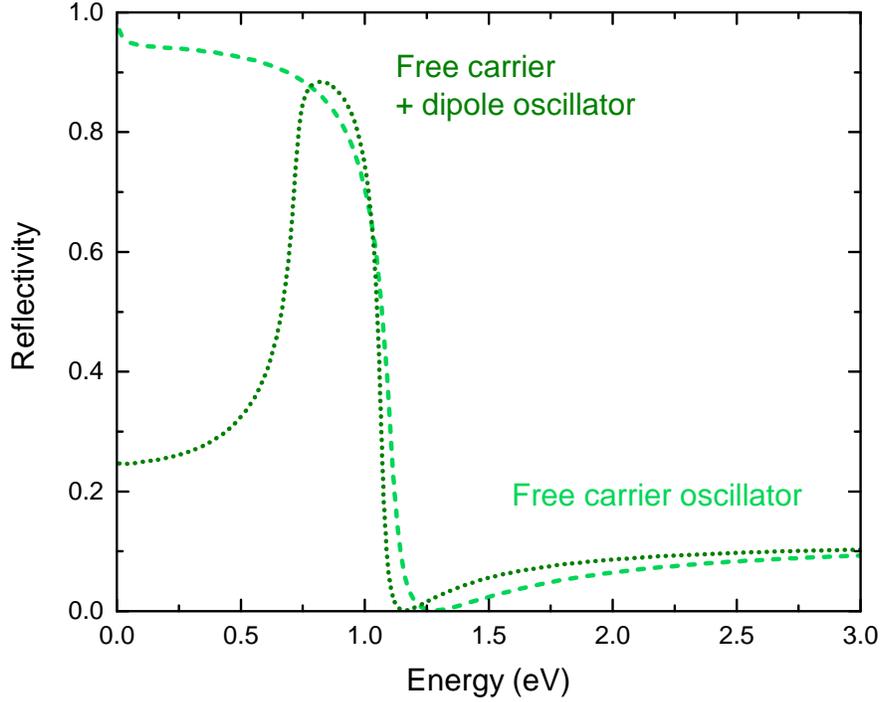
Figure 1.7: Reflectivity of free carriers in a semiconductor (one oscillator model) and free carriers and lattice oscillations (two oscillator model).

The relative dielectric function given in equation 1.3 pertains to free carriers only (ignoring the effects of the lattice vibrations i.e. phonons), but can be easily adjusted to account for phonons (by simply summing the lattice and free carrier contributions[46]). This can be described by the following equation

$$\varepsilon(\omega) = \varepsilon(\infty) + \frac{(\varepsilon(0) - \varepsilon(\infty))\omega_{ph}^2}{\omega_{ph}^2 - \omega^2 - i\gamma_{ph}\omega} - \frac{\varepsilon(\infty)\omega_p^2}{\omega(\omega + i\gamma_p)} \tag{1.6}$$

which accounts for the oscillations of phonons with resonant frequency ω_{ph} and damping γ_{ph}, as well as free carriers. In general it seems reasonable to account for two oscillators when modelling the reflection edge of highly doped materials or metals, as

reflection onset is usually in quite a narrow energy region. However, real materials will have many oscillators which require models with multiple oscillators. Using the formalism set out, it is intuitive to create a multi-oscillator model such as:

$$\varepsilon(\omega) = \varepsilon(\infty) + \sum_j \frac{f_j}{\omega_j^2 - \omega^2 - i\omega\gamma_j} - \frac{\varepsilon(\infty)\omega_p^2}{\omega^2 + i\omega\gamma_p} \tag{1.7}$$

where f_j is the dipole strength ($\sum_j f_j = 1$). Because the main concern is with ω_p for TCOs, and not the lattice effects etc. there is no need to go beyond equation 1.6 in this thesis.

Electron oscillations may also occur at the surface of a material. Using Maxwell's equations, it can be shown that the ratio between the surface plasmon frequency and the bulk plasmon frequency can be given as[26]

$$\frac{\omega_{sp}}{\omega_p} = \left[\frac{\varepsilon(\infty)}{\varepsilon(\infty) + 1} \right]^{\frac{1}{2}}. \tag{1.8}$$

Typically for metals $\omega(\infty) = 1$, so this ratio is $\omega_{sp} = \frac{\omega_p}{\sqrt{2}}$ but is often larger for TCO materials as $\varepsilon(\infty)$ is larger.

1.2.2 EFFECTIVE MASS

One of the most fundamental properties relating not only to the structure, but both the electronic transport and optical properties of a material is the effective mass of the electrons and holes. Here the focus is on the electron effective mass as this is most relevant to n-type TCOs. When talking about the effective mass of an electron in a crystal system it is important to be fully aware of which effective mass is being talked about, and which effective mass is important to which physical situation.

When an electron is in the periodic potential of a crystal, it is accelerated relative to the lattice by an external perturbation e.g. applied electric or magnetic field, tem-

perature or stress. The properties of that semiconductor are very often determined by the response of this electron to the external stimulus. The acceleration follows as if the electron has a mass equal to the effective mass m^*. It can be shown using Newton's second law (i.e. the acceleration of an object relates to its mass and the net force acting upon it) and acknowledging that an electron can be treated as a wave and so a group velocity can be determined, that the effective mass (which is both energy and time dependent in the most general case) is given by[47]

$$\frac{1}{m^*} = \frac{1}{\hbar^2}\frac{d^2E}{dk^2} = \frac{1}{\hbar^2 k}\frac{dE}{dk}. \tag{1.9}$$

where E is the energy and k the wavevector. Often a parabolic dispersion equation ($E(k) = \frac{\hbar^2 k^2}{2m^*}$) is sufficient to approximate the curvature (and so the effective mass) at the band edge for many semiconductors. This can be determined assuming an energy independent effective mass. Of course, this is a special case of band dispersion which does not truly apply to any material, no band dispersion is ever truly parabolic. However, parabolic band dispersion can be a powerful approximation, and is extremely effective at predicting m^* very close to the Γ-point for many materials.
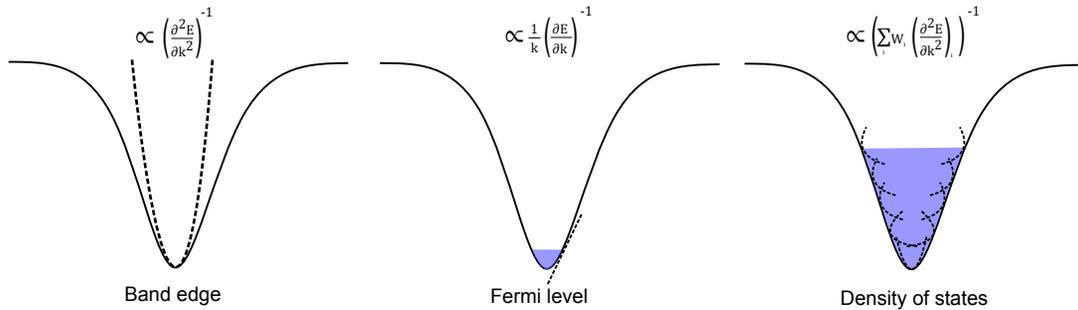
Figure 1.8: Effective mass approximations replicated from Whalley *et al.*[48]: left- the effective mass is inversely proportional to the the curvature of the electronic dispersion for a parabolic band; middle- effective mass is inversely proportional to the gradient of the electronic dispersions; right- Density of states averaged effective mass takes into account the average curvature.

Figure 1.8 shows how the different approximations give different masses. Equation

1.9 is often used to approximate the band edge in relatively non-parabolic band structured materials. To be more generally applicable, non-parabolicity away from Γ must be accounted for in the expression for $E(k)$. In this case it is beneficial to consider the energy and momentum dependence of the electron to give the curvature at a certain Fermi energy:

$$\frac{1}{m_F^*} = \frac{1}{\hbar^2 k_F} \left| \frac{dE}{dk} \right|_{k=k_f}.$$

(1.10)

Here the wavevector is equal to the Fermi wavevector k_F. This is very important for degenerate materials such as TCOs but depends heavily on the choice of approximation for $E(k)$.

Finally, there is one more variant of the effective mass that will be discussed here. This is referred to as the density-of-states-averaged effective mass, or occasionally the optical mass. This is important for materials displaying metallic behaviour, i.e. materials with free carriers. When exciting the collective free carriers, in a reflectance measurement for example, the effective mass measured is a summation over all occupied states and so the relevant effective mass is that averaged over the occupied states.[49,50] Here, a simple carrier statistics model is chosen in the treatment of the average effective mass given as

$$\langle m^*(E) \rangle = \frac{\int m^*(E) g(E) f(E, E_F) dE}{n(E)} = \frac{\int m^*(E) g(E) f(E, E_F) dE}{\int g(E) f(E, E_F) dE}$$

(1.11)

where $f(E)$ is the Fermi-Dirac distribution function describing the distribution of fermions in a system:

$$f(E, E_F) = \frac{1}{exp(\frac{E-E_F}{k_B T}) + 1}.$$

(1.12)

Equation 1.11 shows the density-of-states (DOS) averaged effective mass, where $g(E)$ is

the density-of-states above the CBM, $f(E, E_f)$ is the Fermi-Dirac distribution function, and $n(E)$ is the energy dependent carrier density[34]. This is essentially the average effective mass per electron, which within appropriate limits ($k(0 \rightarrow \infty)$) can be limited to all occupied states above the CBM to E_F.
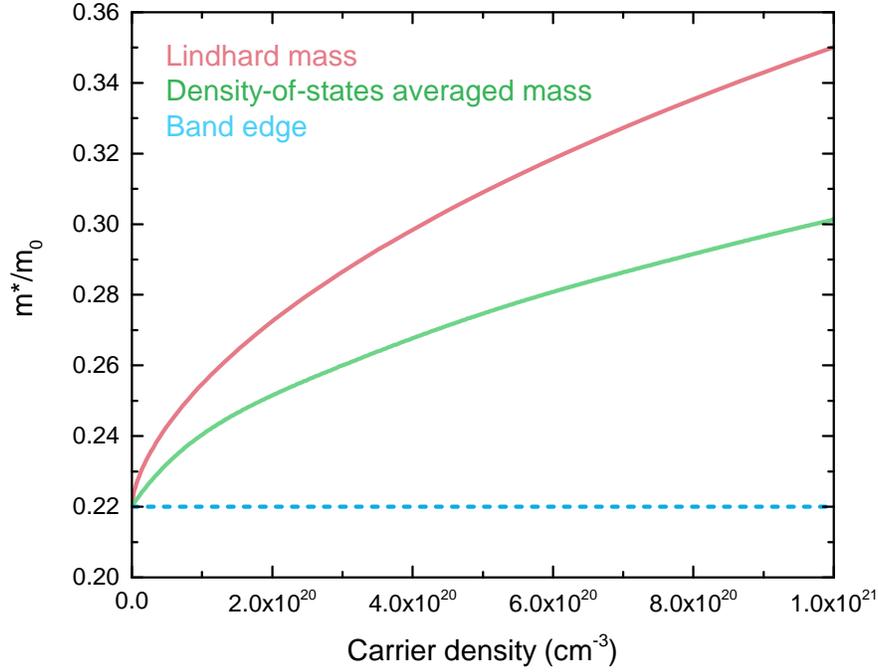
Figure 1.9: Comparison between band edge effective mass ($m^* = 0.22 m_0$), density-of-states averaged effective mass found using equation 1.11 and Lindard optical mass determined using equation 1.10.

It is worth noting that there exists some ambiguity over the correct formalism for the optical effective mass[51]. Another definition of the optical mass is given by Wooten[50], as the wavevector dependent effective mass (equation 1.10) averaged over all occupied states

$$\frac{1}{\langle m^*(E) \rangle} = \frac{2}{n} \sum_{l} \sum_{k}^{occ.} \frac{1}{m_l^*(k)} \quad (1.13)$$

where l represents a given band with mass m_l^*. In the case of only one occupied branch in an isotropic material at T=0 K it has been shown that this simplifies to equa-

tion 1.10[52], in agreement with the results of Inaoka et al.[53] who used the Lindhard dielectric function to arrive at the same optical mass formula. A comparison between these expressions is shown in figure 1.9. Since the mass given by equation 1.11 is averaged across the band structure, we expect this mass to become progressively lower than that given by equation 1.11. Notice the similarity between the two curves, especially at low carrier density. Hence, it should not make a difference which formalism is chosen so long as the carrier density is sufficiently small. Furthermore, in this thesis equation 1.11 is used to describe the optical effective mass, which may give slightly different results to that obtained using equation 1.10. However, this will prove inconsequential in this thesis as the analysis is only performed for the comparison of two material systems, in which case it is highlighting a difference and not commenting on the accuracy of the mass value obtained.

1.2.3 ELECTRONIC STRUCTURE

Knowledge of the electronic structure of a material is extremely important, providing information on both the optical and transport properties of a material. A powerful tool for probing the electronic structure is through first principles calculation methods, such as density functional theory (DFT), which is discussed in section 3.2. For now it is instructive to look at what the electronic structure of a TCO can tell us.

Taking In_2O_3 as our example again, figure 1.10 shows the unit cell of In_2O_3 (space group Ia$\bar{3}$, No. 206), which is a body centred cubic (bcc) bixbyite structure. The Brillouin zone of this structure is also seen in figure 1.10. The high symmetry points are superimposed onto the Brillouin zone, connected by the paths of high symmetry.

The electronic structure from DFT calculations associated with this structure is shown in figure 1.11. This figure shows the electronic density of states (a), calculated band structure (b), a simplified molecular orbital diagram (c), and finally the calculated
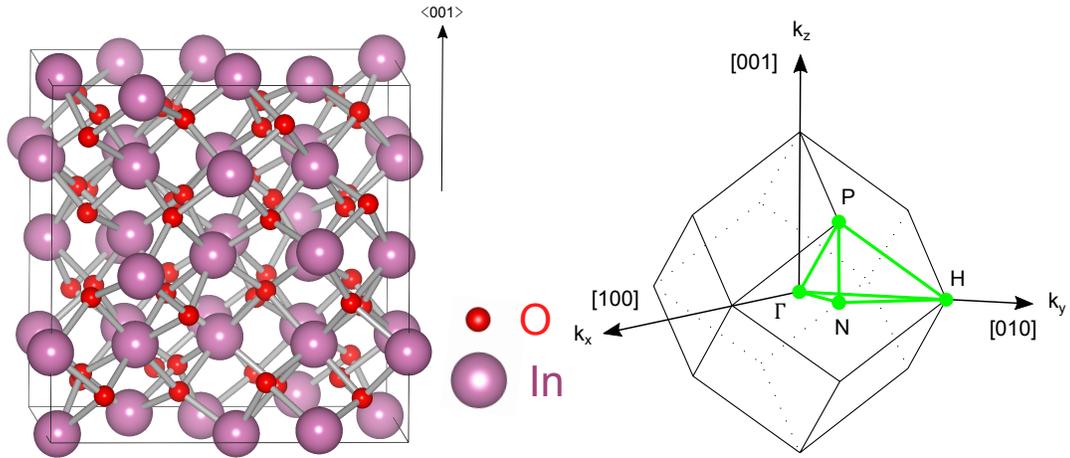
Figure 1.10: Left: Unit cell of the bixbyite crystal structure of In_2O_3. Right: Brillouin zone associated with the bcc lattice.

thermodynamic transition levels (in this thesis, thermodynamic transition levels and defect formation energies are used interchangeably as they amount to the same thing) associated with important defects in In_2O_3 (d).

The DOS allows for a visualisation of the orbital contributions to the band structure. This is shown for In_2O_3 in figure 1.11 (a) around the VBM (set to 0 eV here), with the conduction band at positive energy. Orbital contributions from each atom are displayed in colour and are usually referred to as the partial electronic DOS (PDOS). These orbital contributions can be further split into the different directional components (e.g. p_x, p_y, p_z for an orbital), or it is also common to see the PDOS for each orbital type grouped, i.e. In s and O s components summed to give one s-orbital component. The summation of all the PDOS gives the total density of states or TDOS, represented by the black line. Calculated DOS can be directly compared to photoemssion data after cross section-correction and broadening. This gives a good method to verify the accuracy of calculations and help interpret experimental data.

The DOS is closely linked to the band structure of a material, which is shown for In_2O_3 in figure 1.11(b). Whilst DOS is concerned with orbital contributions at par-
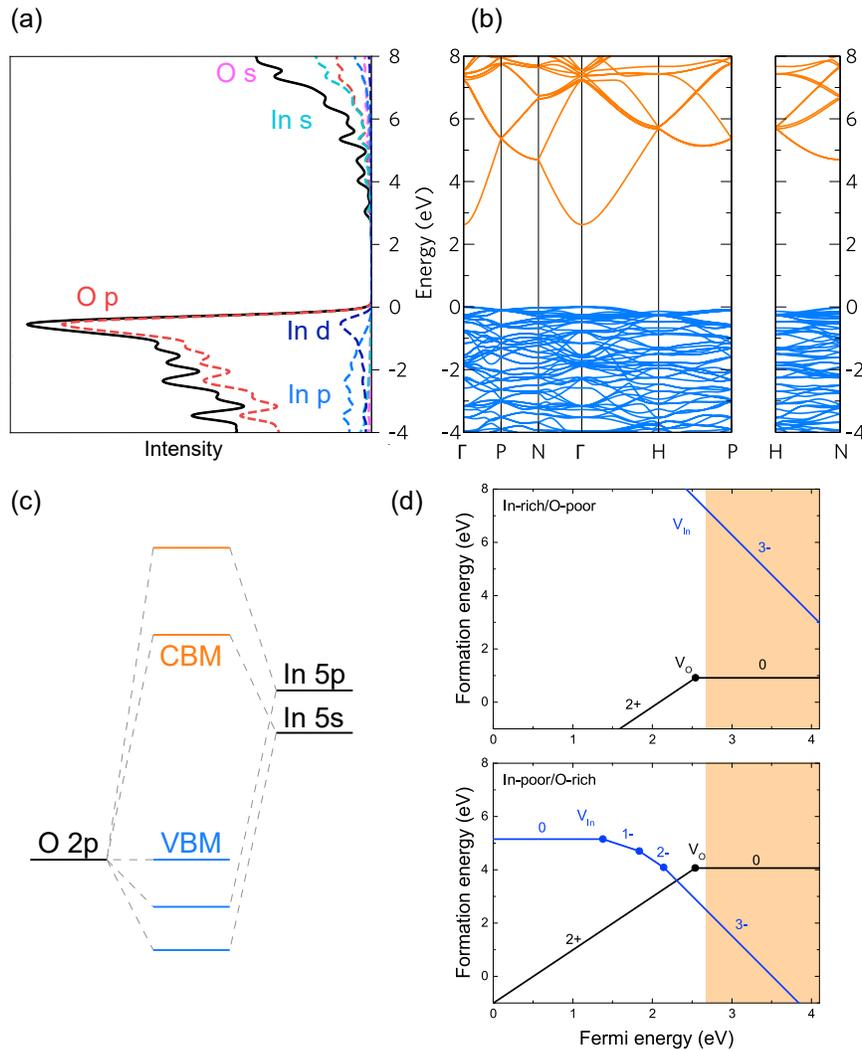
Figure 1.11: (a) Calculated density of electronic states in In_2O_3. (b) Calculated band structure of In_2O_3. (c) Simplified molecular orbital diagram showing the prominent In s character with small O p contribution of the CBM and large contribution from the non-bonding O p dominated VBM. (d) Calculated defect formation energies as a function of Fermi level position for In-rich/O-poor and In-poor/O-rich growth conditions for In_2O_3. The onset of the orange box indicates the position of the conduction band minima at 2.7 eV, with the conduction bands extended to higher energies.

ticular energies, the band structure displays the distribution and evolution of energy bands over wavevector k. In the ideal case the band structure would span the whole of k-space, but this is extremely computationally expensive and so it is conventional to plot the band structures along the high symmetry directions within the Brillouin zone. High degrees of energy band degeneracy are displayed at high symmetry points

and along high symmetry axis which is a consequence the translational symmetry of the Bloch wave function, hence these high symmetry lines are extremely useful in providing a less expensive representation of the whole Brillouin zone.[54] In figure 1.11, the conduction band states are coloured orange while the valence band states are blue. The energy axis is aligned with that of the DOS in (a) so it is clear that the relationship between the two is very close. In regions with highly dispersive bands, such as at the CBM, a low intensity in the DOS is expected, and for low dispersion (such as the VBM) higher DOS intensity is expected. The intensity of the DOS is inversely proportional to the band dispersion. Some features of the band structure that are characteristic and extremely important in TCO materials are: a large band gap (2.7 eV in the case of In_2O_3) which allows the material to be transparent, a flat VBM which gives a high hole effective mass, and a highly dispersed conduction band derived from In 5s and O 2p states giving a low CBM effective mass. This configuration is shown schematically in figure 1.11c as a molecular orbital diagram. This helps explain why TCOs are excellent n-type conductors with highly mobile electrons.

Finally, TCOs are inherently n-type and are highly n-type dopable which is another reason they display excellent electronic properties. This property is heavily influenced by the intrinsic and extrinsic defects in the material. Figure 1.11(d) shows the thermodynamic transition levels of the most important (lowest formation energy) defects in In_2O_3 for two extreme material growth conditions, i.e. In rich/O poor and In poor/O rich. The realistic scenario will be somewhere in between. The CBM in these diagrams are represented by the onset of an orange box, with the VBM at 0 eV. Defects have a certain charge state which is represented by the slope of the line (donors have positive slopes, neutral defects have zero gradient, and acceptors have negative slopes.) In both cases in figure 1.11, the oxygen vacancy V_O is a donor until the Fermi level is in the conduction band, where it becomes neutral. In the In-poor case, the indium

vacancy V_{In} which acts as an acceptor, crosses the V_O line just below the CBM, where it then begins to compensate the donor behaviours of the V_O. If this point is close enough to the CBM ($k_B T \sim 0.03$ eV) then electrons can be excited into the conduction band at room temperature. This may not be the case in the In-poor regime, however the realistic scenario is somewhere between the two regimes.

1.2.4 CONDUCTIVITY IN TCOs

INTRINSIC DEFECTS

Conductivity in TCO materials can be caused by intrinsic point defects, or most commonly for commercial application, by doping i.e. replacing host atoms with other elements to modify the properties of the material, in this case to gain extra electrons. A purely stoichiometric intrinsic material is expected to be insulating but due to high non-stoichiometry and a great degree of size mismatch between anion and cation in metal oxides (often associated with low defect formation energies), the native defects cause n-type conductivity. Conventionally it was thought that oxygen vacancies were the primary source of conductivity in undoped TCOs[55-57], which were often based upon indirect experimental observations of varying conductivity due to different oxygen partial pressure. Hence, this claim has been disputed over the years and is still a contentious issue today. Particle irradiation experiments have been previously performed showing native defects can certainly increase the number of free carriers in a oxide material[58-61]. However the nature of the native defect is very difficult to determine in these experiments and in reality a number of different defects likely contribute to increased conductivity giving inconclusive results to whether V_O are responsible for conductivity in TCOs. With more advanced and accurate DFT calculation now available, large strides towards better understanding the conductivity in TCOs have been made. It has been shown that the donor ([0/2+]) transition level is typically deep (well

below the CBM) in many oxide semiconductors such as In_2O_3, SnO_2 and ZnO, and so unlikely to be the sole source of high conductivity in TCO materials[62-64]. V_O may still contribute to conductivity through the persistent photoconductivity model proposed by Lany and Zunger[65,66] whereby the excited V_0^* state has a metastable conductive state. They suggest an energy barrier prevents the immediate relaxation of excited electrons from the resonant metastable state to the VB, giving enough time for the store of conduction electrons to be constantly replenished. However, it is not clear whether this energy barrier is in fact large enough to achieve persistent photoconductivity in reality[67]. Similarly, cation interstitials are too high in formation energy to give the high carrier concentrations observed[65,68].

Hydrogen is very difficult to remove from growth environments and so H defects are present in materials from growth. H defects are particularly difficult to experimentally observe due to the small size of the H atom amongst other things. However, it has been shown that H interstitials acts as either a donor or acceptor in the 1+ or 1- charge state, the neutral charge state not forming due to H^0 always being higher in energy[70]. The transition level between H in the 1+ and 1- charge states (H(+/-)) for many conventional semiconductors (e.g. GaN, GaAs, etc.) lies somewhere near the middle of the gap, meaning hydrogen will act as a compensating donor (acceptor) when E_F sits closer to the VBM (CBM). H(+/-) sits directly between the H(0/-) and H(+/0) transitions as shown in figure 1.12. Therefore, in all conventional semiconductors H counteracts the conductivity at a given E_F. Conversely, TCOs display a H(+/-) transition near to or even above the CBM, meaning the donor H^+ state dominates up to this point. This suggests H is a source for conductivity in TCO materials[69-71]. Indeed, this is supported by increased conductivity due to H diffusion in ZnO[72,73], SnO_2 and In_2O_3[69]. Subsequently, H is often used as a dopant or co-dopant in TCOs to achieve enhanced electrical properties[74]. It should be noted that H defects can form on substi-
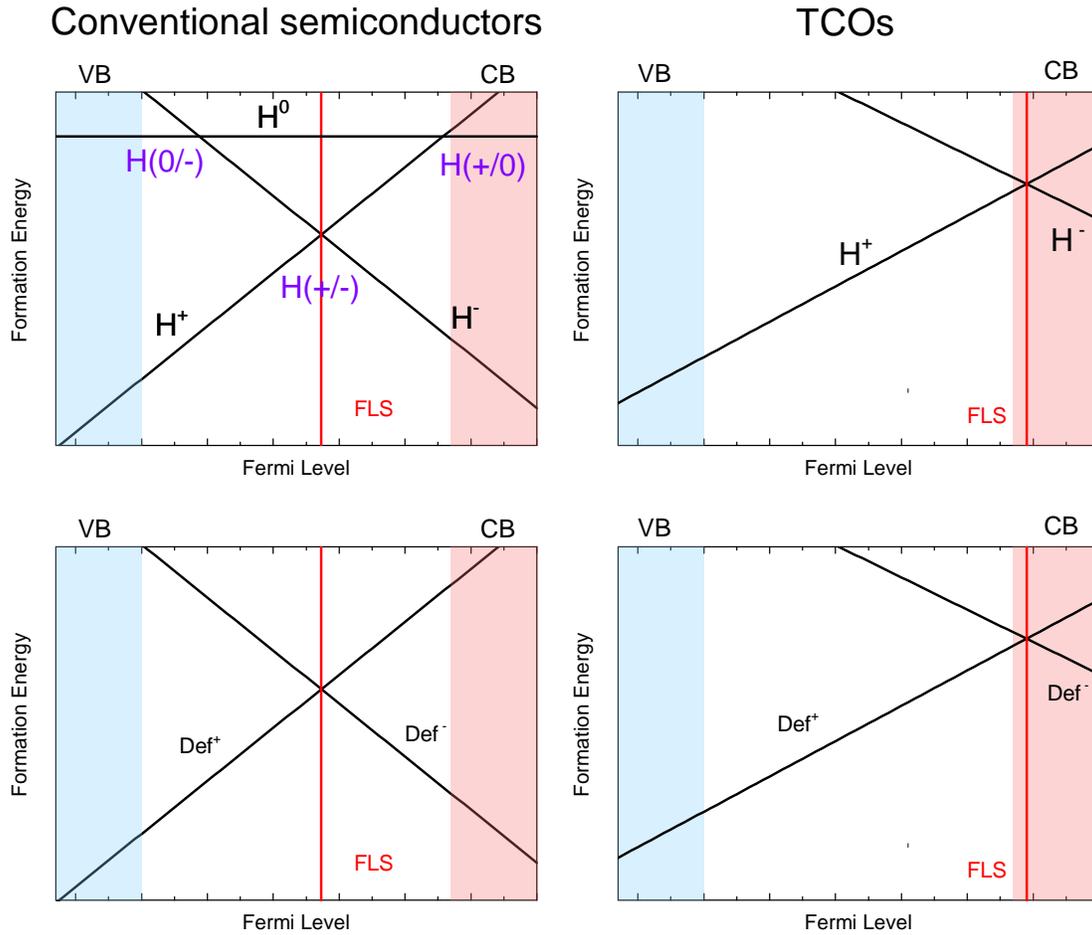
Figure 1.12: Schematic for the formation energies for the interstitial H and the dominant native defects in conventional semiconductors (left) and TCOs (right). Figure adapted from King *et al.* [69]

tutional or interstitial sites, may sit on multiple sites in the lattice and can also form multiple defect complexes or multi-centre bonds, with some defect environments or bonding configurations being far more thermally stable than others[75–77]. It is also worth pointing out that in reality, whilst evidence suggests H defects are very important for determining the conductivity in materials, it cannot be ruled out that other dominant defects as the source. A more intuitive picture may be that many defects actually contribute to conductivity, and the extent of their contribution is measured by their relative energy of formation.

At The Surface

Above, point defects in the bulk of materials are considered, but the behaviour of carriers at the surface of a material needs also be considered. This can be understood by looking at the position of the charge neutrality level (CNL), branch point energy, or Fermi level stabilisation (FLS) point in materials[78,79] (for a more detailed discussion about the history and significance of the CNL, see appendix F). At the material surface, the electronic state exponentially decays into the vacuum due to the breaking of translational symmetry at the terminated surface. This is in analogy to broken symmetry at a defect in an otherwise perfect lattice[80,81]. Because these states decay into the vacuum, they have real energies but display complex wavevectors. In bulk materials, complex wavevectors are not meaningful as their Bloch wavefunctions grow exponentially as $z \to \infty$. Because these surface states can have imaginary wavevectors, they can exist in the band gap of a material, hence they are called virtual gap states (ViGS). ViGS can occur wherever periodicity is broken in the lattice, such as at defects, surfaces and interfaces. The ViGS have donor (acceptor) character closer to the valence band (conduction band) and the charge neutrality level occurs at the energy where equal donor and acceptor character occur. Free carriers actually rearrange themselves in the vicinity of the surface in order to minimise energy and screen the surface charge, leading to the well known space-charge regions where the carrier density varies dramatically compared to the bulk. For TCO materials, it has been found that the common surface charge state displays electron accumulation, and so downward surface band bending is seen. The Fermi level sits higher relative to the band-edges at the surface compared to the bulk. The net negative charge from the accumulated carriers at the surface balances the net positive surface charge pinning the Fermi level and maintaining charge neutrality. This has been shown to be the case for In_2O_3[82] and CdO[59] and is shown schematically in figure 1.12. Since the FLS is above

the CBM, general material properties may be expected such as unintentional n-type conductivity, high n-type dopability and difficulty in achieving p-type conductivity.
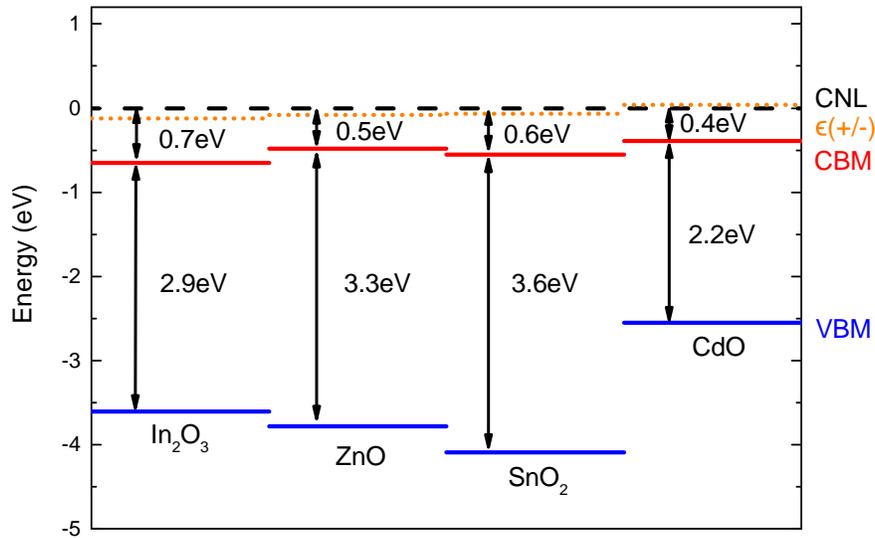
Figure 1.13: Band lineup of important TCO materials with respect to the CNL. The H(+/-) transition levels are shown in orange. Positions of the CNL relative to the band edges are taken from various sources.[59,82–84]

The CNL concept can be related back to H defects under the ViGS framework. Van de Walle *et al.*[70] suggested H creates a positive charge state when creating a cation dangling bond (i.e. H^+ bonding to the anion), or a negative charge state when creating an anion dangling bond (H^- bonding to the cation). This H(+/-) is material specific as it involves bond breaking, and the transition level between cation and anion dangling bond formation (given as the Fermi level position where positive and negative charge states have equal energy $\epsilon(+/-)$[70]) may be expected to occur at the CNL[59]. This is shown schematically in figure 1.12 where the top and bottom panels show the H transition level and native defect levels respectively. This is generally quite a good estimate of the CNL assuming the correct H bonding environment is selected in the calculation. Figure 1.13 show band line ups for important TCO materials, displaying the determined CNL and the calculated H(+/-) energy positions, which are in good agreement. The CNL above the CBM means the donor ViGS remain the important

defect even when E_F is above the CBM.

The reason for the high CNL in TCO systems stem form the large mismatch between anion and cation size, and high electronegativity of the oxygen (giving low O 2s orbital energy) meaning a very low-lying CBM state juts into the gap at Γ, while the VBM states are not very dispersive. The ViGS are a result of defect centres that are localised in real space and so are non-localised in reciprocal space, i.e. they are well spread out in the Brillouin zone, not just localised at Γ. The energy gap averaged over the Brillouin zone is larger than the VBM to CBM at the Γ-point in TCO materials and so the CNL comes above the CBM[71].

IMPURITY DOPING

While intrinsic defects may cause a reasonably high level of conductivity in unintentionally doped TCOs and can play an intricate but fundamental role in determining the levels of dopability achievable in a material (for example through cation vacancies acting as compensating acceptors), for most applications undoped TCOs are not conductive enough to be of particular use. Dopant impurities in the form of substitutional dopants are usually incorporated into TCO materials in order to increase the carrier density and so the conductivity of the film or crystal. This is usually done using one element to the right in the periodic table of one of the host elements, which usually ensures a similar size atom substitutes the host (minimal disruption to the lattice), and one extra electron is incorporated per dopant atom, giving the desirable conditions for high conductivity. To achieve metallic like electrical behaviour, i.e. to dope past the insulator-metal transition, the Mott criterion must be surpassed. It was already established that $n_c \sim 1.6 \times 10^{18}$cm^{-3} for In_2O_3. However, there are many other factors to consider when doping TCOs which include the solubility of the dopant atoms in the lattice; the energy positions of the defect levels in comparison to the host CBM; any

natural compensation mechanism inherent with the dopant; the charge state of the impurity; the balance between doping and alloying and its affect on the material electronic and physical structure.[85] Some of these properties are extremely dopant specific while others depend as much on the host material.
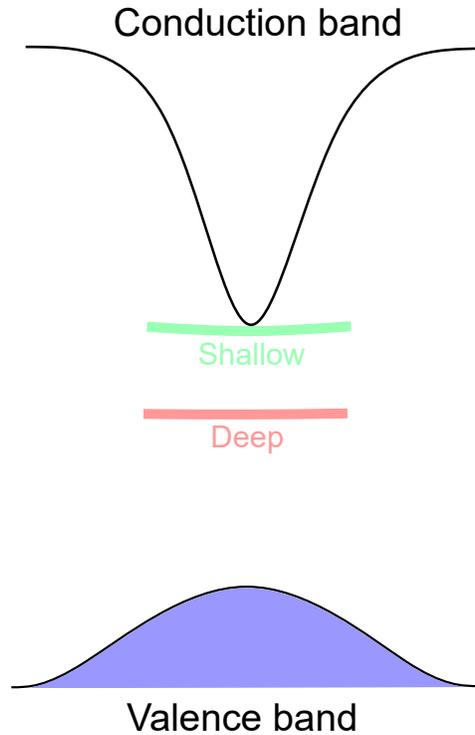
Figure 1.14: Schematic diagram of the relative positions of different impurity dopant states relative to the band edges in a TCO.

Figure 1.14 shows a schematic of different energetic positions that impurities or dopants can take relative to the band edges in a TCO. A deep donor is one which is usually considered to be energetically far from the CBM, and so takes too much energy to thermally ionise. Hence, deep donors do not contribute to conduction. A shallow donor sits no more than ~ 0.03 eV from the CBM and so can thermally ionize into the conduction band, giving free carriers at room temperature.

For an n-type material, conductivity can be given by equation
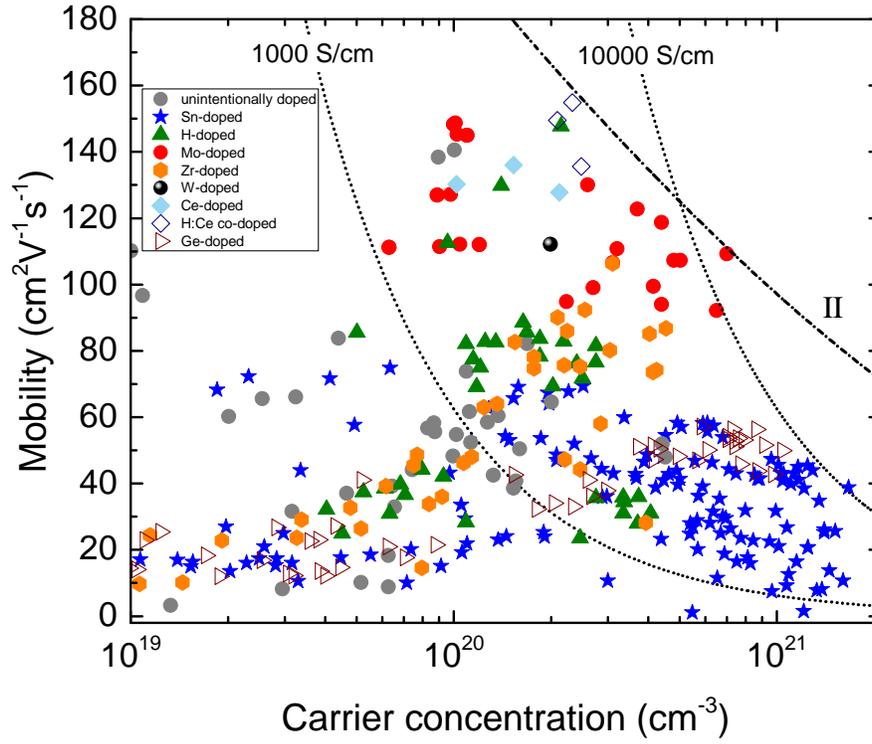
Figure 1.15: Mobility vs free carrier concentration for In_2O_3 with a range of dopants. Figure adapted from Frischbier *et al.* including data taken for this thesis.[86,87] The line labelled II represents the mobility limit due to ionized impurity scattering.

$$\sigma = ne\mu \qquad (1.14)$$

where n is the number of free carriers in the material and μ the mobility of those free carriers. The mobility relates to the electron effective mass as

$$\mu = \frac{e}{m^*}\bar{\tau} \qquad (1.15)$$

where $\bar{\tau}$ is the average scattering time of an electron. Equation 1.14 depends on both the carrier density and the electron mobility and it is often assumed in the rigid band model that the former is controlled by the amount of dopant incorporated into the material whilst the latter is a property solely of the host material. There is some truth in

this, the electron mobility is intimately linked with the effective mass which depends on the curvature of the host conduction band. This helps explain why TCOs are such good conductors, as the host conduction band typically displays a very low effective mass and so has mobile carriers and a high conductivity. Whilst achieving a high number of free carriers through doping is a successful route to higher conductivity, issues do arise with increased free carrier absorption and reflection, and ionized impurity scattering which can limit the transparency and conductivity of a film. Hence, maintaining a high electron mobility in a material may be a more favourable way to achieve high conductivity in TCOs.

Figure 1.15 shows Hall mobilities as a function of free carrier concentration for a number of In_2O_3 films with different dopants incorporation. It may seem surprising that despite the host material remaining the same, the mobility at a given carrier density is very varied depending on the dopant. If it is assumed that the carrier scattering time (i.e. the average time that the carrier is ballistically accelerated by an electric field before colliding) is held constant, then the effective mass is the only property that can vary the mobility. Classically, the band edge effective mass is a property of the host band structure only and so will not change under doping. Variations in the band structure in the context of TCO materials were first observed by Walsh *et al.*[38] when discussing band gap renormalization in Ge and Sn doped In_2O_3, where Ge and Sn modify the host band structure by different amounts when incorporated into the material. However, comparative studies using different dopants are very much lacking in the literature. This sort of electronic structure variability would have massive implication to the electrical as well as optical properties in TCOs (as evidenced by figure 1.15), and greater understanding of the doping mechanisms and dopant selection is greatly required if progress is to be made in creating better TCOs.

It is worth mentioning that it may not necessarily be the case that the average scat-

tering time $\bar{\tau}$ is a constant, and in fact it seems likely that both m^* and $\bar{\tau}$ vary. Indeed, it has been previously suggested that for materials such as Mo:In$_2$O$_3$[88] and La:BaSnO$_3$[89], where the dopant energy levels lie high in the conduction band, that the conduction band wavefunction in the vicinity of the conduction electrons does not greatly overlap with that of the dopant atom. This results in lower scattering from the ionized impurity atoms, which are energetically remote. This situation is analogous to modulation doping such as that seen in GaAs/Al$_x$Ga$_{1-x}$As superlattice structures[90], whereby the dopant atom is spatially separated from the donated conduction electrons. Because the effects of either m^* or $\bar{\tau}$ changing will have similar effects on the mobility, and it is difficult to determine the extent to which each property is changed, for simplicity here the main focus is on changes in the effective mass and the subsequent modification of the conduction band in this thesis.

1.3 Organisation of this thesis

The remainder of this thesis aims to investigate the fundamental properties of TCO materials, paying special attention to advancing the current understanding regarding doping in commercially relevant or interesting TCOs. The next two chapters will discuss the basic experimental and theoretical methodology carried out which facilitated this study. Techniques used throughout this thesis include photoemission, Hall effect, optical spectroscopy and first principles calculations under the density functional theory formalism. The main results chapters are split into three studies, chapter 4 giving new results and understanding of an extremely commercially relevant TCO, fluorine-doped tin dioxide (F:SnO$_2$), looking at its electrical properties and highlighting issues with a classical dopant approach. Chapter 5 gives a direct comparison of a novel doped system and its conventionally dopant counterpart (Mo and Sn:In$_2$O$_3$), demonstrating the electronic and optical gains possible by selecting non-conventional dopants. Fi-

nally, chapter 6 investigates β-Ga$_2$O$_3$, a topical material which is currently generating a lot of interest in the field of high power electronics research. Comparing photoemission to DFT calculations allows for a full investigation into the electronic surface properties of this not well understood material. The findings are related back to the concept of the CNL and H transition level. Finally, the main findings of this work are summarised in chapter 7.

2

Experimental Techniques

Throughout this thesis the experimental techniques of photoemission spectroscopy, optical spectroscopy including transmission, reflection and absorption spectroscopy, and single-field Hall effect, were used extensively. These techniques lend themselves nicely to studying the electronic structure and properties of materials. This chapter presents an introduction to both the theoretical background of these techniques, and also experimental and technical details general to the experiments performed. Since so much of the experimental work performed during my PhD focussed heavily on photoemission, it seems only fitting that much of this chapter is taken up by discussing the subject.

2.1 Photoemission Spectroscopy

Photoemission spectroscopy (PES) is a technique based around the photoelectric effect, whereby an electron is emitted from a solid upon light illumination, and its energy measured. Influenced largely by the work of Siegbahn[91] for which he won the 1981 Nobel prize, PES has become an extremely powerful tool for studying electronic properties, the electronic structure, and chemical environments present in a material. It is also an attractive technique to the experimentalist due to its simplicity, and the breadth of information available from a measurement. XPS is used throughout this thesis, and the discussion regarding primary and secondary structure will be relevant to all results chapters. Discussion of plasmon loss satellites will be of particular interest for chapters 4 and 5. Finally, discussion regarding photoionization cross-sections are relevant to all results chapters, and an example displaying how the are used is shown at the end of section 2.1.4.

2.1.1 PRINCIPLE AND THEORY

In the simplest terms, the photoemission process can be broken down into three stages[92]. Firstly, photons of a characteristic energy $h\nu$ (e.g. Al Kα sources give x-ray photons of ~1486.6eV) incident on a material, interact with electrons in atomic orbitals generating photoelectrons. Secondly, these electrons move through the sample to the surface, most being subjected to various scattering processes which form the background counts (known as the secondary electrons), while some do not suffer any inelastic collisions (primary electrons, which result in distinct spectral features mirroring the samples density of states). Lastly, electrons reaching the surface are then emitted into the vacuum (overcoming the materials work function). In this simple picture, the electrons are emitted from valence and core levels via the photoelectric effect, and if

no other energy loss mechanisms occur they are liberated with some kinetic energy (E'_k), relating to the photon energy ($h\nu$), the binding energy of the orbital E_B, and the material work function (ϕ_s):
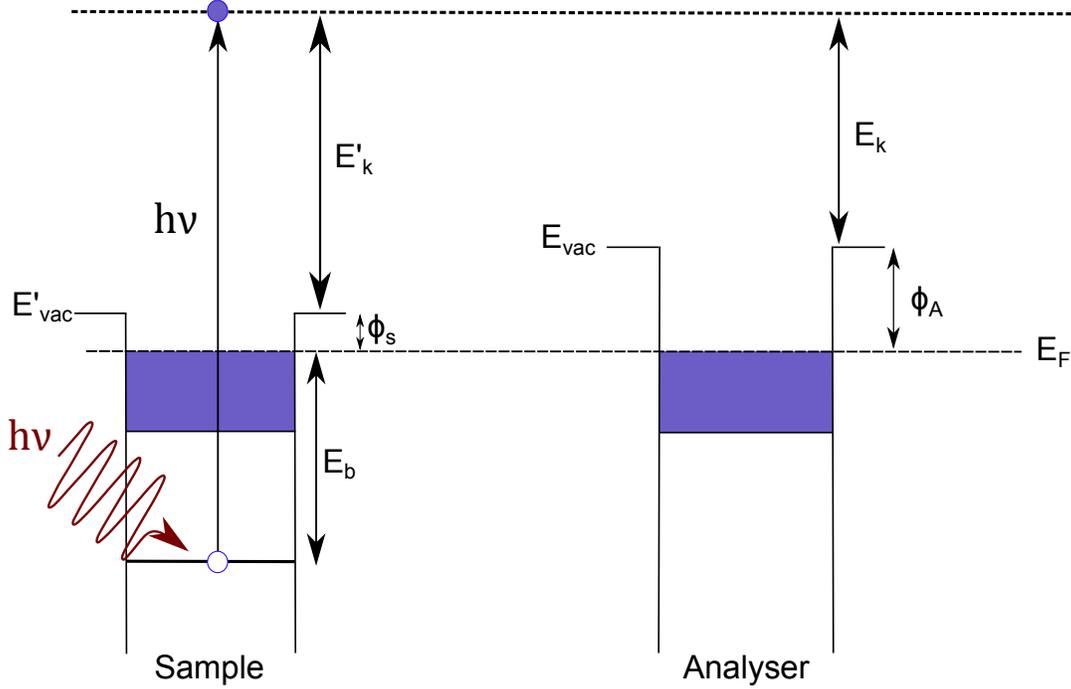
Figure 2.1: Schematic band alignment between sample and analyser in electronic contact in a PES measurement. Energies are aligned to the Fermi level. The purple shaded region indicated occupied density of states just bellow E_F, and the uppermost dashed line represents the vacuum level.

$$E'_k = h\nu - \phi_s - E_b. \tag{2.1}$$

All energies are referenced with respect to the Fermi level E_F. The sample work function can be described as simply the energy difference between the Fermi level E_F of a material, and the vacuum level. Physically this is the minimum energy required to take an electron out of the material into vacuum, where the liberated electron is at rest. This potential barrier exists due to contributions from the bulk inner potential of the crystal under investigation, as well as the surface dipole due to the termination of the

perfect crystal periodicity allowing the electron density to spill into the vacuum. This means for photoemission the concept of the vacuum level is actually material specific as it is defined near the sample surface and different materials display different surface dipoles. This is in contrast to the vacuum level at infinite distance fom the sample which is fixed.

When a sample is connected to an analyser (i.e. they share the same ground and so are in electrical contact), the Fermi levels align and the situation displayed schematically in figure 2.1 is arrived at. Because the sample is connected to the analyser, and the analyser has its own work function (ϕ_A), a contact potential is set up between the sample surface and analyser ($\phi_A - \phi_s$), equation 2.1 can be written as

$$E_k = E'_k - (\phi_A - \phi_s) = h\nu - E_b - \phi_s - (\phi_A - \phi_s)$$
$$= h\nu - E_b - \phi_A \tag{2.2}$$

which is conventionally rearranged to give

$$E_b = h\nu - E_k - \phi_A. \tag{2.3}$$

Equation 2.2 is independent of the sample work function, and since the analyser work function is a constant, it is possible to convert the measured electron kinetic energy to the more useful binding energy. In practice however, ϕ_A is difficult to evaluate accurately, and so the energy axis requires calibration by other means. Fortunately the Fermi energy of a metal surface (which has occupied density of states up to E_F) can be used to overcome this issue and calibrate the energy scale. If the measured energy of the Fermi edge of a metal (in electrical contact to the analyser) is set to $E_k = h\nu$ and a shift of the entire spectrum is applied accordingly, then the binding

energy scale becomes $E_b = h\nu - E_k$ which is independent of ϕ_A, and makes E_F the zero of the binding energy scale. Subsequent to calibration, any sample can be measured with a calibrated energy scale assuming the sample is electrically in contact with the analyser (and does not charge).
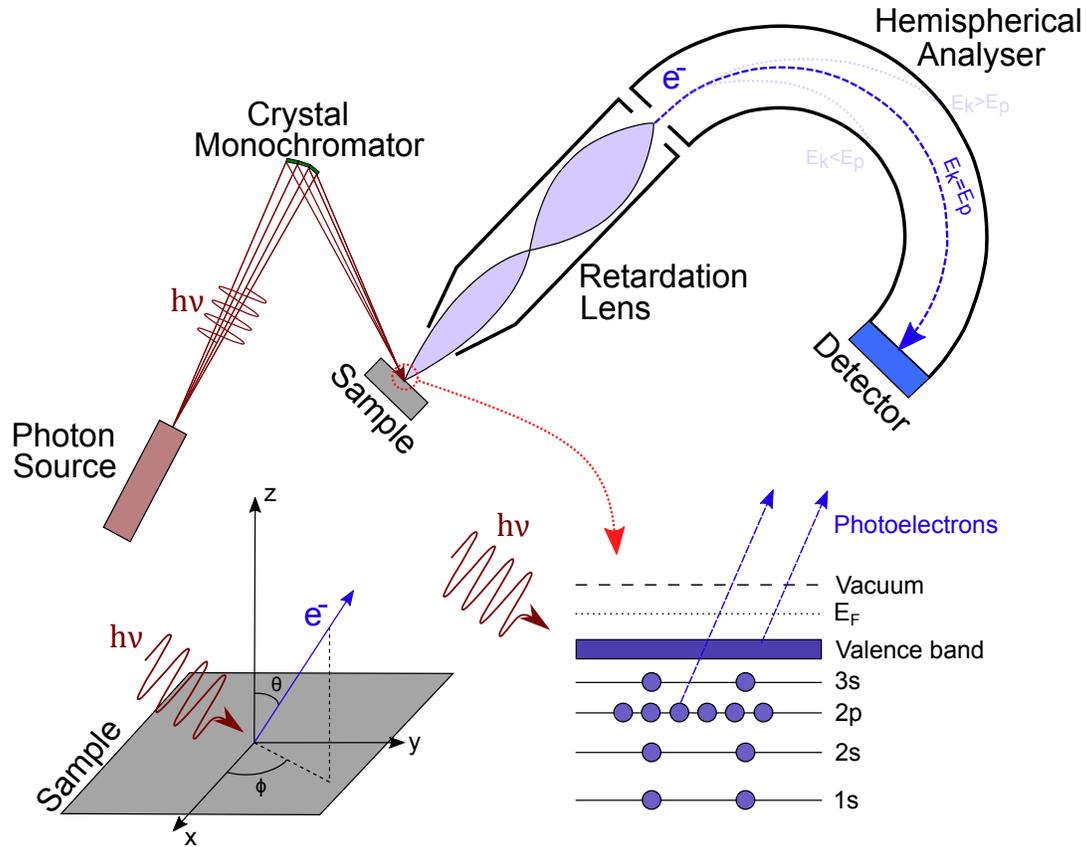
Figure 2.2: Schematic representation of a photoemission experiment showing the generation and path of a photo-electrons through the experimental set-up, an energy level diagram showing the photoemission of electrons form core and valence levels, and the geometry of the photoionization process with respect to the samples coordinate system.

The experimental set-up for a photoemission experiment is shown in figure 2.2. There are a few main features common in most PES experiments that will be discussed here. The first requirement is a photon source which typically takes the form of a lab based x-ray gun, gas discharge lamp, or synchrotron source depending on the required energies for the experiment (and the set-up available to the experimen-

talist). A monochromating crystal (often (1010) oriented quartz in lab based systems) is extremely common in the modern PES experiment. Whilst not strictly necessary, this gives the benefits of avoiding satellite features due to multiple x-ray wavelengths (not only Al Kα_1), and improves on resolution as a result of a more narrow beam line width (independent of the line width from the source, instead the crystal determines the line width at the sample). A retarding electrostatic lens is then required to focus electrons onto the entry slit of the analyser, as well as retard electrons to an energy required to pass through the hemispherical analyser. The retarding input lens usually consists of multiple cylindrical electrostatic/magnetic lenses and apertures which limit the beam angle minimising aberrations as well as improving on resolution. The retarded electron beam enters a concentric hemispherical analyser, designed to filter electrons using an applied potential called the pass energy between the two concentric hemispherical plates, so that electrons with too great or low energy cannot make it to the detector. By sweeping the retarding voltages put on the lenses, it is possible to generate a spectrum with a constant energy resolution. By changing the potential across the plates it is possible to change the pass energy and hence control energy resolution and intensity of electrons measured. The counting process is done in the detector which can consist of a channel and resistive plate combination giving amplification and position sensitivity in the measurement, or a channeltron system utilising many multichanneltron detectors for detection and spatial (angular) resolution.

2.1.2 CHEMICAL SHIFTS

Because PES involves probing the occupied density of electronic states, a lot of information regarding the chemical and bonding environments of different elements can be obtained from the data. A change in the bonding environment of an element is reflected in changes in the binding energy, peak width, peak shape, or a combination

of all of the above. The binding energy of an electron is due to the electrostatic inter-action between the electron and positive nucleus of the atom. Other electrons around the atom may also contribute, causing a screening effect of the nuclear potential, thus lowering the electron binding energy. A simple but reasonably good model to describe the shift in binding is given by the charge potential model[93]

$$E_i = E_i^0 + kq_i + \sum_{i \neq j} \frac{q_i}{r_{ij}} \tag{2.4}$$

where E_i and E_i^0 are the binding energies (for the sample and reference) of a core level on atom i, kq_i is the contribution due to the valence charge density ($k \sim \frac{1}{r_v}$ the reciprocal of the valence shell radius $kq = \int \frac{\rho(r)}{r} dr$), and the final term $\sum_{i \neq j} \frac{q_i}{r_{ij}}$ represents the potential on atom i due to point charges on atom j, closely related to the Madelung potential.

This model assumes the atom is a hollow sphere with electrons on the surface of the sphere, meaning the potential inside the sphere is the same at all points and given by $\frac{q_i}{r_v}$. Whilst simple, the power of this model can already be seen for describing the PES experiment. For example, a change in valence electron charge (density) Δq_i leads to a change in the potential in the sphere by $\frac{\Delta q_i}{r_v}$. Therefore the binding energy of all core levels will change by this amount. In addition, if the radius of the valence orbital r_v increases, then the binding energy shift for a given Δq_i will decrease. Indeed for materials where the core-valence interaction is negligible, it is found that the binding energy shifts decrease between equivalent compounds moving down the columns of the periodic table (similar charges but r_v increases).

$$\Delta E_i = k(q_i^1 - q_i^2) + (V_i^1 - V_i^2) = k\Delta q_i + \Delta V_i \tag{2.5}$$

Focussing on a core level of an atom i in two different chemical environments (1 and

2) and simplifying equation 2.4 by coupling the summation term to V_i, gives equation 2.5, where the first term on the right hand side gives the contribution from changes in valence charge while the second term is the Coulombic interaction between photoelectrons and the surrounding charged atoms. This model while instructive, is based on many assumptions. Often a further term is included ΔR accounting for the relaxation of the final state due to the polarizing effect from the core hole on the surrounding electrons.
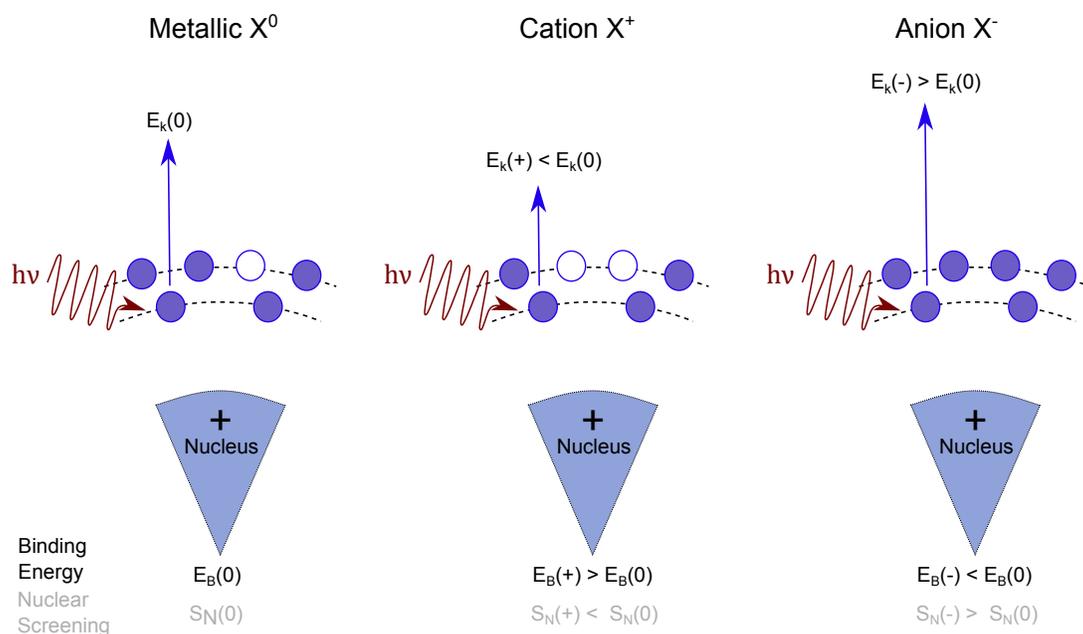
Figure 2.3: Schematic showing the effect of nuclear screening (S_N) in different oxidation state chemical species on the kinetic energy and so binding energy of liberated electrons. The elemental species displayed are a metallic state with zero oxidation state (X^0), a cation state with a positive oxidation (X^+), and an anion state with a negative oxidation (X^-).

Figure 2.3 displays the origin of chemical shifts in XPS schematically (note that XPS and PES are often used interchangeably throughout this thesis as XPS is the most common PES experiment performed today). A metallic (oxidation state of 0) reference element is shown on the left hand side, which is usually used as a baseline to compare to elements in different chemical environments. A change in electron density around

the atom changes the amount of nuclear screening as displayed in the middle and right hand panels of figure 2.3. Considering ionic bonding in solids, one atom is bonded to another and one has a higher electronegativity (the anion). A charge transfer happens from the cation to the anion, and so the cation has less electron density around the atom. The loss of electrons causes the atom to have a positive effective charge, decreasing the nuclear screening and so increasing the binding energy of the electron (measured electrons have a lower kinetic energy). The opposite applies to the case of a more electronegative anion which gains charge density upon forming ionic bonds. In general, the higher the oxidation state of the element around 0, the greater the binding energy shift from the metallic case (positive oxidation meaning $E_B > E_B^0$ and lower oxidation meaning $E_B < E_B^0$). It is worth noting that these general rules are quite oversimplified in the sense that other effects can contribute to chemical shifts, and may counteract these effects. Final-state effects occur after the photoemission process while initial-state effects are always present and can depend on the number of bonds, crystal field effects or ionic potentials. These effects can complicate XPS analysis.

2.1.3 SPECTRAL FEATURES

PRIMARY STRUCTURE

XPS spectra are displayed as the intensity (number of electrons counted) as a function of binding energy, which is a more useful quantity than kinetic energy. However, kinetic energy is customarily plotted from low energy to high, and since binding energy is the negative analogue of kinetic energy it is the convention to plot binding energy from high to low energy. A reference spectrum taken from a single crystal of In_2O_3 is shown in figure 2.4 to provide an example.

In a spectrum, the well-defined peaks are known as core levels and originate from
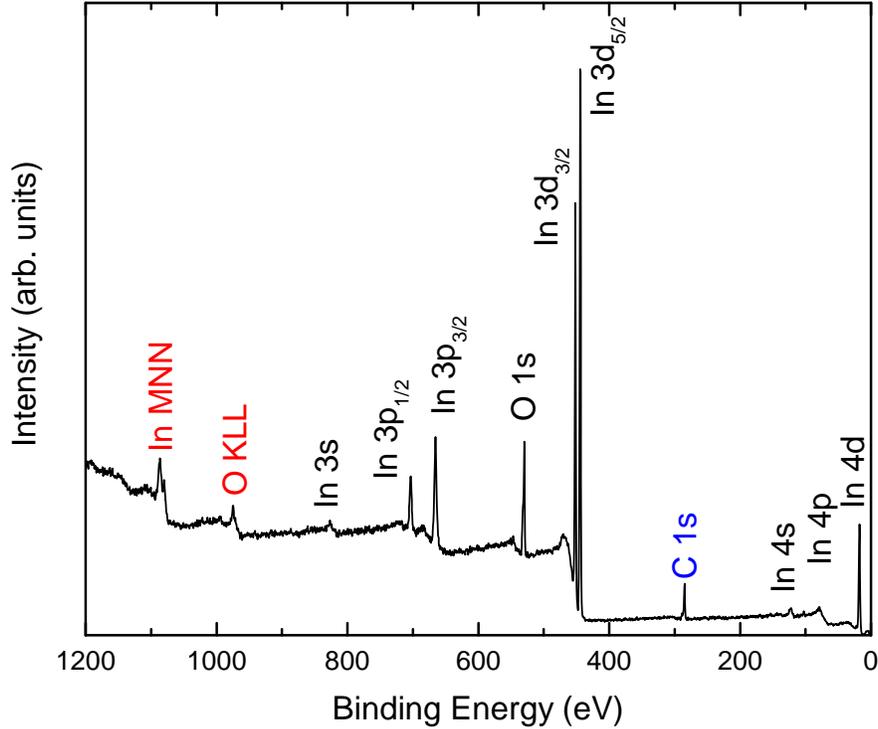
Figure 2.4: XPS survey spectrum from a single crystal In_2O_3 sample measured by the author. Core-level features including spin-orbit split components where visible are indexed in black, while red marks the Auger lines, and blue the carbon contamination.

electrons which do not inelastically lose energy when they emerge from the sample. These lines reflect the shell structure of the electrons around the atom are identified using spectroscopic notation $n\ell_j$ representing the principal (n), azimuthal or orbital (ℓ), and total (j) angular momentum quantum numbers. The total angular momentum can be decomposed to $j = \ell \pm s$ where s is the spin angular momentum and can take two values $s = \frac{1}{2}$. The azimuthal quantum number is customarily represented by a letter also (e.g. *s, p, d...*), which is displayed along with the quantum numbers in table 2.1. For s levels ($\ell = 0$), it is found that the observed peak is a singlet as there is only one value of j. For all other values of ℓ, two values of j are obtained and so two peaks are seen, referred to as a doublet. The two possible states are separated in energy by the spin-orbit splitting, which originates from the unpaired electron left in an orbital after

ℓ, letter	$\jmath = \lvert \ell + s \rvert$	Occupancy $(2\jmath+1)$	Doublet intensity ratio
0, s	1/2	2	
1, p	1/2, 3/2	2, 4	1:2
2, d	3/2, 5/2	4,6	2:3
3, f	5/2, 7/2	6,8	3:4

Table 2.1: XPS spin-orbit doublet intensity ratios.

the photoionization process having two possible spin and orbital angular momentum vector orientations, either parallel or anti-parallel.[94] The magnitude of the spin-orbit splitting ΔE is proportional to the spin-orbit coupling constant $\xi_{n\ell}$ which depends on the expectation value $\left\langle \frac{1}{r^3} \right\rangle$ of the involved orbital, where r is the orbital radius. Therefore ΔE increases with increasing atomic number Z for a given subshell, and decreases for increasing ℓ (if n is constant)[94]. An example of well defined spin-split peaks is the In $3d_{3/2}$ peaks at ~ 445 eV in figure 2.4, which have a distinct energy separation of 7.6 eV. The relative intensity of the two peaks, the doublet intensity ratio, are determined by the ratio of their occupancy, seen in table 2.1.

The valence levels are those with the lowest binding energy (usually $\sim 0 - 10$ eV), and these orbitals are involved in bonding. Therefore the spectrum for these levels usually displays many broad feature reflecting the many closely spaced energy levels. Because the valance band states are situated at low binding energy, they occupy a relatively flat part of the continuum background, meaning they can quite accurately be compared to the calculated electronic density of states. This type of comparison will be a very important theme in this thesis.

Other than the electrons that do not suffer inelastic scattering on the way out of the sample, i.e. those forming distinct peaks, there are the electrons that do lose energy on the way to their departure. There is a finite probability of an inelastic scattering event taking place and this forms the background counts seen in the spectra as a smoothly

increasing level of counts (increasing as binding energy increases). The energy loss processes are random and so the background forms what is called a continuous spectrum. The core levels then sit on this background level of counts. Noise is a spectral feature that occurs due to the the counting process of electrons, and not instrumental in origin. The signal-to-noise ratio is proportional to the square root of the counting time.
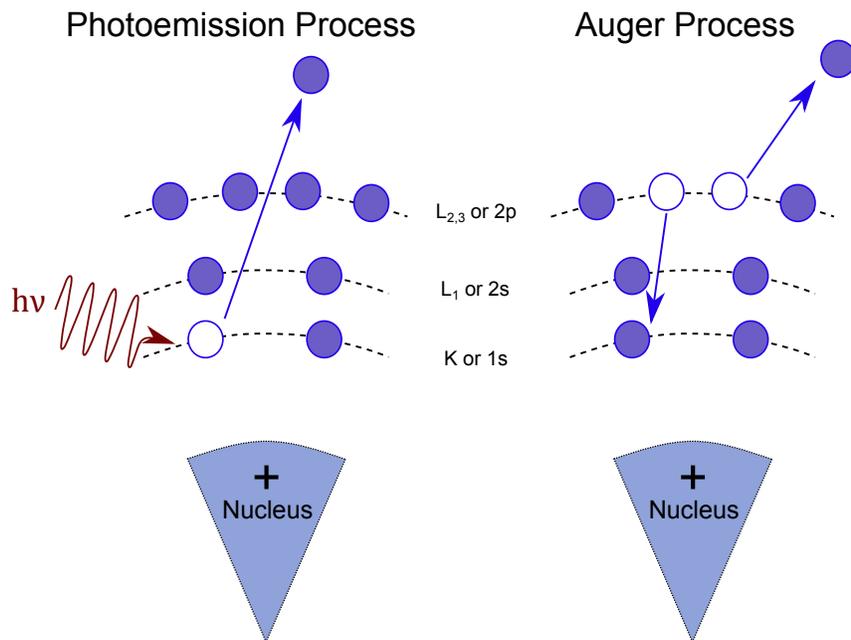
Figure 2.5: The photoemissin process (left) for a model atom with incoming photon causing the ejection of a photoelectron, and the relaxation or Auger process (right) showing the resulting emission of an electron after a higher orbital has filled the core hole and transferred its energy.

Auger features are another spectral line, or more commonly a series of peaks, that comes about from the de-excitation of an outer shell electron filling a core-hole state following the photoemission process forcing the system into a more neutral state. This releases a discrete amount of energy and can excite another electron which may be emitted, see figure 2.5. The kinetic energy of the emitted electron from an Auger process can be described using the following equation

$$E_{Kin} = E_B^{CoreLevel} - E_B^1 - E_B^{2*} \qquad (2.6)$$

where $E_B^{CoreLevel}$ is the energy level of the photoionized core level, E_B^1 is the binding energy of the level from which an electron moves to fill the hole, and E_B^{2*} is the binding energy of the level from which the Auger electron is emitted (where * indicates this level is also affected by the hole in level (1) and so not the same as E_B^2 in the ground state). Auger spectroscopy traditionally used x-ray notation to denote the relevant transition levels involved in the process, e.g. the KLL series involves the initial vacancy in the K shell and final double vacancy in the L shell. The symbol V (such as in the KVV Auger process) indicated the final vacancies are situated in the valence band. Auger lines have a kinetic energy which is independent of the ionizing radiation used, and so appear at different binding energies when different energy photon sources are used (see equation 2.3). The indium MNN Auger feature for In_2O_3 can be seen in the survey in figure 2.4 at $\sim$1075 eV. It is a wide set of features, with an energy expanse of at least 30 eV, each peak much broader than typical core-lines. This is due to many transitions occurring, as well as the different lifetimes of these transitions. Studying the Auger spectra of an element in a material can provide complementary but unique information to XPS core-level analysis.

SECONDARY STRUCTURE

Other than these features, many other physical phenomena can take place that cause other peaks to occur, often referred to as secondary structure in a spectrum (the spectral features already discussed making up the primary structure). A brief mention of the most commonly seen features will be given here. Arguably one of the most important secondary spectral features relevant to TCO materials as well as metals are plasmon loss lines[95]. These plasmon satellite features have fairly complex structures,

so much so that the two physical phenomena that cause plasmon losses can hardly be distinguished experimentally in many cases. These two phenomena are referred to as "intrinsic" and "extrinsic" plasmons.

Extrinsic plasmons are excited by electrons travelling through the material, which have a certain probability to interact with the free conduction electrons in a conductive sample, exciting the free electron gas causing them to oscillate with a characteristic frequency (energy), forcing energy loss in the photoelectrons which emerge with a higher binding energy than the initial core level. To explain the intrinsic plasmon features it is important to remember that in the photoemission process a positive photo-hole is generated. The conduction electrons then act to screen the core-hole potential that has been set-up by the photoemisison process, causing quantized excitations in the conduction-electron system at a distinct energy. Since this excitation process is an intrinsic property of the photoemission process these are known as intrinsic plasmon excitations. The generation of the photohole in a core level can be thought of as an instantaneous switching on of a strong local potential, which forces a coupling between the hole-electron system, whose interaction leads to excitations. It may be expected that the smaller the energy transfer, the greater the coupling.

The distinct energy given to excite the collective oscillation of the conduction-electron system is the plasma energy ω_p given previously by equation 1.5 (the surface plasmon energy, equation 1.8, is also important for surface electrons). In the case of extrinsic plasmons, multiple loss events can occur from the same core level (or valence level) at the same energy spacing from the peak involved, essentially generating a cascade of loss features. In the case of intrinsic plasmons, several losses may be possible but due to the quickly diminishing nature of the loss intensities a cascade of features is not usually seen. Whilst energy loss features in metals can be as much as 20 eV from the main core level, usually in TCO materials a much lower plasma frequency of around

$\omega_p \sim 0.5 - 1.0$ eV is observed, meaning the plasmon loss can overlap heavily with the core level of interest easily causing confusion[96]. Usually only one, very strong plasmon loss feature is seen in TCO spectra, which would suggest the intrinsic plasmon mechanism is dominant for these materials, consistent with the picture of low energy transfer leading to strong features. Extrinsic plasmon losses still occur in dilute electron systems, however due to the much lower number of free carriers relative to typical metals, these events are much less probable.

Plasmon loss features have proven extremely difficult to model in photoemission spectra, even for simple metals. Metals benefit in a sense by having very high numbers of free carriers giving rise to large plasmon energies (meaning their plasmon loss features are distinct and well separated), and allowing their losses to be understood by treating the free electrons as a homogeneous electron gas[97-101], i.e. quantized charge density oscillations. One result of such works is that the core level in a metallic material should contain a series of plasmon satellite peaks following a form of Poisson distribution, whose intensities can be described by[51,102-104]

$$I(n) = \frac{e^{-\beta}\beta^{n_p}}{n_p!} \tag{2.7}$$

where n_p is the order of the satellite feature (n=0 is the elastic core peak, n=1, 2, 3... are the 1st,2nd, 3rd plasmon loss peaks). The coupling constant β is give approximately by

$$\beta = \frac{1}{6}\frac{r_s}{a_0} = \frac{1}{6a_0}\left(\frac{3}{4\pi n}\right)^{\frac{1}{3}} \tag{2.8}$$

where r_s is the distance between conduction electrons and n is the carrier density.

In the weak coupling limit ($\beta < 1$) it is expected that the intensity of the plasmon satellites will increase with decreasing carrier density. Indeed this is seen by plot-
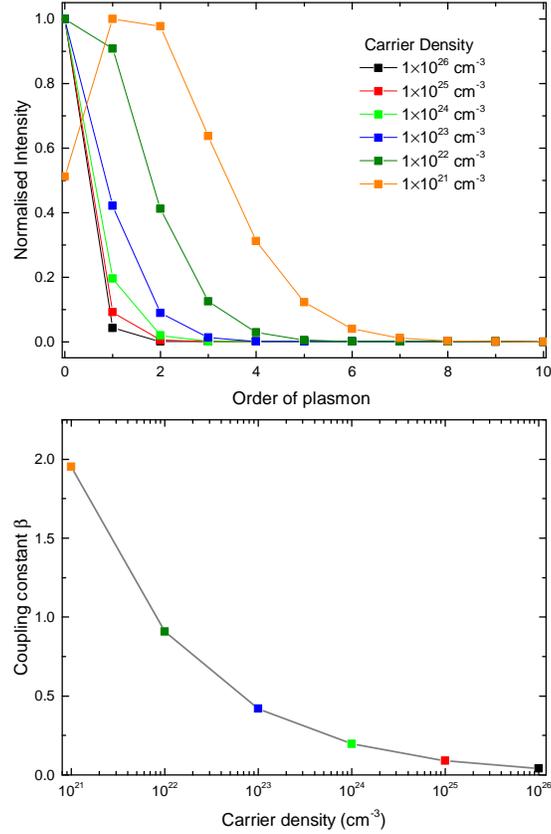
Figure 2.6: Top: The normalised intensity calculated using equation 2.7 as a function of the order of plasmon loss features. Bottom: The coupling constant as a function of carrier density as calculated using equation 2.8.

ting equations 2.7 and 2.8 as seen in figure 2.6, which shows the intensities of a given order plasmon for different carrier density materials, and the magnitude of the coupling constant β as a function of carrier density. Beginning with an unphysically high carrier density of $n = 1 \times 10^{26}$ cm^{-3}, the intensity of the first order plasmon is only around 4% of the elastic peak and the other orders can barely be discerned. Now at $n = 1 \times 10^{22}$ cm^{-3} this ratio increases to around 90% with several orders observable. This is consistent with work previously done on sodium tungsten bronzes[101], which have lower carrier densities than metals but display much stronger plasmon losses. However, when we can no longer assume we are within the weak coupling limit (i.e. $\beta > 1$) these equations fail to describe the physical scenario. This is clear in figure 2.6

when the 1st, 2nd, and 3rd loss feature all have much higher intensity than the elastic peak intensity, and infact if the density of carriers is reduced further the elastic peak can disappear entirely. This is not seen in measurements of dilute electron systems such as TCOs and so this model does not describe them well.

In low carrier density systems, the plasmon loss can be enveloped by the elastic peak, making it very difficult to discern, and can reside at similar energies to chemically shifted components, making identification and quantification very challenging. Furthermore, while metals typically display many loss features, low carrier density systems may only have one loss satellite. It remains a large challenge to accurately describe these loss features in TCO materials. R. Egdell *et al.* have performed some in-depth investigations into this subject over many years. Focussing mainly on Sb-doping of SnO_2 [105,106] they found that the core levels in more heavily doped systems apparently shift by less energy than the VBM. This was described using the Kotani screening model[107], seen schematically in figure 2.7. This screening model describes how different screening effects can be present in the final state when a core-hole is created. This will be described briefly below.

In the model developed by Kotani, two competing final states are available after a core-hole is created following the photoemission of a core level electron. Their model system consisted of a metal with an incomplete shell (denoted as an empty level in figure 2.7), which is forced below the Fermi level when a core hole is created after photoemission as a mechanism to minimise energy. Then the two final states consist of an electron from the (occupied) conduction band dropping into the vacant level now below E_F, or this level remaining vacant. In the former scenario, the level is neutralised (hole is filled) and so the photohole is screened by the localised electron in the now filled band, hence the designation of screened. In the second case, the system is now in a two hole state, the local positive charge is screened by the contraction of the
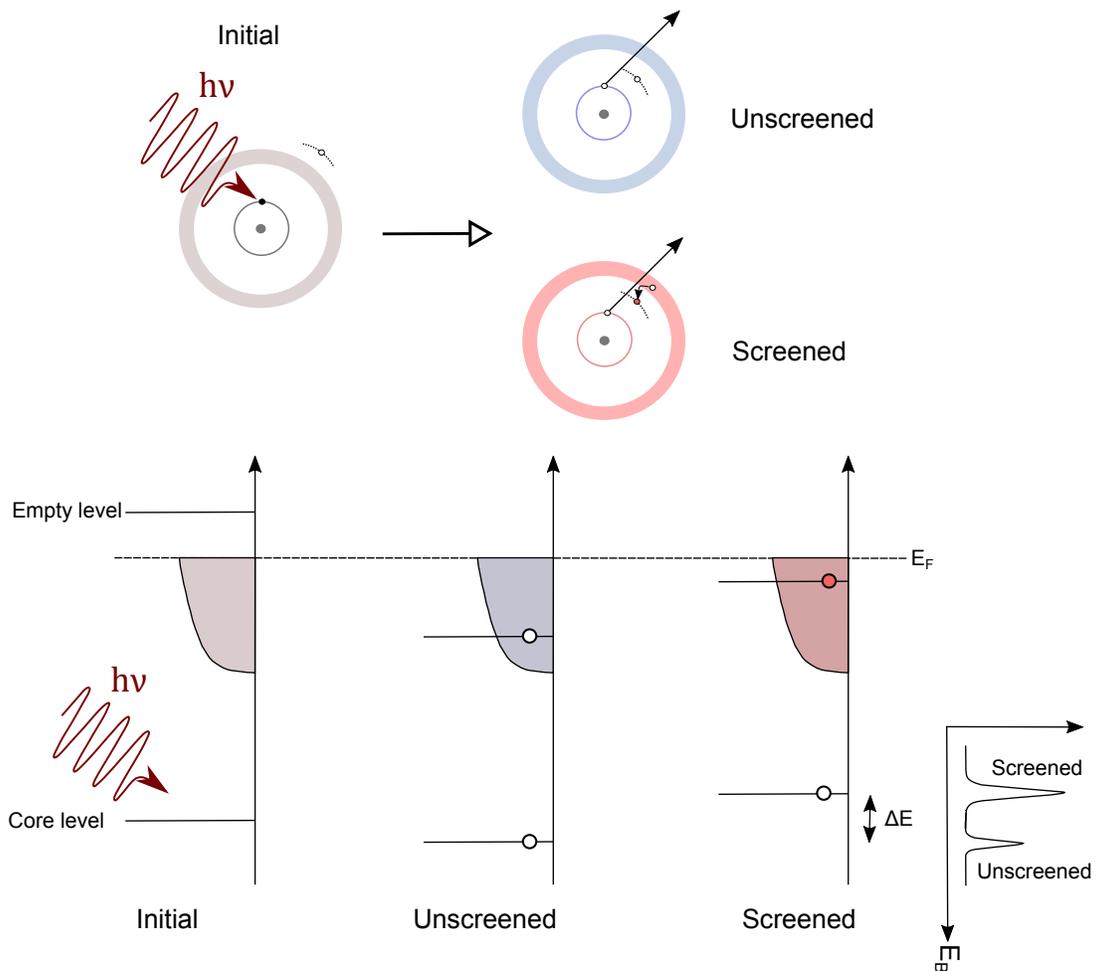
Figure 2.7: Schematic of the final states of the screened and unscreened atomic configurations according to the Kontani-Toyazawa screening model. Figure recreated from Borgatti *et al.*[108] and [109]. The unscreened final state refers to that after relaxation when a localised state pushed below E_F remains empty (meaning its electrical interaction with the core hole is stronger), while the screened states refers to when this hole is filled with a conduction electron (having stronger electron-hole interaction). The difference in the two core level positions of these final states ΔE are equal to the plasmon energy ω_p.

conduction electrons around the hole. The two scenarios differ in relaxation energy which is reflected in the position of the photoelectron energy. Because the screening effect lowers the systems total energy, the escaping electron leaves with higher kinetic energy, and so lower binding energy. In the language of the plasmon model, the unscreened scenario gives the high binding energy peak due to energy transfer to the free electron gas. The width of this peak reflects the plasmon lifetime. Egdell

et al.[105,106] found two component fits reflecting the screened (or no-loss component) and unscreened (plasmon loss component) gave more than adequate fits to core level data for Sb-doped SnO_2 and Sn-doped In_2O_3 dilute electron systems. The difference in energy shift between core levels and VBM for increasing carrier density systems was explained using a redistribution of spectral weight, showing the baricentre of the two peak model gave the same shift as the VBM.

Importantly, it was found that the energy distance between the two peaks was between the plasmon energy for the bulk ω_p and surface ω_{sp} of the material. It was also found that different core levels were screened at different strengths, producing different intensity plasmon loss peaks. This was loosely attributed in the case of Sb-doped SnO_2 to conduction states being mainly Sn 5s in character giving rise to more correlation between the Sn core levels than O 1s level. Whilst Egdell *et al.* provide very useful insights into the nature of the plasmons in TCOs, and give a reasonable approach to plasmon satellite fitting taking into account important parameters such as energy position and FWHM, their model does not provide a robust quantitative description of the satellite peaks, relying mainly of free fitting parameters to provide a least-squares fit. The probabilistic nature of a plasmon loss event compared to a no-loss emission of an electron is not considered. They describe their model as heuristic[106] and suggest some of their findings may be due to coincidence rather than design. The free parameter nature of their fits may mean quantitative information extracted from the peaks may be less reliable. However, it seems clear from these works that, with knowledge of ω_p for a dilute electron density material, we are able to fix this as a fitting parameter with confidence that this reflects the physical scenario, and so may be able to extract some quantitative description from the peaks. It is also worth noting that different core levels can display different plasmon intensities, which is not necessarily obvious as the rate that plasmons are created by photoelectrons from different core levels do

not differ significantly in the extrinsic case. Again, no quantitative description of this was provided. While further work is still required to provide a predictive model of the intensity and shape of plasmon loss satellites in photoemission of dilute electron systems, the approach taken by Egdell *et al.* benefits from its simplicity and a very similar one to this is adopted in this thesis.

It is also worth discussing a recent study by Borgatti *et al.*[108] who combine the fitting approach developed by Egdell *et al.* on $Sb:SnO_2$ for high-energy XPS data, with high level first principles calculations to help further discuss the nature of the plasmon with a theoretical approach. Their calculations were performed at the *GW* level allowing them to account for the many-body interactions associated with (among other things) plasmon excitation. They show by calculating the spectral functions of the Sn 4d and valence band regions, and comparing the quasipartical part (coherent part) to the full spectral function, that the high binding energy component in these levels is not present in the coherent part of the spectral function and so is not due to the quasiparticle part of the spectrum, but instead the incoherent part of the calculated spectrum (see appendix C for more details on spectral functions). They attribute this to the collective oscillations of free carriers in the doped system. Then by varying the carrier density in their theoretical material from undoped to highly doped, they observe a further broadening in the incoherent high binding energy part of spectral function with increasing carrier density. This is consistent with experimental observations[105]. However, they also observe a weakening of the satellite intensity as carrier density is increased. This is not supported by experiment[105], and more closely reflects the predictions of the weak-coupling limit approximation used for metals. This may be related to the method used to artificially introduce free electrons into the system. By using a homogeneous free electron gas as the basis for doping, they cannot achieve any local Kotani-like screening mechanism.

The general trends in the data of both the Sn 4d core level and valence band are replicated reasonably using the *ab initio* approach, although the effects of the plasmon loss seem rather more muted in the theory. This approach shows how the satellite structure relates to coupling of core and plasma electrons. However, this requires a fairly advanced level of theory which may be less accessible to most, and this work did not provide new information regarding the fitting of experimental data (e.g. line shapes, intensities, etc.), which is sorely needed to properly account for plasmon loss features. In addition, their findings do not correctly predict the relationship between intensity of the plasmon peak with carrier density, which will require further investigation to determine. They also do not attempt any calculations of true core-levels, only semi-core levels, and so we have no knowledge of how this level of theory extends to such spectra.

Other spectral features which may be of interest include multiplet splitting and shake-up lines. Multiplet splitting come about in systems with unpaired valence electrons. For example, many transition metal elements have unfilled d-shells in a configuration such as nd^5 with all spins aligned. If the ns^2 level is then ionized, a single s electron remains in this level which is unpaired and can have spin either parallel or anti-parallel to that of the nd electrons. This sets up an exchange interaction whereby the parallel spin state has lower energy than the anti-parallel, which splits the s level. Clearly, for other orbitals apart from s-levels the structure can be more complex due to multiple final states being available. Shake-up lines are those that result from the finite possibility that after the photoemssion process occurs, the ion is left in an excited final state (a few eV above the ground state) with a valence electron promoted to an empty conduction state. When this occurs, the kinetic energy of the outgoing electron is reduced (its binding energy increases) which is seen as a satellite to higher binding energy on the core level peak. This event is more likely in elements with un-

paired electrons in the valence states. Similarly, shake-off features are those when a photoelectron excite valence electrons out of the material. These can manifest in metals (with high densities of states at E_F) as a continuum of shake-up features (as there are a continuum of states above E_F and no band gap to overcome). Therefore the core line presents with an asymmetric tailing to the high binding energy side of the peak.

2.1.4 Inelastic Mean Free Path and Photoionization Cross–sections

XPS is primarily used to study the surface of materials, with the surface sensitivity being determined by the escape depth of the generated photoelectrons and so mainly by the inelastic scattering of electrons in a solid. Here the concepts involved with the mean free path of electrons will be briefly introduced. The inelastic mean free path (IMFP) is the most basic quantity for describing electron scattering, and it is defined as the average distance that an electron with a given energy travels between successive inelastic collisions. Although now outdated, the 'universal curve' originally provided by Seah and Dench[110] was an early attempt to model how the inelastic mean free path varied with electron kinetic energy. Utilising a large database of material data, they noticed general trends in the variation of the inelastic mean free path λ with energy. At low energy < 15 eV, λ varies with E_k by a power law of the order E^{-2}, while at energies $> 75eV$ the variation follows much closer to $E^{\frac{1}{2}}$. Following this they propose the empirical model of the form

$$\lambda = \frac{A}{E^2} + BE^{\frac{1}{2}} \tag{2.9}$$

where A and B are constants related to the material class under investigation (sometimes seen related to a parameter $a = \left(\frac{M}{\rho N_A}\right)^{\frac{1}{3}}$ where N_A is Avogadro's number, but the relationship varies depending on whether the materials investigated are elemental or compounds in nature, and if compounds, organic or inorganic.[111]). Equation 2.9 is
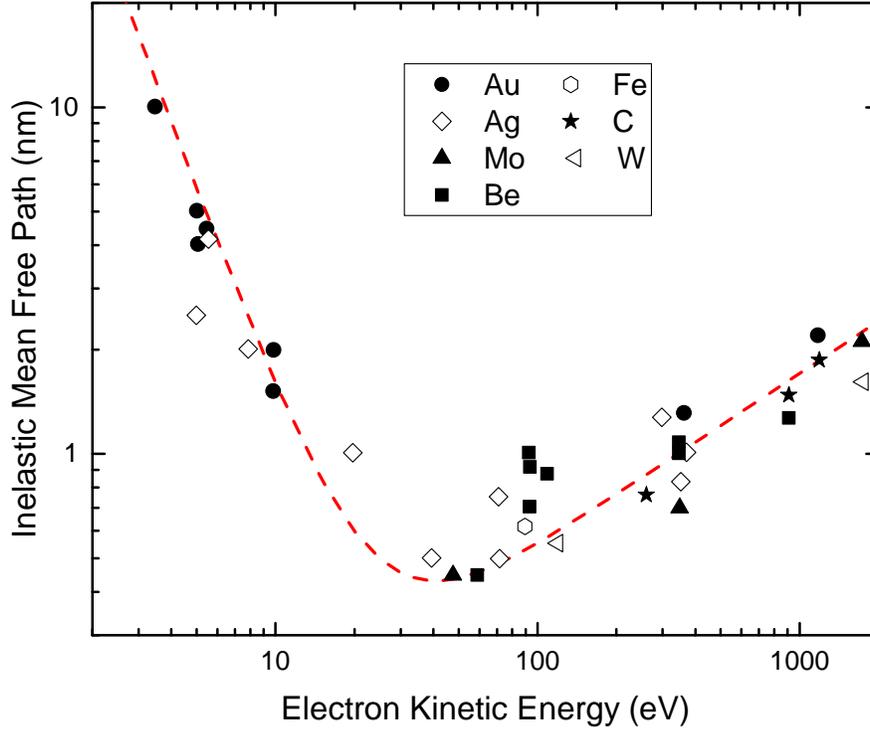
Figure 2.8: The inelastic mean free path (IMFP) of electrons in solids as a function of kinetic energy and the theoretical 'universal curve'. [110].

plotted in figure 2.8 accompanied by some elemental data. Clearly this two component fit takes account of the two general trends seen for low and high energies. However, it has since been found that this empirical approach does not hold for all materials and is quite limited for compounds. What is more, it is now believed that there is no universality between IMFP for all materials. More sophisticated models have since been developed. One of these most commonly utilised is known as the TTP-2M equation (after Tanuma, Powel, and Penn who proposed the model) [112,113]. Here they give the electron inelastic mean free path of an electron of energy $E_k = \frac{\hbar^2 k^2}{2m}$ in a free electron gas as

$$\lambda = \left(\frac{\hbar k}{m}\right)\left(\frac{\hbar}{2|M_I(k)|}\right) \tag{2.10}$$

where $M_I(k)$ is the imaginary part of the electron self-energy, a term related to the Lin-

dard dielectric function which accounts for electric screening of electrons in a solid. We will not discuss Lindhard theory in any detail here but just acknowledge it is a description of how an electron system reacts to a weak potential. This has many parallels to the dielectric equations discussed in appendix B. Hence, the imaginary part of this function is similar to the imaginary part of the dielectric function, which will account for the loss of energy in the system through something similar to the Beer-Lambert law. [114–116] Utilising optical data for $|M_I(k)|$ they calculate the IMFP for a number of elements, as well as organic and inorganic compounds as a function of energy. This approach gave strong experimentally informed trends which they then developed the TPP-2M model to fit.

Their model is based on a simplified form of the Bethe equation for the total cross-section of inelastic scattering of electrons and is given as [113,117]

$$\lambda = \frac{E}{E_p^2 \left[\beta ln(\gamma E) - \left(\frac{C}{E}\right) + \left(\frac{D}{E^2}\right) \right]} \tag{2.11}$$

where

$$E_p = 28.8 \left(\frac{N_v \rho}{M} \right)^{\frac{1}{2}} \tag{2.12}$$

is the free-electron plasmon energy in eV, N_v is the number of valence electrons per atom or molecule, ρ is the density and M the atomic/molecular weight of the material. The parameters β and γ are fitting parameters that require least squares analysis to properly determine, designed to scale the energy dependence of λ. Additional parameters (C and D) were included in the model to account for exchange effects (significant at low energies), and other departures from the Born approximation, on which the Bethe equation is based. Further information regarding the simplifying assumptions to the Bethe equation and the empirical equation 2.11 can be found in references 118–

120. However, it is sufficient to know that equation 2.11 is analytical and does not arise from first principles. Tanuma *et al.* provided empirically determined equations relating these parameters to physical material constants.

$$\beta = -0.10 + 0.944(E_p^2 + E_g^2)^{-\frac{1}{2}} + 0.069\rho^{0.1}$$

$$\gamma = 0.191\rho^{-\frac{1}{2}}$$

$$C = 1.97 - 0.91U \qquad\qquad (2.13)$$

$$D = 53.4 - 20.8U$$

$$U = \frac{N_v\rho}{M} = \frac{E_p^2}{829.4}.$$
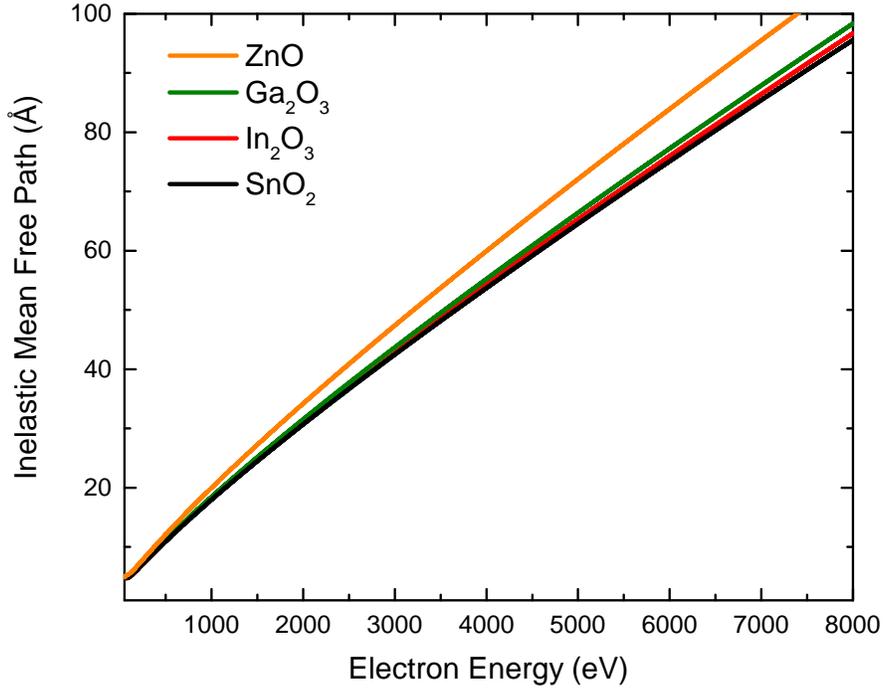
Figure 2.9: IMFP for common TCO materials calculated using the TPP-2M model from equations 2.11-2.13.

However, equation 2.13 are also not from first principles and so should not be confused with true physical constants. Despite this, the empirical model has proved powerful,

deviating by only around 10% from the optical measurements for elemental materials, and 9% for organic compounds. Figure 2.9 shows a plot of the calculated IMFP of several common TCO host materials. All of these materials display similar IMFP as they all have not too dissimilar properties relevant to equation 2.13. Clearly, for electrons in the energy range of $\sim$ 1000 eV, the IMFP is $\sim$ 1.8 nm, which is why the majority of the electrons contributing to an XPS spectrum which do not undergo energy loss, originate from the top few atomic layers of the sample.

It is clear that the IMFP has a large bearing on the surface sensitivity of PES experiments. In fact, we can represent to a reasonable approximation that the intensity of an inner shell photoelectron peak is given by [121,122]

$$I \propto n\sigma\lambda\varsigma \tag{2.14}$$

where n is the concentration of atoms, λ the inelastic mean free path, σ the photoionization cross-section of a given subshell, and ς is used here here to represent the terms that are specific to the spectrometer, any effects from the experimental geometry or illumination area etc., experimental efficiencies and photon flux. These terms can be coupled as they will be held constant in a PES experiment. Two important factors exist in equation 2.14, the IMFP λ which was discussed above, and the photoionization cross-section σ which defines the probability of a specific photoemission event occurring. This quantity is both orbitally and elementally specific. Common practice when using photoionization cross-sections is to look at the elemental cross-sections of the elements in the studied material. When investigating compounds this is not strictly correct, but does seem to provide reasonable results. This simplification is used due to the high complexity in calculating σ for orbitals in compounds.

Collective databases and tables for the theoretically determined cross-sections are available [123–127]. The calculation of the total photoionization cross-sections for a given

atomic shell involves describing the absorption of a photon by an electron. It is then promoted from an occupied initial ground state to an unoccupied final state which is far from the Fermi level (approximated as a free electron). The electrons are described as moving in a Hartree-Fock-Slater central nuclear potential. If the transition rate between two states of energy E_1 and E_2 for a photon energy of $h\nu$ is described by Fermi's golden rule

$$W_{1\to 2} = \frac{2\pi}{\hbar}|M_{12}|^2 \delta(E_2 - E_1 - h\nu) \qquad (2.15)$$

then an appropriate choice of matrix element M_{12} is required to describe the transition. The simplest approach, as used in the formalism of Yeh and Lindau[124], is the dipole approximation, see Ref. 128,129 for more information.* The total orbital photoionization cross-section in the dipole approximation is given as[128]

$$\sigma_{n\ell}(h\nu) = \frac{4\pi^2 \alpha_0 a_0^2}{3}\frac{N_{n\ell}}{2\ell + 1} h\nu \left[\ell R_{\ell-1}^2(E_k) + (\ell + 1)R_{\ell+1}^2(E_k)\right]. \qquad (2.16)$$

This is the same formalism adopted by many, although the powers of the constants seem to vary in the prefactor from source to source[131−134]. Here α_0 is the fine-structure constant $\sim \frac{1}{137}$, a_0 is the Bohr radius, $N_{n\ell}$ is the number of electrons in the $n\ell$ subshell (the occupation), $E_k = h\nu - E_{n\ell}$ is the kinetic energy of the ionized electron, $R_{\ell\pm 1}(E_k)$ are the one-electron radial dipole matrix elements summed over the two (dipole allowed) final continuum (ionic core plus unbound electron) energy states $\phi_{E_k,\ell\pm 1}$, available from the photoemission of a selected initial ground state $\phi_{n\ell}$. These matrix elements have the form

---

*The dipole approximation looks at the matrix element $M_{12} = \int \phi_2^* H' \phi_1 d^3 r$ where H' is a perturbation by light given a value of $H' = \frac{e}{m_0}\mathbf{p} \cdot \mathbf{A}$. The vector potential $\mathbf{A}$ is approximated by $\mathbf{A} = \mathbf{A}_0 e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)} \sim A_0$ which is only valid at low photon energies where the wavelength of light is sufficiently large. At higher energies more terms in the Taylor expansion of the exponent are required[123,130].

$$R_{\ell\pm1} = \int_0^\infty P_{n\ell}(r) \, r P_{E_k, \ell\pm1}(r) \, dr \qquad (2.17)$$

where $\frac{P_{n\ell}(r)}{r}$ and $\frac{P_{E_k, \ell\pm1}(r)}{r}$ are the initial and final radial wave functions $\phi_{n\ell} = R_{n\ell}(r) Y_{nl}(\theta, \phi) = \frac{P_{n\ell}(r)}{r} Y_{n\ell}(\theta, \phi)$ with $Y_{n\ell}(\theta, \phi)$ being the spherical harmonics of the wavefunction. The calculation in equation 2.16 requires a solution to the radial Schrödinger equations for the initial and final state electron

$$\begin{aligned}
\left(\frac{d^2}{dr^2} + V(r) + E_{n\ell} - \frac{\ell(\ell+1)}{r^2} \right) P_{n\ell}(r) &= 0 \\
\left(\frac{d^2}{dr^2} + V(r) + E_k - \frac{\ell'(\ell'+1)}{r^2} \right) P_{E_k, \ell'}(r) &= 0
\end{aligned} \qquad (2.18)$$

where $V(r)$ is there central potential for the nl^{th} subshell and $\ell' = \ell \pm 1$. It can be seen therefore that $P_{n\ell}(r)$ and $P_{E_k, \ell\pm1}$ are solutions to equation 2.18, and the second part of equation 2.18 has the same potential as the first, with a missing electron, a central assumption to these calculations. These wave functions must also satisfy [131,133,134]

$$\int_0^\infty = P_{n\ell}^2(r) \, dr = 1$$

$$P_{E_k, l} \lim_{r \to \infty} (r) = \frac{1}{\pi^{\frac{1}{2}} E_k^{\frac{1}{4}}} \sin \left(E_k^{\frac{1}{2}} r - \frac{1}{2}(\ell \pm 1)\pi - E_k^{-\frac{1}{2}} \ln(2 E_k^{\frac{1}{2}} r) + \delta_{\ell\pm1} \right) \qquad (2.19)$$

where $\delta_{\ell\pm1}$ is a constant phase shift. By evaluating equations 2.18 the radial matrix element of the wave function $R_{E_k \ell}$ can be obtained and so equation 2.16 can be solved to give the total cross-section of a given orbital.

An example of this is shown in figure 2.10 for the carbon atom, looking at the photoemission from the C 2p level. Some points of interest can be taken from figure 2.10. Clearly the final continuum states have oscillatory functions according to equation
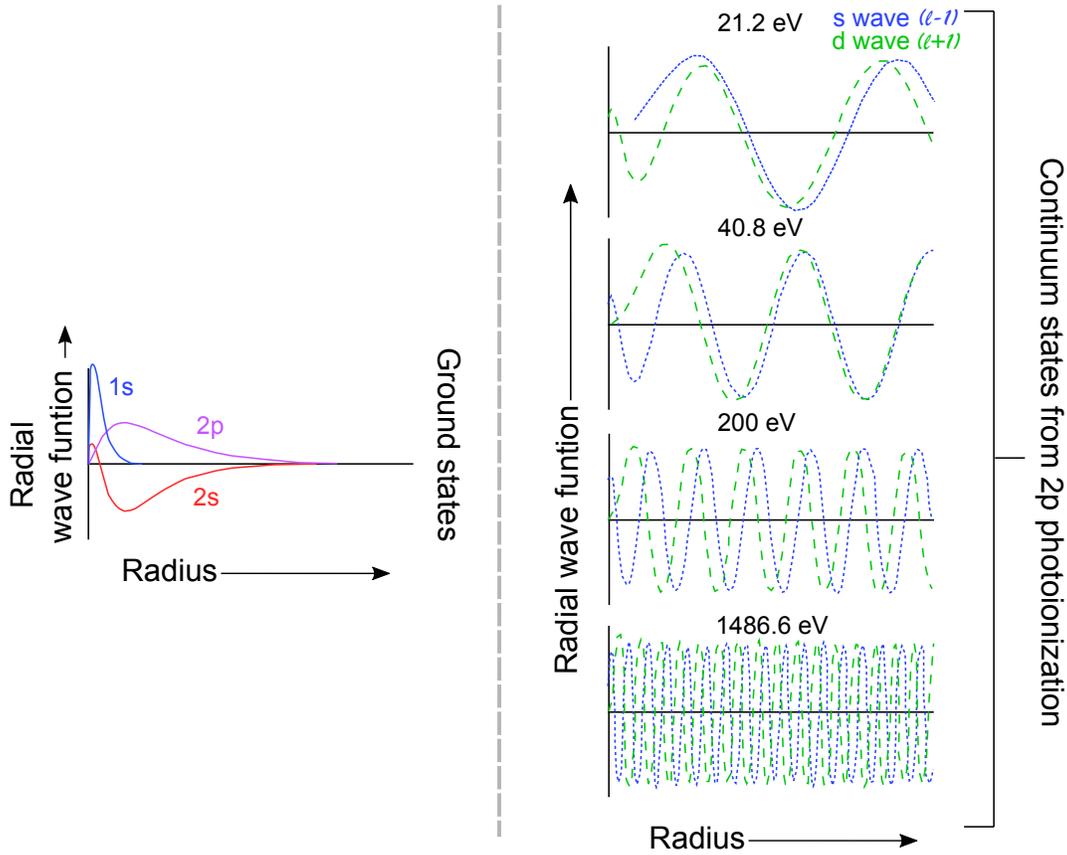
Figure 2.10: Schematic of the radial wave functions $P(r) = rR(r)$ for the occupied ground states of a carbon atom (left), and the continuum states ($\ell + 1$ and $\ell - 1$) for a photoelectorn resulting from the excitation of the C 2p core level as a function of energy (right). Figure replicated from Goldberg *et al.* [131]

2.19, but we also see they are energy dependent with changing E_k and so $h\nu$. As $h\nu$ is increased, the effective wavelength of the radial oscillations decrease and so the oscillations penetrate more deeply into the region of non-zero $P_{n\ell}(r)$ in the atom. There is also a phase shift in equation 2.19, which will persist even when the $\delta_{\ell \pm 1}$ is set to zero due to the different angular momentum quantum number term. For a given photon energy, equation 2.17 is essentially determined by the wave function overlap between the initial and two final continuum state wave functions. The two absorption channels are displayed in figure 2.10, but at high enough energies $h\nu$ the $\ell + 1$ transition becomes more favourable meaning this transition dominated the cross section.
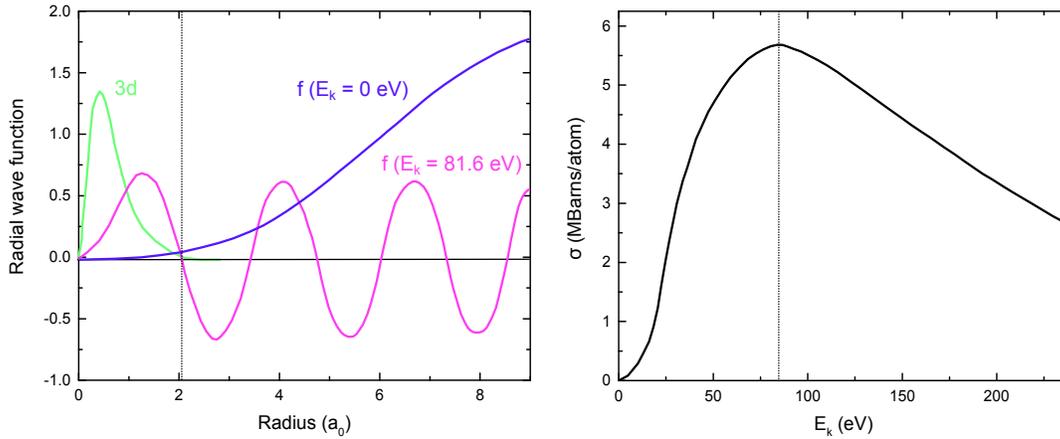
Figure 2.11: (left) Schematic of radial wave function of ground state 3d level and the $\ell + 1$ excited levels at electron kinetic energies of $E_k = 0$ and 81.6 eV for elemental Kr, (right) photoionization cross section for Kr as a function of electron kinetic energy. Image reproduced from Manson *et al.*[134]

As another example, figure 2.11 shows the radial wave function $P_{E\ell}(r)$ for the 3d initial state and the $\ell + 1$ (f wave) state with two electron energies of $E_k = 0$ eV, and 81.6 eV (photon energies of $h\nu = 96.6$ eV and 178.2 eV which were the values used in Ref.[134]) in Kr. These have been overlaid in figure 2.11 for comparison. For $E_k = 0$ eV, the wave function has a long radius. Its intensity is far from the atom, which is due to centrifugal repulsion[134]. Therefore it has very small overlap with the 3d orbital and hence a small dipole matrix element and small cross section. With increasing energy the wave function penetrates deeper into the core of the atom increasing the level of overlap and hence increasing the cross section. In the case of figure 2.11, the $E_k = 81.6$ eV wave function overlaps far more with the 3d level that the $E_k = 0$ eV giving it a higher cross section at this energy. There is also a small amount of cancellation due to the negative wave function of the $E_k = 81.6$ eV at a radius of $\sim 2a_0$. If the energy is increased further, the cancellation becomes stronger, further decreasing the cross section. This explains the shape of the cross section as a function of energy, it is common to see an increase from $E_k = 0$ eV to a nodal point where the maximum overlap occurs, before the cancellation increases and the cross section begins to decrease. Clearly, if

the wavefunctions display more antinodes in a smaller radial range, or the initial wave function has anti-nodal character, this smooth σ relationship can be altered, giving further plateaus etc. Physically, the $P_{E_k,\ell+1}(r)$ wave function requires an energy similar to that of the height of the potential barrier $(V(r))$ in equation 2.18 to penetrate deeply into the atom[134] and give a large cross section. If this barrier is too wide then the cross section will only reach a maximum at high energy. This determines the position of the nodal point in energy.
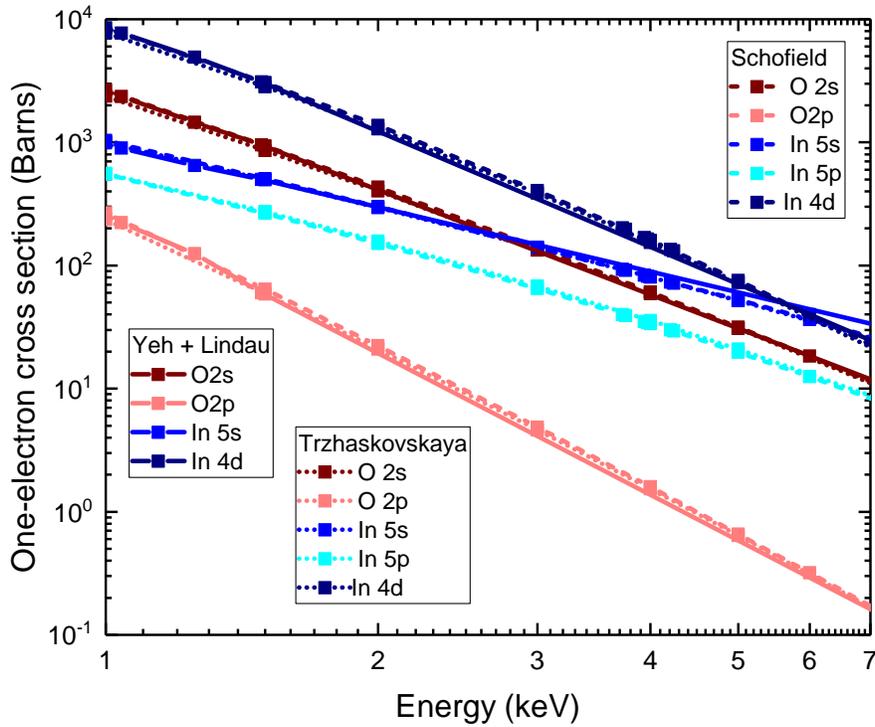
Figure 2.12: One-electron photoionization cross sections as a function of photon energy for the valence states in In_2O_3 using three different theoretical approaches[123,124,126].

The In 5s and O 2p orbitals that make up the conduction band of In_2O_3 behave in a very similar way to those of the Pb 6s and O 2p in PbO, except the In 5s wave radial function possesses one less anti-node that Pb 6s[135]. In this case the radial part of the wave function for the O 2p orbital is relatively contracted close to the nucleus, similar to the 1s peak in figure 2.10. The In 5s displays higher intensity further out

and has both positive and negative radial intensity. Similar to that discussed above, at low photon energies the O 2p orbital displays a higher cross section relative to the In 5s as the overlap with the low energy continuum orbital is greater. Since the In 5s orbital has higher intensity further from the atom, the opposite is true at higher energies, where the continuum wave function displays more rapid oscillations. This explains why the In 5s levels are made relatively stronger in hard x-ray photoemission spectroscopy (HAXPES) while O 2p levels become relatively more intense in standard XPS.

As previously mentioned, a great deal of time and effort has been put into calculating the theoretical photoionization cross sections for a large range of elements and photon energies, which have been published previously. Three examples are presented in figure 2.12 for the valence states in In_2O_3. The calculations of Yeh and Lindau as discussed above, are non-relativistic [124] while Schofield and Trzhaskovskaya include relativistic effects in their calculations of the central field [123,126]. The basic one-electron (corrected for the occupation of the orbital) cross sections for the valence states of In_2O_3 are presented in 2.12, and despite any difference in the calculation method the values of σ are essentially the same. It is worth noting however that Schofield only provides the basic cross sections (although over a very large energy scale), while Yeh and Lindau also calculate the asymmetry parameters using the dipole approximation. Trzhaskovskaya extends this even further by using the quadrupole approximation of the matrix element to calculate the non-dipole parameters. This is done by considering the differential cross section

$$\frac{d\sigma}{d\Omega} = \frac{4\pi^2 \alpha_0}{\omega} |M_{12}|^2. \tag{2.20}$$

where ω is the photon frequency. The differential cross section takes into account the angular distribution of emitted photoelectrons as well as the polarization of light
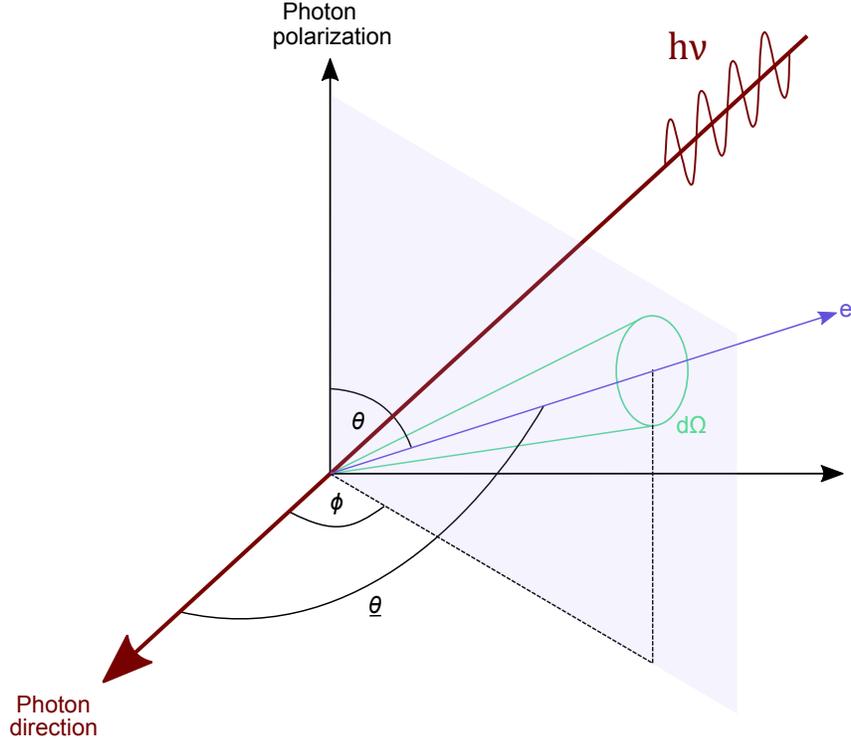
Figure 2.13: Schematic of the conventional geometry for asymmetry corrections to photoionization cross sections. The directions of the emitted electron (e^-), direction of photon propagation, and the photon polarization direction (for linearly polarized light) are labelled, along with the solid angle of electron collection ($d\Omega$), the angle between the polarization and electron emission direction (θ), the angle between the photon propagation direction and the (purple) plane containing the polarization and electron emission direction ϕ, and finally the angle between the photon propagation direction and the electron emission direction $\underline{\theta}$. Figure adapted from J. Cooper[132].

depending on the choice of $|M_{12}|^2$. The conventional designation for the geometry used for the photoemission process here is shown in figure 2.13. For unpolarised/circularly polarised light (e.g. a lab source) and linearly polarized light equation 2.20 becomes[136]

$$\frac{d\sigma}{d\Omega} = \frac{\sigma}{4\pi}\left[\underbrace{1 - \frac{\beta}{2}P_2(cos\underline{\theta})}_{dipolar} + \underbrace{\left(\frac{\gamma}{2}sin^2\underline{\theta} + \delta\right)cos\underline{\theta}}_{non-dipolar}\right] \qquad (2.21)$$

$$\frac{d\sigma}{d\Omega} = \frac{\sigma}{4\pi}\left[\underbrace{1 + \beta P_2(cos\theta)}_{dipolar} + \underbrace{\left(\gamma cos^2\theta + \delta\right)sin\theta cos\phi}_{non-dipolar}\right] \qquad (2.22)$$

where $P_2(x) = \frac{1}{2}(3x^2 - 1)$ is the second order Legendre polynomial and the angles and defined in figure 2.13. β is the dipole asymmetry parameter, and γ and δ are the non-dipolar parameters (associated with the electric quadrupole and magnetic dipole terms terms see Ref. [132,136] for more information). For lab based sources where unpolarised light is used (equation 2.21), the energy is relatively low and so the unpolarised terms can usually be ignored (see figure 2.14), returning us to the dipole approximation.
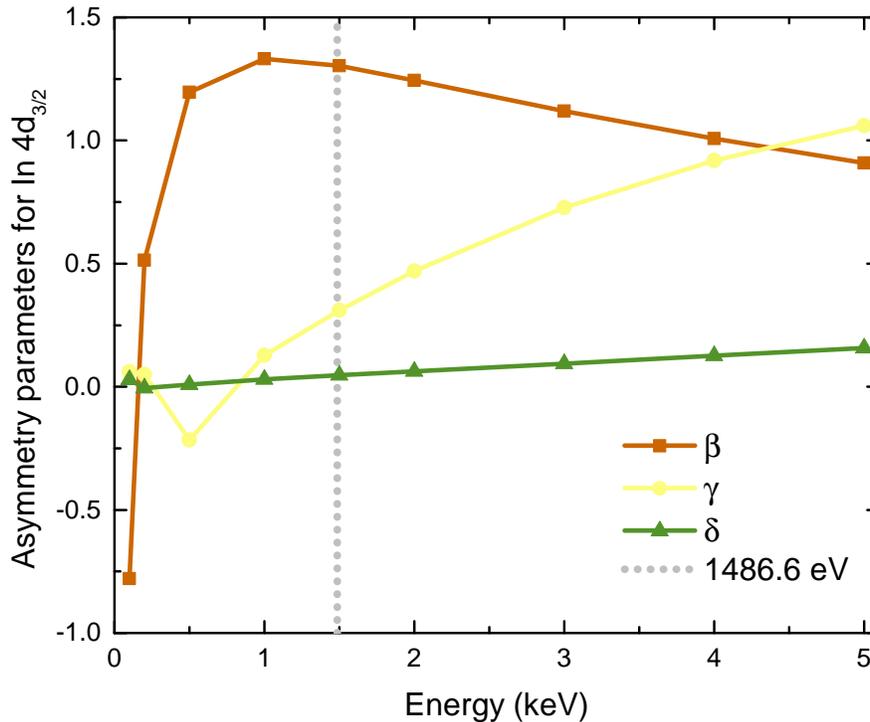
Figure 2.14: Asymmetry parameters for the In $4d_{3/2}$ level [126].

In addition to this it is common for the geometry of a lab based XPS system to give an angle between incident photon beam and electron emission vector of $\underline{\theta} = 54.7°$ known as the "magic angle", where the dipolar term goes to unity. Most lab based systems adopt this geometry, including the lab system used for this work.

It is possible now to show how the variation in photoionization cross-section as a function of energy seen in figure 2.12, impacts the valence band structure of In_2O_3. This is demonstrated in figure 2.15. Applying the cross-sections at both 1486.6 eV (soft
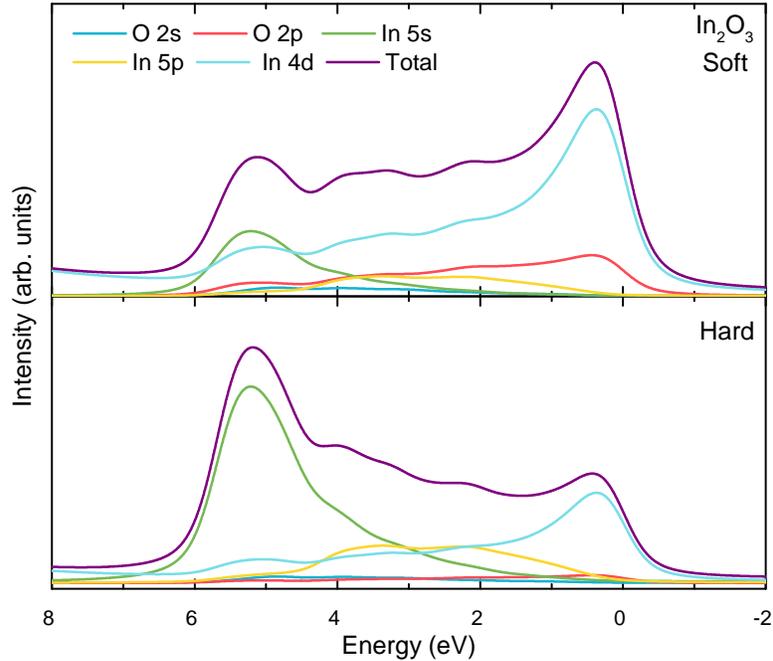
Figure 2.15: Photoionization cross-section corrected partial and total density of states for In_2O_3. Cross-sections were taken from Schofield[123].

x-rays) and 6000 eV (hard x-rays) calculated by Schofield[123] to the DFT calculated PDOS for In_2O_3, clear differences in spectral intensity can be seen. The In 5s level at ~ 5.5 eV in figure 2.15, is much stronger at hard energies relative to soft, while conversely the In 4d level seen at the VBM, becomes far less intense. This is because the In 4d cross section reduces much more quickly with energy compared to the In 5s, demonstrated by it's steeper gradient seen in figure 2.12. These differences have been demonstrated experimentally by Körber *et al.*[137]

2.1.5   SPECTRAL CALIBRATION AND RESOLUTION

One of the factors affecting the resolution of the measurement system in an PES experiment is the analyser that the system uses. For a hemispherical analyser, the resolution is determined by the physical size of both the analyser and the slits. Due to their finite sizes, there is a small range of kinetic energies of photoelectrons that can pass

through the analyser with different trajectories. The uncertainty in this gives rise to a contribution to the peak widths as measured in the spectrum. This width is given as[94]

$$\Delta E_A = E_P \left(\frac{w}{2R_c} + \Delta\alpha^2 \right) \tag{2.23}$$

where w is the average (exit) slit width of the analyser, $\Delta\alpha$ is the acceptance angle (or sometimes called the aperture half angle) of the analyser, and R_c is the mean radius of trajectory for an electron at a given pass energy E_P calculated as the average radius of the inner and outer hemispheres. Different kinetic energy electrons require different pass energies in order to be accepted, and so the resolution would vary over the spectrum if E_P was changed during the experiment. Therefore in order to maintain a constant resolution and also achieve good resolution with an analyser of reasonable size (R_c needs to be small enough in a practical sense), the analyser is held at constant E_P and the electrostatic lenses are used to retard electrons to the right energies to pass. Sweeping the retarding voltages across the range of energies required, generates a spectrum of constant energy resolution.

Equation 2.3 would suggest that in order to convert measured photoelectron kinetic energies into meaningful binding energy values, the work function of the analyser is required. However, as previously mentioned, it is possible to calibrate the measured energy scale using the alignment of the Fermi level of a metal in contact with the system. In practice this is commonly performed using silver or gold due to their relative inertness and ease of cleaning using Ar^+ ion bombardment. A metal has an occupied density of states right up to the Fermi level which when measured appears as a cutoff, past which the intensity is zero. For a metal, by definition this cutoff occurs at $E_B = 0$ eV, and so the Fermi edge of a metal is usually used for spectrometer calibration.

Figure 2.16 shows an experimentally measured Fermi-edge of an Ar^+ sputtered poly-crystalline silver metal foil. To determine the energy offset of the measured spectra, a
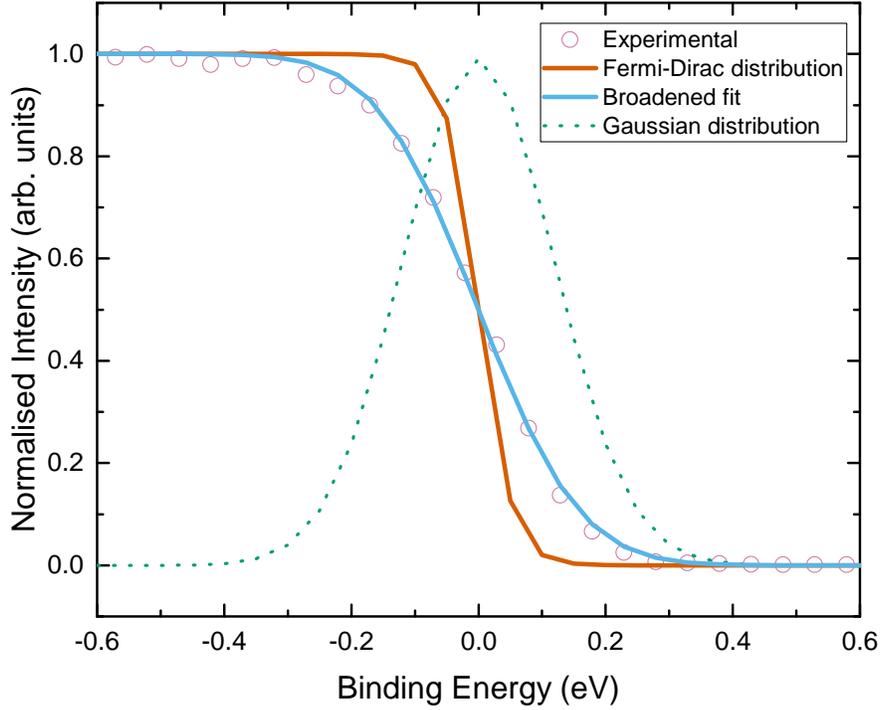
Figure 2.16: Fermi-edge cutoff of a cleaned silver spectrum centred at 0 eV, fitted with a broadened Fermi-Dirac distribution. This spectrum was measured at room temperature using a monochromated x-ray source.

Fermi-Dirac distribution (equation 2.24) is used to model the cutoff

$$F(E) = \frac{1}{e^{\frac{E-\mu}{K_B T}} + 1} \tag{2.24}$$

where k_B is the Boltzman constant, $T = 300$ K at room temperature, and μ is the total chemical potential which in this case $\sim E_F$ which is situated at 0 eV in binding energy. Equation 2.24 is seen as the brown curve in figure 2.16. Because instrumental broadening smears the intensity, this is modelled as a Gaussian function (seen as a green dotted line in figure 2.16) and is described by

$$F_g(E) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{E-\mu}{2\sigma^2}} \tag{2.25}$$

where σ is the standard deviation measuring the variation in the recorded energies.

The full width of the Gaussian is $\Gamma = 2\sqrt{2ln2}\sigma$ and μ can be thought of as the central energy value of the distribution. The two functions are convolved to express the effect of the broadening of the Gaussian upon the Fermi-Dirac distribution. Convolution is performed using

$$(F * F_g)(E) = \int F(\xi)F_g(E - \xi)d\xi \tag{2.26}$$

where ξ is a fixed energy position to allow for a sequential energy shift in E to drag one function over the other. This is plotted as a blue line in figure 2.24 which fits the data well. By extracting the fitting parameter σ it is possible to evaluate the FWHM and so the resolution of the instrument used. Note that the effects of Lorentzian broadening are neglected here. While the Gaussian distribution is usually used to model the distribution of spectral intensity due to the instrumental broadening, a Lorentzian line shape is associated with the broadening due to the inherent core-hole lifetime of an excited state. This contribution is ignored here as this adds extra parameters to the fit, meaning the determined Gaussian width maybe skewed. If the core-hole lifetime reflects the energy spread of the Lorentzian function through the Heisenberg uncertainty principle $\Delta E = \frac{\hbar}{2\Delta\tau}$, then if the lifetime of a state at E_F is sufficiently long this assumption holds. Due to screening and the fact that there are no states above E_F to drop into the core hole, this assumption could be correct, although it is extremely difficult to prove experimentally. Typically, the deeper the orbital, the shorter the core-hole lifetime[138]. Similarly, as atomic number increases, so does the width, as there are more valence electrons to fill the hole and so shorter lifetimes.

In addition, measuring the peak position and width of a strong core level of the metal standard allows for a two-point calibration to be performed, and gives an additional measure of resolution determination (assuming this has been performed using the spectrometer previously) and allows a check of the linearity of the spectrometer, i.e.

whether the binding energy scale has become stretched, resulting in incorrect binding energies being measured. Once the broadening of the Fermi edge is determined one is able to evaluate the overall peak width of a given core level (to first approximation) from

$$\Delta E = \sqrt{\sum_j \Delta E_j^2} \tag{2.27}$$

where ΔE_j represents different contributions to the overall width, e.g. ΔE_X is the natural line width of the excitation emission line, ΔE_A is the analyser resolution, and ΔE_ω represents phonon broadening (also referred to as Franck-Condon broadening), where electronic states can couple to vibrational modes which can provide a large contribution to the peak width (of the order of ~ 1 eV) in non-metallic oxide materials[139,140]. Additional contributions to equation 2.27 can be made by further summing in quadrature. Note that equation 2.27 applies only to the Gaussian components relevant to the resolution, while the widths of convoluted Lorentzians (e.g. the line width ΔE_{CL} of the core level being measured which is element, orbital and chemically dependent, mainly being due to the lifetime of the core-hole) sum in a linear fashion.

2.1.6 SPECTRAL ANALYSIS

To extract useful information from the raw spectra, some analysis is required. This does not necessarily involve applying a fit to the spectrum, but by applying synthetic line shapes and least squares fitting, it is possible to simulate the spectra and extract more information from the raw data.

BACKGROUNDS

Prior to applying any peak mode, it is necessary to model the background intensity. Any applied peak fit will then sit on top of this modelled background. There are nu-

merous ways to model the backgrounds, but some are more physical than others. It is worth pointing out that in reality, no background truly models the background perfectly which is probably indicated by the number of backgrounds available.
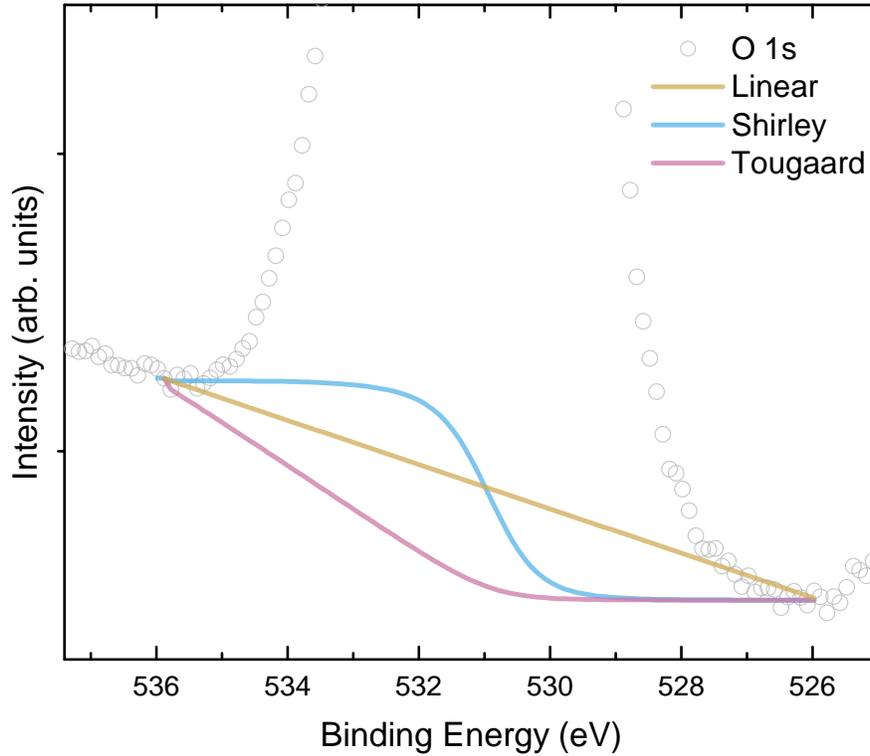
Figure 2.17: The Linear, Shirley and Tougaard backgrounds demonstrated for an O 1s spectrum.

The most basic is the simple linear background (see figure 2.17), which is a line connecting the two end points of the region specified. This is attractive due to its simplicity, however the background is made up of secondary electrons forming a continuum and so is not actually linear in shape. Instead, two common background types are usually adopted in core level fitting in XPS. The first is the Shirley background[92,141,142] which attempts to use spectral information from the data to inform changes in the simulated background.

$$B_{i+1} = k_s \sum_{i=0}^{i_{max}} (I_i - B_i).$$

(2.28)

It is an iterative method which splits the region of interest into many energy steps between the limits $E_{B,max}$ and $E_{B,min}$ defined by the analyst. The background intensity B_{i+1} is given by a fraction k_s (sometimes called the Shirley factor) of the intensity in the previous channel $(I_i - B_i)$. k_s is determined by the condition that at the region edges $I(E_{B,max}) = B(E_{E,max})$ and $I(E_{B,min}) = B(E_{E,min})$ (demonstrated in figure 2.17). The Shirley background is essentially a weighted average method making it computationally easy to implement, but lacks a full physical basis, and does not account for the fact that features have finite width and so often underestimates peak areas, especially towards lower binding energy.

$$B_i = \int_{E^k}^{E^k_{max}} \underbrace{\frac{B(E' - E)}{[C + (E' - E)^2]^2}}_{cross\ section\ term} \underbrace{I(E')}_{spectral\ intensity} \, dE' \tag{2.29}$$

The Tougard background[143] has a firm physical basis as it accounts for the initial energy distribution function and inelastic electron scattering. In equation 2.29, E^k_{max} is the high kinetic energy endpoint, where the background is equal to the measured intensity. B and C are simply fitting parameters ($B \sim 681.2$ eV² and $C \sim 355.0$ eV² were found to be appropriate for many elements). This background assumes an energy loss cross section for the probability an electron at some energy offset $E' - E$ will undergo a loss event and so appear as a contribution to the background.

The Tougaard background may be more physically realistic (a claim many would debate), but even so it has drawbacks. A large energy range is required on the high binding energy side to really achieve a reasonable fit. This is not always achievable. Most commonly the Shirley background is chosen to represent the inelastically scattered electrons under a peak due to its simplicity and ease of implementation.

CORE LEVEL LINE SHAPES

Because of the separate contributions from instrumental and core-hole lifetime broadening, the peak line shape used to fit a core level is most simply represented by a convolution between two line shapes. The Gaussian line shape is described in equation 2.25 while a Lorentzian is given as

$$F_L(E)\frac{1}{\pi}\frac{\frac{1}{2}\Gamma}{(E-\mu)^2+(\frac{1}{2}\Gamma)^2} \qquad (2.30)$$

where $\Gamma = \Delta E = \frac{\hbar}{\tau}$ is the width of the peak. Using equation 2.26 it is possible to combine these to give the convolution between a Gaussian and Lorentzian, usually called a Voigt function. A simple Gaussian, Lorentzian and Voigt line shape can be seen in figure 2.18. Convolution of many Gaussian and Lorentzian peaks can quickly become cumbersome in practice, and so most fitting software available opt for the simpler Gaussian-Lorentzian combination functions,[144] which are much less computationally expensive than the convolution procedure, and can provide accurate replication of the line shape given by a convolution. It is worth pointing out at this point that all of the photoemission spectral analysis done in this thesis was performed using the Casa XPS software, which adopts Gaussian-Lorentzian combination functions. An example of such a function (this particular function being the Gaussian/Lorentzian product form of the combination function, but other forms do exist, such as the sum form) is given by[145]

$$GL(x:F,E,m) = \frac{e^{\left((-4\ln2)(1-m)\frac{(x-E)^2}{F^2}\right)}}{1+4m\frac{(x-E)^2}{F^2}} \qquad (2.31)$$

where x is the energy scale, E represents the peak centre, F its full width at hall maximum, and m is the Gaussian-Lorentzian mixing ratio.

In some cases there is further need for greater flexibility in the model lineshapes,
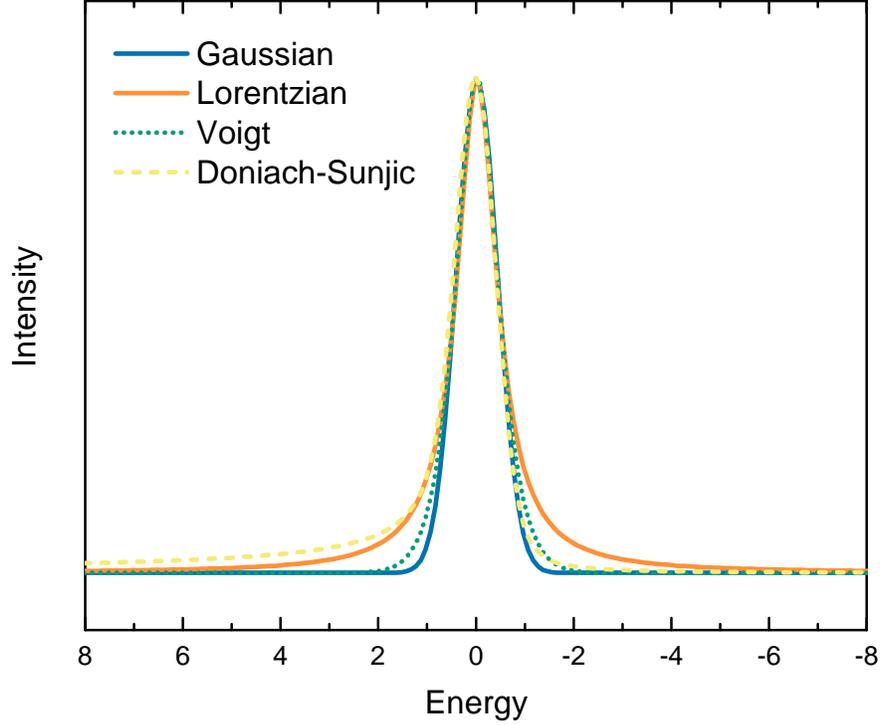
Figure 2.18: Examples of line shapes that can be used to fit peaks in PES.

which demonstrates that the nature of core-levels are often not simple, and may require other considerations.[94] One such example of this that is commonly encountered is the line shape for metallic chemical species. Conducting samples tend to display an asymmetric line shape towards the high binding energy due to the photoelectron interaction with free electrons (as discussed in section 2.1.3). This is often modelled using the Doniac-Sunjic line shape[146]

$$F_{DS}(E) = \frac{(1 - \alpha_{DS}) \cos\left[\frac{\pi\alpha}{2} + (1 - \alpha_{DS}) \tan^{-1}\left(\frac{E_b^0 - E}{\frac{\Delta E_0}{2}}\right)\right]}{\left[(E_b^0 - E)^2 + \left(\frac{\Delta E_0}{2}\right)^2\right]^{\frac{(1 - \alpha_{DS})}{2}}} \qquad (2.32)$$

where E_b^0 is the binding energy of the core level, ΔE_0 is the FWHM of the natural (no loss) line shape, and α_{DS} is the high binding energy asymmetry parameter ($0 \leq \alpha_{DS} \leq 0.5$ where $\alpha_{DS} = 0$ retains a Lorentzian-like lineshape). A Doniach-Sunjic

lineshape is displayed in figure 2.18 for reference. The extent of the tailing is related to the density of electrons and holes at the Fermi edge, as well as the probability of the many-body interaction occurring. As already stated, this line shape is most applicable to metals, where there is an obvious asymmetric tail to high binding energy. However, one issue with this line shape is its asymptotic form means it integrates to infinity. Therefore, the peak area that is extracted is very dependent on the defined energy region. This peak shape could be replicated by a series of Voigt functions or even greater freedoms are afforded by modified asymptotic peak models, all of which need careful consideration when analysing data. Another downside to equation 2.32 is it is hard to extract information regarding the loss feature. A TCO is an awkward material lying somewhere between a semiconductor and a metal, and the applicability of equation 2.32 appears not to be very representative of the core levels of TCOs. As to which line shape to use in which modelling situation, it likely requires a case by case evaluation and really will be down to the analyst's best judgement. It is entirely possible that the correct line shape to accurately describe and quantify the core levels and loss features of TCO materials has yet to be conceived. In this thesis, equation 2.31 is used in the all cases for spectral analysis.

2.2 Optical Measurements

Optical spectroscopies allow one to assess the optical performance of TCO materials. This is extremely important as a TCO is defined by its level of transparency, as well as its conductivity. Transmission of light in the IR, visible and UV spectrum is commonly used to evaluate this metric as well as provide information on the band gap of the material, while IR reflection can be used to obtain information about the properties of the free electrons in TCOs as the plasma frequency (ω_p) is usually in the IR region. In this section, I show how these measurements are implemented using two

simple experimental set-ups, Ultraviolet-visible-near infra-red spectroscopy (UV-Vis-NIR) and Fourier transform infra-red spectroscopy (FTIR). In theory, both of these measurements can be used to give the same information. However, in practice, the experimental equipment may have different source or detector ranges limiting what can be measured.

2.2.1 UV-Vis-IR Spectroscopy

The working principle of the UV-Vis-IR spectrophotometer is fairly simple, and generally the set-up of most modern spectrophotometers looks like the schematic in figure 2.19.
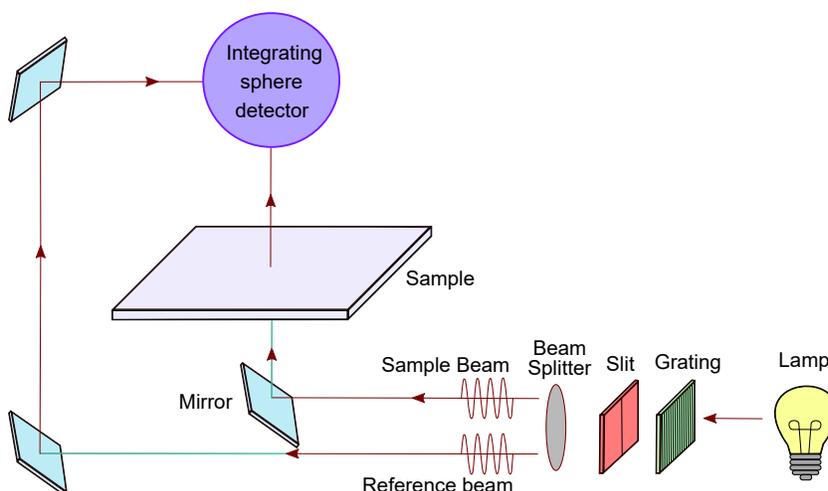
Figure 2.19: Simplified schematic of the typical experimental configuration used for optical measurements using a spectrophotometer.

A light source (in our case the option of either a deuterium lamp or halogen lamp could be chosen to cover a range of wavelengths from 240-~300 nm and from ~300 - 2600 nm respectively) is used to produce the full range of wavelengths of light required, which is directed to the monochromating diffraction grating. The white light is diffracted and dispersed into its constituent spectral parts and a filter slit is used to select the wavelength of light for each part of the measurement. A beam splitter (in

this case a chopper mirror is used) is often used to give a reference beam with 100% transmission, which the sample transmission signal is base-lined to. After passing through the sample, the beam is directed to the integrating sphere detector, which should allow for lower losses in the measured signal due to the diffuse samples. The integrating sphere used here was equipped with three detectors, a photomultiplier tube, an InGaAs detector and a PbS detector. This is simply to span the full spectral range available from the lamps. The measured transmission is given as a function of the wavelength of light.

2.2.2 FOURIER TRANSFORM INFRA-RED SPECTROSCOPY
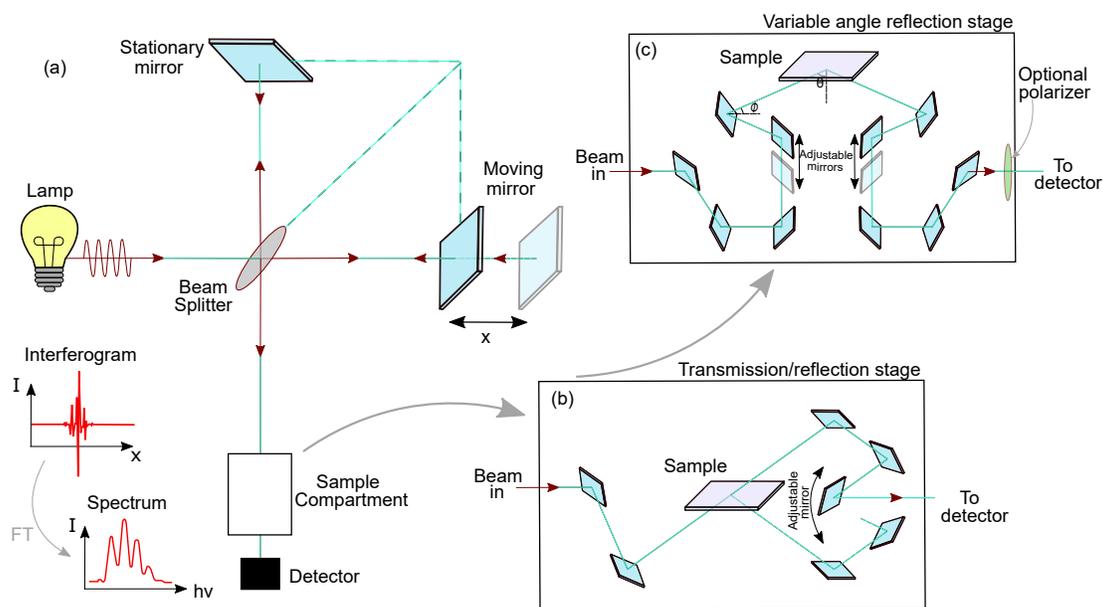
Figure 2.20: Simplified schematic of the typical experimental configuration used for optical measurements using an FTIR. (a) Beam path from lamp to to Michelson interferometer to sample compartment and finally to the detector. (b) One possible configuration in the sample compartment, using a moving mirror to collect both transmission and reflection data. (c) Another possible configuration for specular reflection using moving mirrors to vary the angle of the beam at the sample surface.

An FTIR measurement uses a broadband light source (usually alternating between a mid and near IR source is necessary, depending on the energy range required in the experiment), passing the light through a Michelson interferometer before reaching

the sample. The Michelson interferometer consists of a beam splitter (often a half-silvered mirror) which reflects 50% of the light towards a fixed mirror and transmits 50% towards a moving mirror. The light from the two separate paths then recombines back at the beam splitter, before being directed towards the sample compartment and subsequent detector.

When the moving mirror is at its equilibrium position, the path between it and the beam splitter is equal to that of the fixed mirror to the beam splitter. After moving a distance x, the path between the moving mirror and the beam splitter becomes longer, giving a path difference between the two of $2x$. This path difference gives a phase difference of $2kx$ between the two beams, where $k = \frac{2\pi}{\lambda}$ is the wavevector. When the path lengths are equal the phase difference is equal to zero, all of the light is directed towards the sample, while for a phase difference of π destructive interference occurs and no light is directed to the sample compartment due to wave interference. As the phase difference varies with the motion of the mirror, an interferogram (plot of interference) is generated which is measured at the detector as an intensity versus mirror position.

A spectrum is then generated from the Fourier transform of this signal (the spatial domain of the mirror x is converted into a wavenumber domain with units matching that of $\frac{1}{x}$ which is easily related to energy). Because an FTIR samples all the frequencies from the source at once, it is able to measure much more quickly and achieve a much greater flux per measurement than standard UV-Vis instruments (this is known as the Fellgett's advantage). This means better signal-to-noise can be achieved and much faster scans performed. Also, the position of the mirrors can be determined very accurately using a laser and very good resolution can be achieved by increasing the maximum optical path difference (OPD). The resolution is equal to the reciprocal of the maximum OPD. However, great care must be taken when obtaining data with

an FTIR, as not to introduce additional spectral features. FTIR is a single beam measurement and so sample spectra need to be referenced to the experimental set-up with the sample removed, or to a perfect mirror for reflection measurements to account for the effects that the instrumental set-up (e.g. any energy dependence of the source, beam splitter or variation in detector performance), or the atmosphere in the FTIR etc. have on the spectrum. Of course in reality a perfect mirror does not exist, and so most often either aluminium, gold, or silver mirrors are utilized in practice, which display extremely high reflectance (>90%) for much of the infrared spectrum. Additionally, appropriate choice of beam splitter and detector must be made to cover the spectral range of interest.

2.2.3 ABSORPTION COEFFICIENT FOR DEGENERATE SEMICONDUCTORS

In this section, the basic theory for calculating the absorption coefficient for direct transitions in semiconductor materials is introduced, and a simple method based on the linear extrapolation of an α^2 vs. $\hbar\omega$ plot is discussed. Here the specific case relevant to degenerately doped semiconductors is discussed, and an easy-to-implement analysis is introduced, based on the works of Hamberg *et al.*[5,33] and Dolgonos *et al.*[147], is presented. This model is utilized in chapter 5 in for optical analysis.

One large assumption made in this analysis is that the bands have parabolic character. This was suggested by Hamberg *et al.* in their analysis for Sn-doping of In_2O_3[33] which displays conduction band non-parabolicity. However, assuming parabolic bands simplifies the expressions enough to allow for simple analysis to be performed, i.e. no complicated modelling need be applied which is necessary for the adoption of an analytical method by the majority in the field. Therefore, we will also assume parabolic bands here, acknowledging that the physical case is not best represented this way.

Starting from the transition rate for optical transitions between initial $|i\rangle$ and final
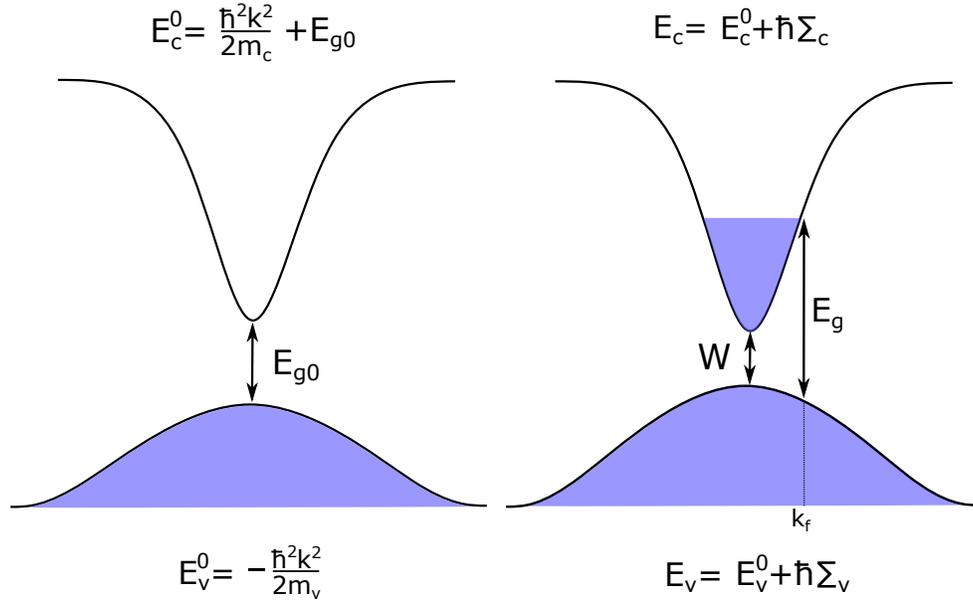
Figure 2.21: Illustration of the band filling and renormalisation process in a doped semiconductor. Left: Direct transition between parabolic bands. Right: Transition from VB to lowest unoccupied CB state for degenerate system with parabolic bands after renormalistion effects. Figure adapted from Hamberg *et al.* [33]

$|f\rangle$ states we can write [33]

$$R = \frac{2}{\hbar^2} \sum_{i,f} |\langle i|V|f\rangle|^2 \frac{\tau^{-1}}{(\omega - \omega_{fi})^2 + \tau^{-2}} (P_i - P_f). \tag{2.33}$$

This expression is essentially the same as in equation D.11 but a broadening term (given by τ) has been added to account for the broadening of the initial and final states, and a occupation probability P of each of the initial and final states is also included, which is important for a degenerate system, and $\omega_{fi} = \frac{E_f - E_i}{\hbar}$. In the limit of $\tau \to \infty$ equation 2.33 goes to the usual Fermi's golden rule expression. The valence band is expected to be filled while the conduction band is only partially filled in a degenerate system. Using an approach similar to that discussed in appendix D, it can be shown that this transition rate can be given as

$$R \propto \int_{x_0}^{\infty} (x + \hbar\omega - W)^{\frac{1}{2}} \frac{\Gamma}{x^2 + \Gamma^2} (1 - P_c) \, dx \qquad (2.34)$$

where $x = \frac{\hbar^2 k^2}{2m_{vc}^*} + W - \hbar\omega$, while $x_0 = \Delta E_g^{BM} + W = \hbar\omega$ (with $\Delta E_g^{BM} = \frac{\hbar^2}{2m_{vc}^*} (3\pi^2 n_e)^{\frac{2}{3}}$),

and $\Gamma = \frac{\hbar}{\tau}$. The minimum energy separation between the valence and conduction band in the doped material is given as

$$W \approx E_{g0} + \hbar\Sigma_c(k_f) - \hbar\Sigma_v(k_f) \qquad (2.35)$$

which is shown in figure 2.21. The term P_c can be represented by a Fermi-Dirac distribution at k_F

$$P_c = \frac{1}{e^{\frac{E_c^0 - \mu}{K_B T}} + 1} \qquad (2.36)$$

where μ is the chemical potential.

To obtain an expression for the absorption coefficient, it is necessary to acknowledge that $\alpha \propto R$ (see appendix D). Hamberg *et al.* proposed that, if Γ is small, the Lorentzian part of equation 2.34 is a sharp peaked function, and therefore the square-root expression is considered as slowly varying and taken outside of the integral. This simplifies to[33]

$$\alpha \propto (\Delta E_g^{BM})^{\frac{1}{2}} \int_{x_0}^{\infty} \frac{\Gamma}{x^2 + \Gamma^2} (1 - P_c) \, dx$$
$$\propto \left(1 - \frac{2}{\pi} \tan^{-1}\left(\frac{W + \Delta E_g^{BM} - \hbar\omega}{\Gamma}\right)\right) \qquad (2.37)$$

With a series expansion of equation 2.37, and replacing $W + (\Delta E_g^{BM})^{\frac{1}{2}}$ with E_g, ignoring higher order terms, we see

$$\alpha \propto 1 - \frac{2}{\pi\Gamma}(E_g - \hbar\omega) \qquad (2.38)$$

Equation 2.38 tells us that, contrary to what is commonly done for degenerate materials, where a linear extrapolation of α (usually to some power depending on whether the transition is predicted to be direct or not) against $\hbar\omega$ is expected to give an estimate of E_g, it in fact gives $E_g - \frac{\pi\Gamma}{2}$. Performing another series expansion of α^2 in equation 2.37 we see

$$\alpha^2 \propto 1 - \frac{4}{\pi\Gamma}(E_g - \hbar\omega) \tag{2.39}$$

is determined and so the linearly extracted value of α^2 with $\hbar\omega$ is $E_g - \frac{\pi}{4\Gamma}$. Extrapolation from both α and α^2 can be combined to find the value of E_g in the method shown in figure 2.22 for a degenerate semiconductor.
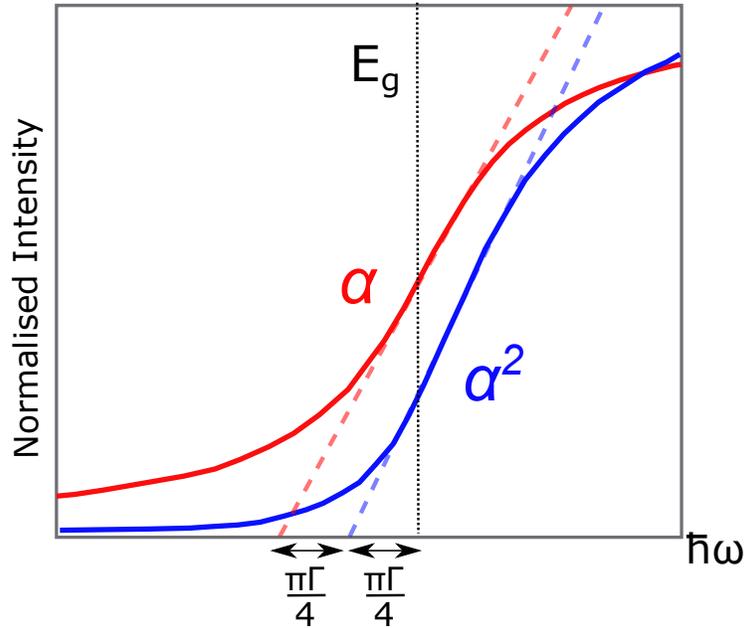
Figure 2.22: Schematic of the difference between the two extrapolation methods for α and α^2 to $\hbar\omega$. The difference between the two intercepts can be calculated and added to the α^2 intercept to give an estimate of E_g. Figure replicated from Dolgonos *et al.* [147]

This method of determining the band gap from optical data of a degenerate system is easy to implement and based on good physical backing. However, many simplifica-

tions are taken in the expansion of the above equations to finally arrive at equations 2.38 and 2.39 (such as ignoring all higher order terms in the expansions). Therefore, it seems possible this approach can obtain an estimate for the optical gap of a degenerate semiconductor which is more accurate than the Tauc method in the case of degenerate materials. However, it is unlikely to perfectly match the actual optical gap due to the assumptions made and information disposed of in the simplifications. Hence, in reality both the Tauc method and the one presented here are both flawed for different reasons and so either can be used with this knowledge in mind when applied to degenerate systems.

2.2.4 TRANSFER MATRIX METHOD FOR OPTICAL MODELLING

When investigating the optical properties of multilayer stacks it is often most convenient to use the transfer matrix method to determine optical coefficients, rather than the long hand ray tracing method. Commercial products incorporating TCO materials most often consist of at least two layers (a TCO deposited on glass for example), and so the ray tracing method quickly becomes cumbersome and inefficient. In this thesis, the transfer matrix method formalism set out by Katsidis and Siapkas is used.[148] This method is described schematically in figure 2.23.

The superscript on the electric field amplitudes $E^{+/-}$ denotes the direction of travel of the electric field (the prime symbol E' denotes a wave directly on the right of an interface), n_i denotes a medium with a new dielectric constant, and the numbers underneath the interfaces represents the interfaces between two medium. In figure 2.23 there is an arbitrary number of layers N which is specific to the stack being investigated.

The amplitude of the electric field on the left $E_{m-1}^{+/-}$ and right $E_m^{+/-}$ of an interface can be related by the dynamical matrix that describes both reflection and transmission at
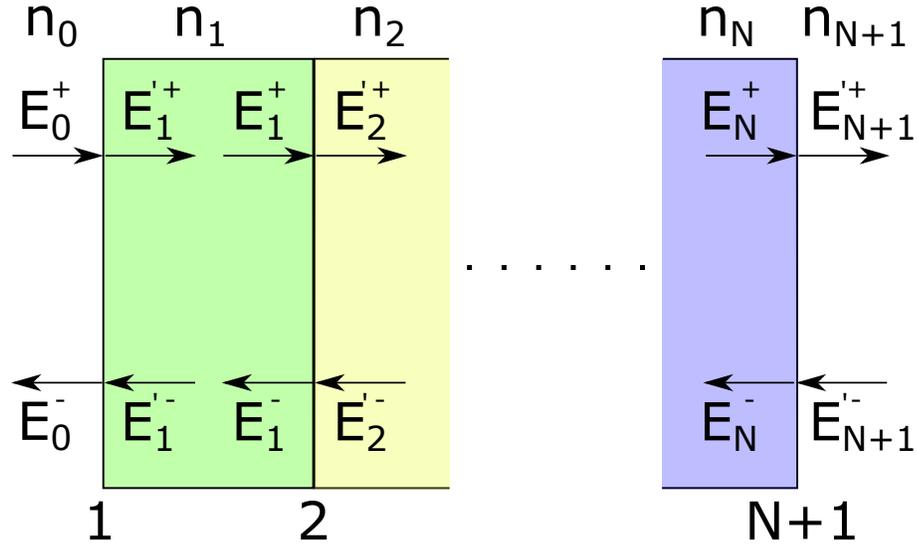
Figure 2.23: Schematic multilayer structure with arbitrary N number of layers and $N+1$ interfaces. The corresponding field amplitudes are shown.

the interface. If the waves reflected and transmitted from a boundary between material $m-1$ and m can be described as a linear combination of two plane waves

$$E_m^+ = t_{m-1,m}E_{m-1}^+ + r_{m,m-1}E_m^- \qquad \text{and} \qquad E_{m-1}^+ = r_{m-1,m}E_{m-1}^+ + t_{m,m-1}E_m^- \qquad (2.40)$$

here $r_{m-1,m}$ and $t_{m-1,m}$ are the Fresnel equations that can be written for both s and p polarised light [149]

$$r_{m-1,m}^s = \frac{n_{m-1}\cos\theta_{m-1} - n_m\cos\theta_m}{n_{m-1}\cos\theta_{m-1} + n_m\cos\theta_m} \qquad (2.41)$$

$$t_{m-1,m}^s = \frac{2n_{m-1}\cos\theta_{m-1}}{n_{m-1}\cos\theta_{m-1} + n_m\cos\theta_m} \qquad (2.42)$$

$$r_{m-1,m}^p = \frac{n_m\cos\theta_{m-1} - n_{m-1}\cos\theta_m}{n_m\cos\theta_{m-1} + n_{m-1}\cos\theta_m} \qquad (2.43)$$

$$t^p_{m-1,m} = \frac{2n_{m-1}\cos\theta_{m-1}}{n_m\cos\theta_{m-1} + n_{m-1}\cos\theta_m}. \tag{2.44}$$

θ_i is the angle of incidence/transmission in medium $i = m - 1, m$. Now solving equation 2.40 leads us to

$$\begin{pmatrix} E^+_{m-1} \\ E^-_{m-1} \end{pmatrix} = \frac{1}{t_{m-1,m}} \begin{bmatrix} 1 & -r_{m,m-1} \\ r_{m-1,m} & t_{m-1,m}t_{m,m-1} - r_{m-1,m}r_{m,m-1} \end{bmatrix} \begin{pmatrix} E^+_m \\ E^-_m \end{pmatrix} \tag{2.45}$$

This equation can be simplified using Stokes relations[†]

$$\begin{pmatrix} E^+_{m-1} \\ E^-_{m-1} \end{pmatrix} = \frac{1}{t_{m-1,m}} \begin{bmatrix} 1 & r_{m-1,m} \\ r_{m-1,m} & 1 \end{bmatrix} \begin{pmatrix} E'^+_m \\ E'^-_m \end{pmatrix} = D_m^{-1}D_{m+1} \begin{pmatrix} E'^+_m \\ E'^-_m \end{pmatrix}. \tag{2.46}$$

The effect of the bulk on the field amplitudes of the wave are accounted for in the propagation matrix

$$\begin{pmatrix} E'^+_m \\ E'^-_m \end{pmatrix} = \begin{bmatrix} e^{i\delta_m} & 0 \\ 0 & e^{-i\delta_m} \end{bmatrix} \begin{pmatrix} E^+_m \\ E^-_m \end{pmatrix} = P_m \begin{pmatrix} E^+_m \\ E^-_m \end{pmatrix}. \tag{2.47}$$

where δ_m is the phase change of light travelling through the medium m of thickness d_m, given by Snell's law

$$\delta_m = \frac{2\pi}{\lambda} n_m d_m \cos\theta_m. \tag{2.48}$$

For a multi-layered system, repeated application of equations 2.46 and 2.47 for N layers and $N+1$ interfaces gives the product of N+1 2×2 matrices, and allows for the

---

[†]Stokes relations are derived using time-reversal arguments and are given as $rr + tt' = 1$ so $tt' = 1 - r^2$ and $tr' + rt = 0$ so $r = -r'$.

determination of the total transfer matrix T

$$
\begin{aligned}
\begin{pmatrix} E_0^+ \\ E_0^- \end{pmatrix} &= T \begin{pmatrix} E_{N+1}'^+ \\ E_{N+1}'^- \end{pmatrix} = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} \begin{pmatrix} E_{N+1}'^+ \\ E_{N+1}'^- \end{pmatrix} \\
&= D_0^{-1} \left[\prod_{m-1}^{N} D_m P_m D_m^{-1} \right] D_{N+1} \begin{pmatrix} E_{N+1}'^+ \\ E_{N+1}'^- \end{pmatrix}
\end{aligned}
\tag{2.49}
$$

The total (complex) reflection and transmission coefficients of the multi-layered system are given by the terms in the transfer matrix elements T_{ij} as

$$
r = r_{0,N+1} = \left. \frac{E_0^-}{E_0^+} \right|_{E_{N+1}'^-=0} = \frac{T_{21}}{T_{11}}
\tag{2.50}
$$

$$
t = t_{0,N+1} = \left. \frac{E_{N+1}'^+}{E_0^+} \right|_{E_{N+1}'^-=0} = \frac{1}{T_{11}}
\tag{2.51}
$$

$$
r' = r_{N+1,0} = \left. \frac{E_{N+1}^-}{E_{N+1}^+} \right|_{E_0^+=0} = -\frac{T_{12}}{T_{11}}
\tag{2.52}
$$

$$
t' = t_{N+1,0} = \left. \frac{E_0^-}{E_{N+1}^+} \right|_{E_0^+=0} = \frac{\det T}{T_{11}}
\tag{2.53}
$$

where $\det T = T_{11} T_{22} - T_{12} T_{21}$. The total front and back reflectance and transmittance can be determined from the square of the magnitudes of r, r', t, and t' respectively. For unpolarised light the total reflectivity is the averahe of the incoherent polarised reflection coefficients

$$
R = \frac{1}{2} (R^s + R^p) = \frac{1}{2} (|r^s|^2 + |r^p|^2).
\tag{2.54}
$$

If the multi-layer structure contains a layer with a rough interface it is relevant

to include partial coherences into the transfer matrix model. This is done through modifying the Fresnel coefficients of the rough interface. The modification represents the phase difference in the reflected and transmitted beams. The modified coefficients for a rough m^{th} layer take the form of

$$r_{m-1,m} = r_{m-1,m}^{(0)} e^{-2(2\pi Z n_{m-1}k)^2} \tag{2.55}$$

$$r_{m,m-1} = r_{m,m-1}^{(0)} e^{-2(2\pi Z n_m k)^2} \tag{2.56}$$

$$t_{m-1,m} = t_{m-1,m}^{(0)} e^{-\frac{1}{2}(2\pi Z k)^2 (n_m - n_{m-1})^2} \tag{2.57}$$

$$t_{m,m-1} = t_{m,m-1}^{(0)} e^{-\frac{1}{2}(2\pi Z k)^2 (n_{m-1} - n_m)^2} \tag{2.58}$$

where (0) represents the smooth interface Fresnel coefficient, k is the wavenumber, and Z is the root-mean-square height of the roughness.

Finally, in the case of complete incoherence, such that would be found for very thick layers (e.g. a thick glass substrate) where light can reflect from the sides of the substrate, we can treat the incoherent layer similarly to before, with a modification to the transfer matrix. If there is incoherence in layer m, one layer in an N layer system, the transfer matrix up to this layer and beyond this layer is

$$
\begin{aligned}
T_{0,m} &= \frac{1}{t_{0,m}}
\begin{bmatrix}
1 & -r_{0,m} \\
r_{0,m} & t_{0,m}t_{m,0} - r_{0,m}r_{m,0}
\end{bmatrix} \\[2ex]
T_{m,N+1} &= \frac{1}{t_{m,N+1}}
\begin{bmatrix}
1 & -r_{m,N+1} \\
r_{m,N+1} & t_{m,N+1}t_{N+1,m} - r_{m,N+1}r_{N+1,m}
\end{bmatrix} .
\end{aligned}
\tag{2.59}
$$

These are then converted to intensity matrices by squaring

$$
T_{0,m}^{int} = \frac{1}{|t_{0,m}|^2}
\begin{bmatrix}
1 & -|r_{0,m}|^2 \\
|r_{0,m}|^2 & |t_{0,m}t_{m,0}|^2 - |r_{0,m}r_{m,0}|^2
\end{bmatrix}
$$

$$
T_{m,N+1}^{int} = \frac{1}{|t_{m,N+1}|^2}
\begin{bmatrix}
1 & -|r_{m,N+1}|^2 \\
|r_{m,N+1}|^2 & |t_{m,N+1}t_{N+1,m}|^2 - |r_{m,N+1}r_{N+1,m}|^2
\end{bmatrix}.
\tag{2.60}
$$

The total transfer matrix is given by

$$
T^{int} = T_{0,m}^{int} P_m^{int} T_{m,N+1}^{int}
\tag{2.61}
$$

where

$$
P_m^{int} =
\begin{bmatrix}
|e^{i\delta_i}|^2 & 0 \\
0 & |e^{-i\delta_i}|^2
\end{bmatrix}.
\tag{2.62}
$$

is the intensity propagation matrix of the incoherent layer. Essentially this treats the coherent parts surrounding the incoherent layer as separate, and then includes the intensity of the propagation matrix directly. Splitting the matrices this way allows for the representation of multiple reflections inside a thick substrate. The reflectivity in this case is given as $R = \frac{T_{21}^{int}}{T_{11}^{int}}$.

2.3 Hall Effect

Hall effect measurements are used for determining carrier densities and electrical properties for a material. Therefore this technique is extremely useful in the field of TCO materials, where the electronic properties often determine the commercial relevance of the material.

A schematic of the experimental set-up of a Hall effect measurement of a simple conducting bar is shown in figure 2.24. In a Hall experiment, an electric field $\vec{E_x}$ is applied to the material in the x-direction, causing a current to flow j_x. A magnetic field $\vec{B_z}$ is applied in the z-direction resulting in the Lorentz force. As electrons build up in the $-y$-direction against the side of the sample, an electric field $\vec{E_y}$ builds in the y-direction to oppose the electrons motion and try to stop further accumulation of charge.

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \tag{2.63}$$
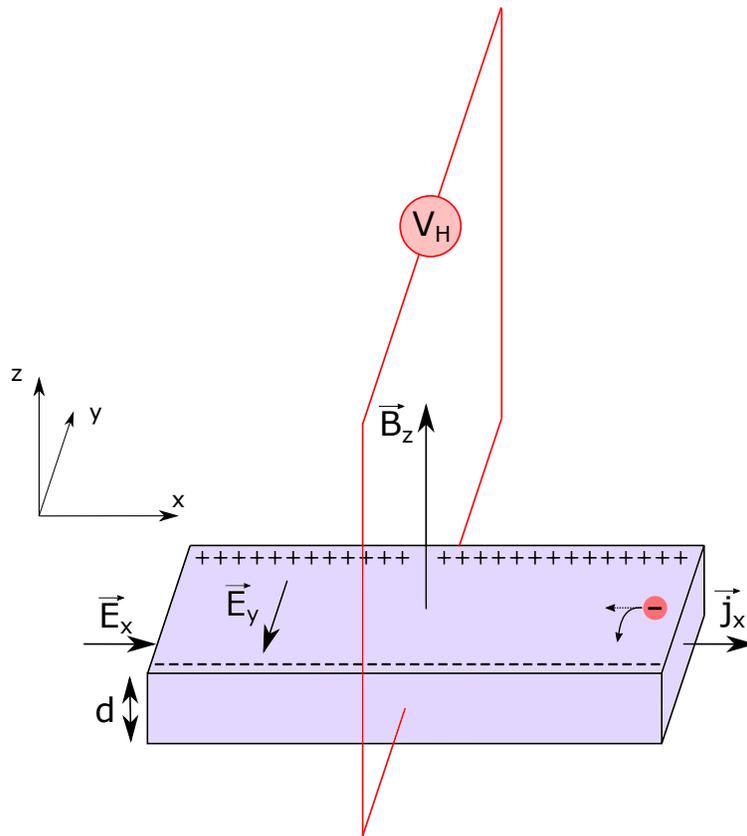
Figure 2.24: Illustration of the Hall effect. $\vec{E_i}$ and $\vec{B_i}$ represent the electric and magnetic field vectors in the direction of $i = x, y, z$, the current density is given by $\vec{j_i}$, the hall voltage by V_H, and d represents the sample thickness.

where q is the charge on the particle and $\mathbf{v}$ is the particle velocity. The Lorentz force

acts to deflect electrons in the direction of $-y$. Hence, in equilibrium, the $\vec{E_y}$ field balances the Lorentz force, and electrons will flow in the x-direction only. To extract valuable information from this, the motion of an electron in a magnetic field needs to be considered.

The equation of motion for an electron in a magnetic field (electrons are chosen here as the most relevant case to TCOs) being acted upon also by friction, which is represented by collisions at a rate $\frac{1}{\tau}$, is given by Newton's equations as[47,150]

$$m\left(\frac{d}{dt} + \frac{1}{\tau}\right)\mathbf{v} = -e\left(\mathbf{E} + \mathbf{v} \times \mathbf{B}\right) \tag{2.64}$$

where m is the effective mass of the carrier, and τ its scattering rate. If the magnetic field is in the z-direction then equation 2.64 becomes

$$
\begin{aligned}
m\left(\frac{d}{dt} + \frac{1}{\tau}\right)\mathbf{v_x} &= -e\left(\mathbf{E_x} + B\mathbf{v_y}\right) \\
m\left(\frac{d}{dt} + \frac{1}{\tau}\right)\mathbf{v_y} &= -e\left(\mathbf{E_y} - B\mathbf{v_x}\right) \\
m\left(\frac{d}{dt} + \frac{1}{\tau}\right)\mathbf{v_z} &= -e\mathbf{E_z}.
\end{aligned}
\tag{2.65}
$$

In the steady-state in a static electric field the time derivatives $\frac{d\mathbf{v}}{dt}$ is equal to zero and so we can write

$$\mathbf{v_x} = -\frac{e\tau}{m}\mathbf{E_x} - \omega_c\tau\mathbf{v_y}; \quad \mathbf{v_y} = -\frac{e\tau}{m}\mathbf{E_y} + \omega_c\tau\mathbf{v_x}; \quad \mathbf{v_z} = -\frac{e\tau}{m}\mathbf{E_z} \tag{2.66}$$

where $\omega_c = \frac{eB}{m}$ is the cyclotron frequency. Hence, equation 2.66 can be written as

$$\mathbf{v_y}(1 + \omega_c^2\tau^2) = -\frac{e\tau}{m}\left(\mathbf{E_y} - \omega_c\tau\mathbf{E_x}\right). \tag{2.67}$$

94

Now in a Hall measurements as in figure 2.24, the Hall field $\vec{E}_y$ is in the direction $\mathbf{j} \times \mathbf{B}$ where current $\mathbf{j}$ is flowing from left to right in our diagram. If the current cannot flow out of the material in the y-direction then a scenario where $\delta\mathbf{v_y} = 0$ is achieved. Therefore, equation 2.67 becomes

$$\mathbf{E_y} = -\omega_c \tau \mathbf{E_x} = -\frac{e\mathbf{B}\tau}{m}\mathbf{E_x}. \tag{2.68}$$

The Hall coefficient is given by

$$R_H = \frac{\mathbf{E_y}}{\mathbf{j_x B}} \tag{2.69}$$

where $\mathbf{j_x}$ is the current density given by $\mathbf{j_x} = ne\mathbf{v_x} = ne\mu\mathbf{E_x} = \frac{ne^2\tau\mathbf{E_x}}{m}$. Equation 2.69 can thus be evaluated using equation 2.68 as

$$R_H = -\frac{1}{ne}. \tag{2.70}$$

Note that in a similar way, the Hall coefficient for either a p-type semiconductor, or one where contributions form holes and electrons are important, can be given as

$$R_H = \frac{1}{pe}; \qquad R_H = \frac{p\mu_h^2 - n\mu_e^2}{e(p\mu_h + n\mu_e)^2} \tag{2.71}$$

where p is the concentration of holes and μ are the carrier mobilities.

Electrical (conducting) probes are required to measure the electronic properties of a material, which are attached to the surface of the sample. This means that, in reality, the measurable quantities in a Hall effect measurement are the current and the Hall voltage, rather than the current density and the Hall field. Therefore, an areal carrier density is actually determined from the measurement. This is then converted to a volume density through division of the film/sample thickness.

Some useful quantities are given from $\mathbf{j} = ne\mathbf{v} = \frac{ne^2\tau\mathbf{E}}{m}$ such as the conductivity (σ) and the resistivity (ρ). Using this expression it is found that $\mathbf{j} = \frac{ne^2\tau\mathbf{E}}{m} = \sigma\mathbf{E} = \frac{\mathbf{E}}{\rho}$. Therefore, it is seen that $\sigma = \frac{1}{\rho} = \frac{ne^2\tau}{m}$ and if the mobility, which is simply the ratio of the drift velocity to an external electric field, is given by $\mu = \frac{e\tau}{m}$ we can write

$$\sigma = n\mu e \qquad \rho = \frac{1}{n\mu e}. \tag{2.72}$$

These quantities are frequently used when describing electrical conduction in TCO materials. Note also that in the above it is assumed that all electrons have the same properties. However, the scattering probabilities are not really energy independent and electrons will have a distribution of energies. This means R_H is not exactly equal to $-\frac{1}{ne}$. This has the impact of the mobility obtained from a Hall measurement, the Hall mobility μ_H has to be distinguished from the actual drift mobility $\mu = \frac{e\tau}{m}$ defined above. They are related by a quantity known as the Hall scattering factor $\mu = \frac{\mu_H}{r_H}$ where $r_H = \frac{\langle\tau^2\rangle}{\langle\tau\rangle^2}$ and is usually $1 \leq r_H \leq 2$. In the same way $n = r_H n_H$ and $R_H = -\frac{r_H}{ne}$.

Fortunately, for TCOs as well as other degenerate materials (where $\tau = \tau(E_F)$), it is generally seen that $r_H \approx 1$ at high carrier density, as shown for In_2O_3 in figure 2.25. Because r_H is dependent on the scattering time τ, each scattering mechanism will in theory have an independent r_H parameter. The scattering times are combined using Mathiessen's rule

$$\frac{1}{\tau} = \frac{1}{\tau_{impurities}} + \frac{1}{\tau_{lattice}} + \frac{1}{\tau_{defects}} + \cdots \tag{2.73}$$

and since ionized impurity scattering is usually the dominant scattering mechanism in degenerate materials at high carrier concentrations, and its r_H is ≈ 1 for TCOs[152] this explains why we can neglect this term when analysing Hall data for TCOs.
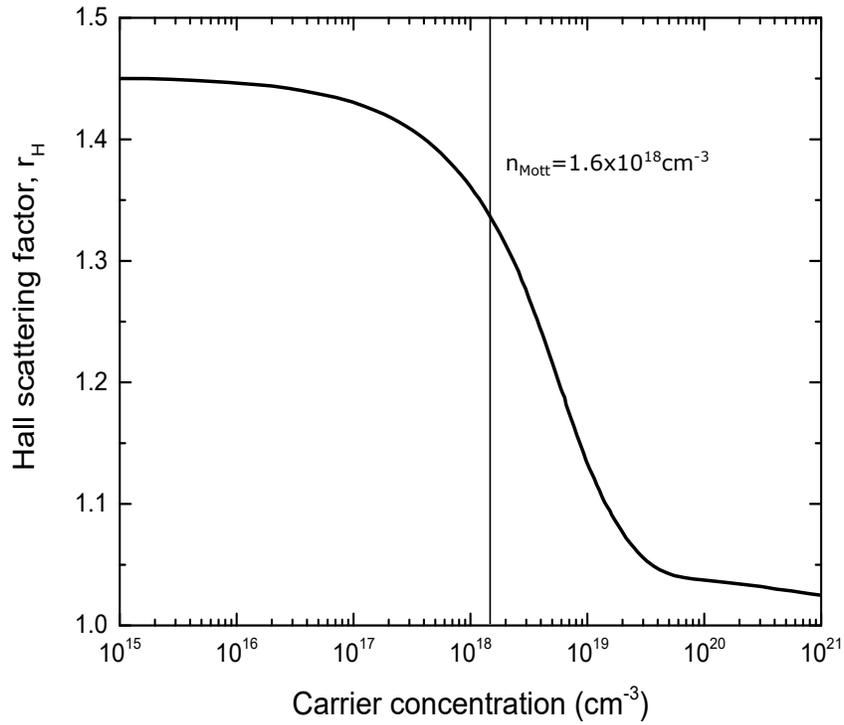
Figure 2.25: Hall scattering factor for In_2O_3 for singly charged donors. Figure replicated from Preissler *et al.* [151]

2.3.1 VAN DER PAUW GEOMETRY

Whilst the Hall bar configuration shown in figure 2.24 is perfect for illustrating the Hall effect simply, and is the simplest way of measuring important electrical properties, it is not the preferred geometry for the measurement. Instead, we opt for what is known as the van der Pauw geometry, named after its founder L. J. van der Pauw. [153,154] This measurement geometry is popular as it allows the measurement of the electrical properties of a solid of arbitrary shape, so long as the solid is much thinner than it is wide, and is homogeneous, isotropic, possesses no holes, and the electrodes placed at the perimeter connected as point contacts on the surface (at least an order of magnitude between the dimensions of the sample and contact).

Figure 2.26 shows the typical van der Pauw geometry for electrical measurements. A current I_{12} can be passed between contacts 1 and 2 and a voltage V_{34} measured across
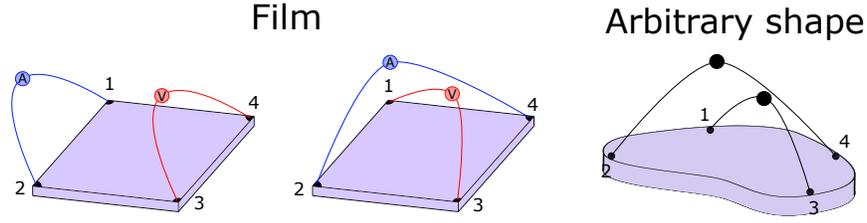
Figure 2.26: Schematic of the van der Pauw geometry for Hall effect measurements. Left: typical set-up for measuring films. Right: a set-up for materials with arbitrary dimensions.

contacts 3 and 4. The resistance can be calculated as $R_{12,34} = \frac{V_{34}}{I_{12}}$. If what is measured at the contacts is then switched, another resistance can be determined according to the equation

$$R_{ij,kl} = \frac{V_{kl}}{I_{ij}} \tag{2.74}$$

where the current enters contact i and leaves j, and the voltage is measured as $V_{kl} = V_k - V_l$. Van der Pauw showed using geometrical arguments that that the sheet resistance ($R_s = \frac{\rho}{t}$ which is a measure of the resistance along (parallel to) a two-dimensional film, not to be confused with resistance through (perpendicular to) a film $R = \frac{\rho}{t}\frac{L}{W}$ where t, L and W are the thickness, length and width of the film) relates to the these measured resistances in equation 2.74 by

$$e^{-\frac{\pi R_{12,34}}{R_s}} + e^{-\frac{\pi R_{23,41}}{R_s}} = 1. \tag{2.75}$$

In addition, the reciprocity theorem gives [153,154]

$$R_{ij,kl} = R_{kl,ij} \tag{2.76}$$

meaning many quantities can be measured and averaged just by switching the contacts to give a more accurate measure of the sheet resistance, helping minimize the effects of offset voltages such as those due to magnetoresistance. Further improvement is made

by switching the polarities of the current source and voltage meter. The measured resistances become

$$
\begin{aligned}
R_{vertical} &= \frac{R_{12,34} + R_{34,12} + R_{21,43} + R_{43,21}}{4} \\
R_{horizontal} &= \frac{R_{23,41} + R_{41,23} + R_{32,14} + R_{14,32}}{4}.
\end{aligned}
\tag{2.77}
$$

R_s can still be determined from equation 2.75 replacing the resistances with those in equation 2.77.

In a similar way, the Hall voltage can be determined more accurately using the reciprocity theorem. If, for example, a current is applied between contacts on opposite corners, e.g. I_{13}, and the magnetic field polarity is flipped and the measurement repeated, the Hall voltage measured will be $V_{24}^{\pm}$ where $\pm$ denotes the magnetic field polarity. The average Hall voltage is then

$$
V_H = \frac{V_{13}^+ + V_{13}^- + V_{24}^+ + V_{24}^- + V_{31}^+ + V_{31}^- + V_{42}^+ + V_{42}^-}{8}.
\tag{2.78}
$$

By combining the sets of currents and voltages, the Hall voltage and sheet resistance can be determined as well as the Hall sheet carrier density $n_s = \frac{IB}{e|V_H|}$ and Hall mobility $\mu_H = \frac{1}{en_s R_s}$ in terms of measurable quantities.[‡]

2.3.2 ELECTRICAL TRANSPORT

The final part of this section is concerned with scattering theory and electrical transport. This is essentially an approach to understanding the results of transport measurements, i.e. why majority carriers behave the way they do under the application of an

---

[‡]Note that if the velocity of a carrier is $v = -\frac{I}{nAe}$ then in the steady-state $E = vB = -\frac{IB}{nAe}$. If the Hall voltage is simply the electric field multiplied by the width of the material $V_H = \int_0^W E dl = EW$ then $V_H = -\frac{IBW}{enA} = -\frac{IB}{ent} = -\frac{IB}{en_s}$, therefore $n_s = \frac{IB}{e|V_H|}$. Finally, since the sheet resistance is defined as $R_s = \frac{\rho}{t}$ and $\rho = \frac{1}{ne\mu}$ we obtain the majority carrier mobility $\mu_H = \frac{1}{en_s R_s} = \frac{|V_H|}{IBR_s} = \frac{|V_H|t}{IB\rho}$.

external field to the material. It is most useful to be able to compare the measured mobilities to a model based on the band structure, and parameters such as the dielectric constant, as then our understanding of the material is informed by the measurement. Amongst the many difficulties in transport modelling, determining the important scattering mechanisms involved in carrier transport and deciding the appropriate ways to model them may be paramount. This is emphasized in figure 2.27 displaying a number of scattering mechanisms relevant to semiconductors. Electronic transport and scattering has a very long history of works aimed at accurately describing processes, both classically and quantum mechanically. Since equations of a closed form are often desirable for their ease of applications, great effort has been paid to providing simple equations describing the transport relationships for a given scattering mechanism. One example of this, which is very important for degenerate TCOs is ionized impurity scattering, and carrying on from the formalism set out in appendix E, the transport equation for this mechanism will be provided as an example. This is used in chapters 4 and 5.
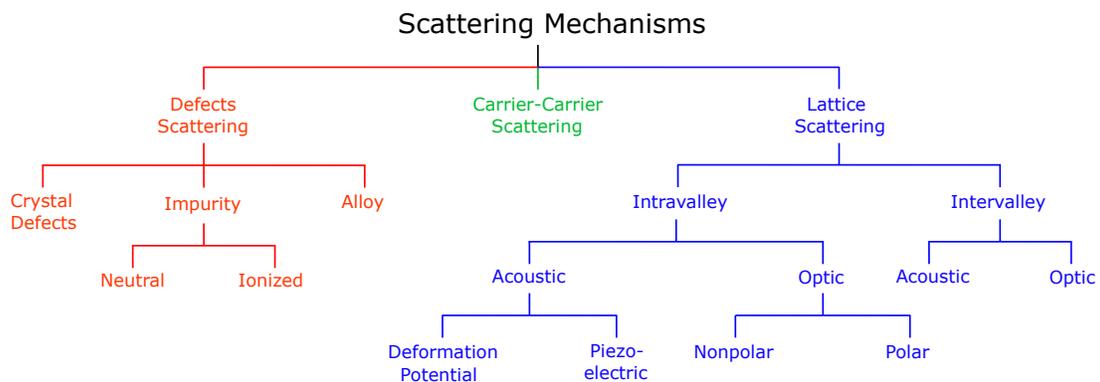
Figure 2.27: Various carrier scattering mechanisms possible in semiconductors. Figure reproduced from B. R. Nag[155].

As previously stated, one difficulty in calculated transport coefficients is deciding which scattering mechanisms are most important, and which can be neglected. It is necessary to neglect some mechanisms simply due to the shear quantity of scattering

types, some of which are shown in figure 2.27. For example, in heavily degenerate systems ($n > 10^{18}$ cm⁻³), the effects of electron-electron scattering can be ignored according to Bate *et al.* due to screening effects[156,157]. Arguably the most important scattering mechanism relevant to TCO materials is that due to ionized impurities. This is because at the high carrier densities ($n > 10^{20}$cm⁻³) present in doped TCOs, there is significant interaction between free carriers and ionized dopant atoms. Hence, this mechanism will be used as an example to get a better understanding of how electrical transport data is modelled.

The discussion begins with a general solution to the Boltzmann equation for carrier scattering by defects. The scattering rate under the relaxation-time approximation $\left(\frac{df}{dt}\right)_{collision} = \frac{f(\mathbf{k})-f_0(\mathbf{k})}{\tau}$, (where $f(\mathbf{k})$ is the distribution function describing the occupation of electrons with equilibrium distribution function $f_0(\mathbf{k})$, and τ is the relaxation time, a similar formula is given in equation E.28), is given by

$$\frac{1}{\tau_{ii}} = \frac{V_c}{(2\pi)^3} \int_{\mathbf{k}} \frac{2\pi}{\hbar} |M(\mathbf{k}, \mathbf{k}')|^2 (1 - cos\theta) \delta(E_{\mathbf{k}} - E_{\mathbf{k}'}) d\mathbf{k}' \tag{2.79}$$

where $M(\mathbf{k}, \mathbf{k}')$ is the matrix element described by

$$M(\mathbf{k}, \mathbf{k}') = \int_{\Omega} \phi_{\mathbf{k}'}^*(\mathbf{r}) \Delta V \phi_{\mathbf{k}}(\mathbf{r}) d\mathbf{r}. \tag{2.80}$$

Here the integration is performed over the unit cell volume Ω, and $\phi(\mathbf{r})$ may be described as a Bloch function with wave vector $\mathbf{k}$ and $\mathbf{k}'$ before and after the scattering event respectively. V_c is the crystal volume, $\frac{1}{(2\pi)^2}$ is the density of states in $\mathbf{k}'$ space, θ is the angle between $\mathbf{k}$ and $\mathbf{k}'$, and $\delta(E_{\mathbf{k}} - E_{\mathbf{k}'})$ is the Dirac delta function ensuring energy conservation. For Ionized impurities scattering, electrons are scattered by their interaction with the screened coulomb potential of the impurity. In this case we may write our perturbing potential as

$$\Delta V = \frac{Ze^2}{4\pi\varepsilon r}e^{-\frac{r}{\lambda_D}} \tag{2.81}$$

where Ze is the ionic charge, ε the dielectric permittivity, r is a point in space, and λ_D is the Debye screening length which essentially sets the length scale that the potential drops off at. The screened potential comes about from the electrons responding to an applied electric field, effectively screening the potential. Without the screening length included, the potential has an infinite distance range. Other approaches to deal with this issue include the predecessor by Conwell and Weisskopf[158] who essentially use a cut-off distance (the average seperation between two atoms) much in the same thinking as with Rutherford scattering, rather than a smoothly transitioning one[159].

The matrix element from equation 2.80

$$M(\mathbf{k}, \mathbf{k}') = \frac{1}{V_c}\int_\Omega \phi_{\mathbf{k}'}^*(\mathbf{r})\frac{Ze^2}{4\pi\varepsilon r}e^{-\frac{r}{\lambda_D}}\phi_{\mathbf{k}}(\mathbf{r})d\mathbf{r}. \tag{2.82}$$

Solving this using the orthogonality of the Bloch wave functions (see appendix D, equation D.8) gives[155,160,161]

$$\begin{aligned} M(\mathbf{k}, \mathbf{k}') &= \frac{1}{V_c}\int_\Omega u_{\mathbf{k}'}^*(\mathbf{r})e^{-i(\mathbf{k}'\cdot\mathbf{r})}\frac{Ze^2}{4\pi\varepsilon r}e^{-\frac{r}{\lambda_D}}u_{\mathbf{k}}^*(\mathbf{r})e^{i(\mathbf{k}\cdot\mathbf{r})}d\mathbf{r} \\ &= \frac{Ze^2}{\varepsilon V_c}\frac{1}{|\mathbf{k}'-\mathbf{k}|^2 + \lambda_D^{-2}}. \end{aligned} \tag{2.83}$$

The conservation of momentum for an electron scattering event is pictured in figure 2.28. Because we have the delta function in equation 2.79 this would imply conservation of energy and so $\mathbf{k} = \mathbf{k}'$.

This leads us to the relation

$$|\mathbf{k}' - \mathbf{k}|^2 = 2k^2 - 2k^2\cos\theta. \tag{2.84}$$
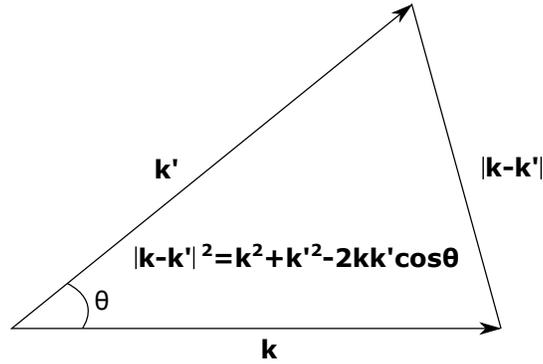
Figure 2.28: Schematic of the conservation of momentum in a scattering event.

Plugging equations 2.84 and 2.83 into equation 2.79 we arrive at

$$
\begin{aligned}
\frac{1}{\tau_{ii}} &= \frac{2\pi}{\hbar} \frac{Z^2 e^4}{\varepsilon^2 V_c} \frac{1}{(2\pi)^3} \int \frac{1 - cos\theta}{\left(2k^2(1 - cos\theta) + \lambda_D^{-2}\right)^2} \delta(E_{\mathbf{k}} - E_{\mathbf{k}'}) d\mathbf{k}' \\
&= \frac{2\pi}{\hbar} \frac{Z^2 e^4}{\varepsilon^2 V_c} \frac{1}{(2\pi)^3} \int \frac{1 - cos\theta}{\left(2k^2(1 - cos\theta) + \lambda_D^{-2}\right)^2} \delta(E_{\mathbf{k}} - E_{\mathbf{k}'}) k^2 dk' sin\theta d\theta d\varphi.
\end{aligned}
\tag{2.85}
$$

The total potential requires the number of impurities N_I over the volume of the crytal giving $\Delta V_{tot} = N_I V_c \Delta V$ and so equation 2.85 becomes

$$
\frac{1}{\tau_{ii}} = \frac{2\pi}{\hbar} \frac{Z^2 e^4}{\varepsilon^2} \frac{N_I}{(2\pi)^3} \int \frac{1 - cos\theta}{\left(2k^2(1 - cos\theta) + \lambda_D^{-2}\right)^2} \delta(E_{\mathbf{k}} - E_{\mathbf{k}'}) k^2 dk' sin\theta d\theta d\varphi. \tag{2.86}
$$

integration over φ gives a factor of 2π and we can apply a simple parabolic band model (the mathematics becomes more involved for non-parabolic bands but this has been previously performed, see for example Zawadzki *et al.*[162,163]), where $dk' = \frac{m^*}{\hbar^2 k} dE'$ to give

$$
\frac{1}{\tau_{ii}} = \frac{Z^2 e^4}{\hbar^3 \varepsilon^2} \frac{N_I}{8\pi} \frac{m^*}{k^3} \int_0^\pi \frac{(1 - cos\theta)}{\left((1 - cos\theta) + \frac{1}{2k^2 \lambda_D^2}\right)^2} sin\theta d\theta. \tag{2.87}
$$

This integration can be solved by substitution to give

$$\frac{1}{\tau_{ii}} = \frac{Z^2 e^4}{\hbar^3 \varepsilon^2} \frac{N_I}{8\pi} \frac{m^*}{k^3} \left[\ln(1 + 4k^2\lambda_D^2) - \frac{4k^2\lambda_D^2}{1 + 4k^2\lambda_D^2} \right] \tag{2.88}$$

which can be given as

$$\frac{1}{\tau_{ii}} = \frac{Z^2 e^4}{\hbar^3 \varepsilon^2} \frac{N_I}{8\pi} \frac{m^*}{k^3} \left[\ln(1 + b) - \frac{b}{1 + b} \right] = a_{ii}^{-1} k^{-3} \left[\ln(1 + b) - \frac{b}{1 + b} \right] \tag{2.89}$$

Equation 2.89 is formally known as the Brooks-Herring-Dingle model for ionized impurity scattering[164,165]. The mobility associated with this model is given by

$$\mu_{ii} = \frac{e\langle\tau\rangle}{m^*} = \frac{e}{m^*} \frac{\int_0^\infty \tau_{ii} E^{\frac{3}{2}} \frac{df_0}{dE} dE}{\int_0^\infty E^{\frac{3}{2}} \frac{df_0}{dE} dE} \tag{2.90}$$

where $f_0 = \frac{1}{e^{\frac{E - \mu_F}{K_B T}} + 1}$ is the Fermi-Dirac distribution, and μ_F can be equated with E_F. In the case of non-degenerate statistics, $\frac{df_0}{dE} \propto -\frac{1}{K_B T} e^{-\frac{E}{K_B T}}$, which leads to the classic Brooks-Herring-Dingle model for ionized impurity limited mobility, which is heavily temperature dependent. However, for degenerate carrier statistics we see that $\frac{df_0}{dE}$ can logically be treated as a delta function $\frac{df_0}{dE} \approx \delta(E - E_F)$ and so equation 2.90 simply collapses down to the usual expression for mobility

$$\mu_{ii} = \frac{e\tau_{ii}}{m^*} = \frac{ea_{ii}k^3}{m^*} \left[\frac{1}{\ln(1 + b) - \frac{b}{1+b}} \right]. \tag{2.91}$$

Again, simplifying for a parabolic band the wave vector at the Fermi level is given by

$$k^2 \approx k_F^2 \approx \frac{2m^* E_F}{\hbar^2} \approx (3\pi^2 n)^{\frac{2}{3}}. \tag{2.92}$$

Inserting this into equation 2.89 and 2.91, we arrive at

$$\mu_{ii} = \frac{24\pi^3 \varepsilon^2 \hbar^3 n}{N_I Z^2 e^3 m^{*2} \left[\frac{1}{ln(1+b) - \frac{b}{1+b}} \right]}. \tag{2.93}$$

Now all that is required to have a closed-form equation for ionized impurity scattering in degenerate semiconductors, is an equation for the screening length λ_D. In the case of non-degeneracy this is simply the Debye screening length $\frac{e^2 n}{\varepsilon K_B T}^{-\frac{1}{2}}$, but for degenerate statistics, the Thomas-Fermi screening length is used, which for a parabolic density of states is given as

$$\lambda_D = \frac{2}{3} \frac{\varepsilon E_F}{e^2 n} = \frac{\varepsilon \pi^{\frac{4}{3}} \hbar^2}{(3n)^{\frac{1}{3}} e^2 m^*} \tag{2.94}$$

giving a screening parameter of $b = \frac{4(3\pi^8 n)^{\frac{1}{3}} \varepsilon \hbar^2}{e^2 m^*}$ to be used in equation 2.93. A very similar expression for ionized impurity limited mobility was given for a non-parabolic density of states (α-approximated band structure in the Kane formalism) by Zawadzk[162,163], and was later simplified by Pisarkiewicz[44]. The main difference is $\frac{dE}{dk}$ is retained in the expression.

In the above, the focus has been on the scattering of electrons due to ionized impurities as an example of how analytical expressions can be derived for the mobility limits. However, there are many scattering mechanisms relevant to semiconductors and a wealth of information is available in the literature on these subjects. The main challenges are to derive an expression for the scattering probability appropriate to the mechanism under investigation (which chiefly involves finding a form for the perturbing potential), and relating this expression in the collision term of the Boltzmann transport equation to the scattering time (through an approximation such as the relaxation-time approximation).

2.4 Thin Film Growth

A number of materials were investigated during the course of this thesis. The thin films deposited on glass investigated in chapters 4 and 5 were deposited by some variant of the chemical vapour deposition (CVD) process, while the bulk crystals in chapter 6 were grown using the edge-defined film-fed growth method[166]. The materials in chapter 4 and 6 were donated and purchased from NSG Group and Novel Crystal Technology, Inc. respectively, and so were not directly grown as a result of the work presented here. Hence, no more information regarding the deposition of these materials will be presented here, further information should be obtained from the manufacturer. However, the materials in the study presented in chapter 5 were designed and deposited specifically for this project by S. Sathasivam in the group of I. Parkin at University College London. Therefore, the deposition of these films will be discussed in some more details.

2.4.1 CHEMICAL VAPOUR DEPOSITION

CVD is a technique which is used for depositing gas phase precursors onto a solid to produce a thin solid film. This is different to physical vapour deposition (PVD) methods such as sputtering and pulsed laser deposition, as CVD uses volatile precursors which react or decompose onto a substrate, while PVD relies on converting a source material into its gas phase and subsequently condensing this onto a substrate. There are advantages and disadvantages associated with both CVD and PVD methods, but for large scale deposition relevant to many TCO applications, the cost effectiveness (not requiring high vacuum) of CVD is extremely attractive.

CVD takes place in some sort of reactor, a typical design is shown schematically in figure 2.29. This simple schematic shows many of the important components required
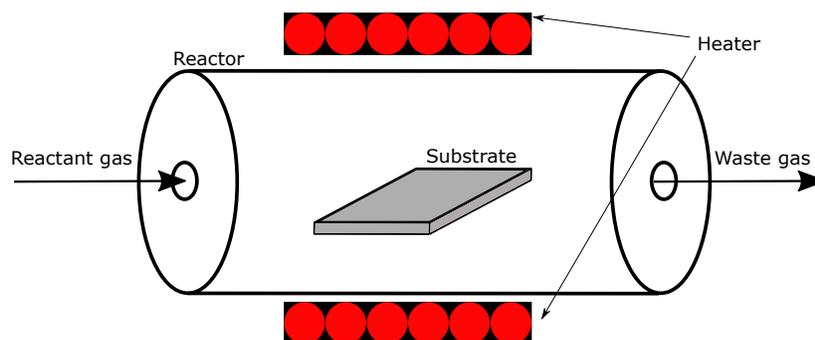
Figure 2.29: Schematic of a typical CVD reactor.

by a CVD reactor, including gas inlet and outlet to allow reactant gas into the chamber and waste gases to escape, a quartz tube to house the substrate and gasses which is both chemically inert and has a high melting point, and some form of heating element to allow for the reaction to take place.

The conventional CVD process can be broken down into a few steps shown schematically in figure 2.30: i) Transportation of precursor molecules to substrate, which is often done using an inert carrier gas such as nitrogen. ii) The gaseous precursors undergo chemical reactions either on the surface of the substrate, or in the gas phase itself, facilitated by some form of environmental catalyst e.g. heat. iii) Physisorption of precursors onto the substrate surface is followed by decomposition of the molecules, allowing for the formation of the desirable material and by-product waste material which is removed from the chamber. iv) Adatoms diffuse on the substrate surface to find the lowest energy sites. These then act as nucleation sites, which promote film growth[167].

There are many variations on the conventional CVD process, one of which was adopted in chapter 5 to deposit films of In_2O_3 doped with Mo and Sn. The process used is called aerosol-assisted chemical vapour deposition (AACVD), which uses the same basic principle discussed above with some important differences. Unlike during conventional CVD, precursors are first dissolved in solvent and an aerosol is formed
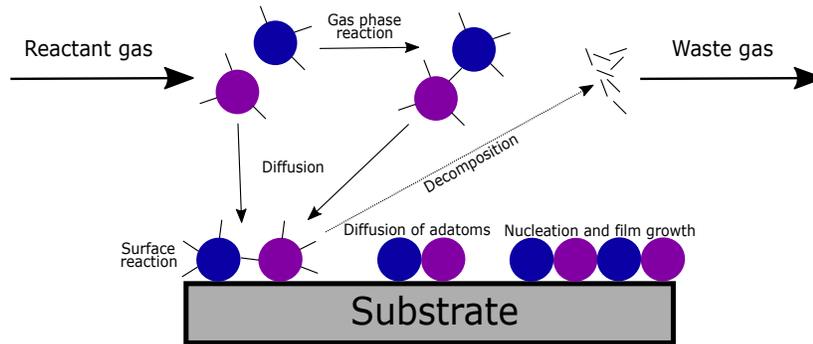
Figure 2.30: Schematic of the conventional CVD process. Figure replicated from Marchand *et al.* [167]
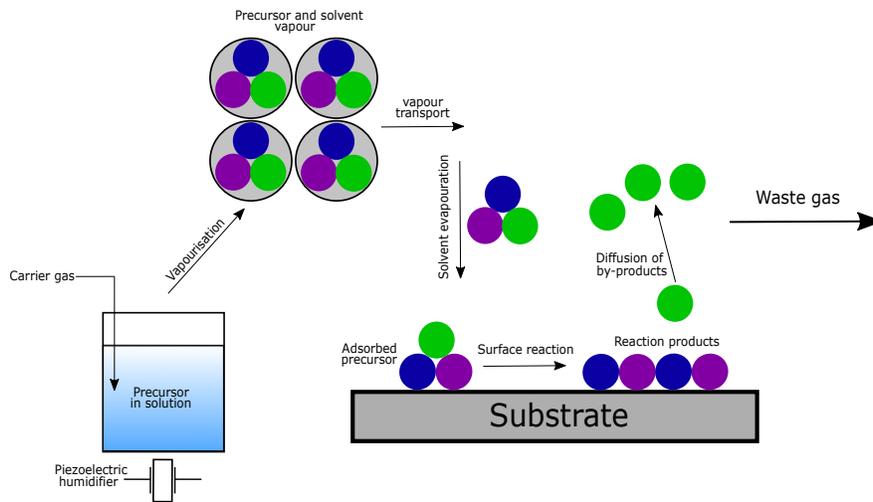
Figure 2.31: Schematic of the AACVD process. Figure replicated from Marchand *et al.* [167]

with use of a humidifier (see figure 2.31). This aerosol is transported into the CVD reactor using a carrier gas, often N_2 is used. The solvent undergoes evaporation when the the aerosol reaches the heated substrate, leaving the precursor in its gas phase. Finally, the precursor gas vapour can take two routes to form the final film. It can decompose in the gas phase, forming an intermediate product which adsorbs to the substrate surface. This undergoes further decomposition and heterogeneous reaction to form the desired film. Alternatively, decomposition and chemical reaction can occur predominantly in the gas phase (usually if the substrate is too hot), which leads to homogeneous nucleation of particles. However, the absorption of these preformed

particles will often lead to higher levels of film porosity. AACVD has some benefits over conventional CVD methods, including a high mass transport rate meaning higher deposition rates, and AACVD does not require volatile precursors meaning there is far greater choice in the available precursor options.

3

Theoretical Background

TCOs display non-parabolic conduction bands and electronic degeneracy and so expressions for the band structure should reflect this. We also rely heavily on the insights gained from density functional theory, where calculations can be complicated by large complex unit cells and dopants breaking perfect translational symmetry. Furthermore, surface electronic space charge properties can be understood by solving the Poisson equation within the surface layer of the material. All of the above will be discussed in this chapter, with the emphasis on the underlying fundamentals of the theory, with occasional insight into its application to the field of TCOs.

3.1   Electronic structure calculations

3.1.1   BAND STRUCTURE APPROXIMATIONS

The electronic band structure can be approximated from a solution of the Schrödinger equation, with the Hamiltonians[26,161] :

$$H_{1e}\Phi_n(\mathbf{r}) = \left(\frac{p^2}{2m} + V(\mathbf{r})\right)\Phi_n(\mathbf{r}) = E_n\Phi_n(\mathbf{r}) \tag{3.1}$$

where H_{1e} is the one-electron Hamiltonian, $\Phi_n(\mathbf{r})$ is the wave function of an electron in eigenstate n, p is the momentum operator for the electron, $V(\mathbf{r})$ is the periodic potential of the Bravais lattice, and E_n is the energy of that electron. Equation 3.1 is a simplified version of the full many-bodied atomic Hamiltonian. Simplifying assumptions include grouping non-valence electrons with the ion cores, the Born-Oppenheimer approximation (ions are essentially spatially fixed as they are much heavier than electrons), the one-electron approximation (where the electron-electron interaction is averaged providing a constant repulsive potential to the Hamiltonian), and the mean-field approximation (which assumes every electron experiences the same average potential $V(\mathbf{r})$). It can be shown using the translational symmetry of the lattice, that for a crystal with periodic potential $V(\mathbf{r}) = V(\mathbf{r} + \mathbf{R})$, the eigenstates to equation 3.1 are Bloch functions[26,47,161]

$$\Phi_{n\mathbf{k}}(\mathbf{r}) = u_{n\mathbf{k}}(\mathbf{r})e^{i\mathbf{k}\cdot\mathbf{r}} \tag{3.2}$$

where n labels the band, $\mathbf{k}$ the wave vector of the electron in the first Brillouin zone, and $u_{n\mathbf{k}}(\mathbf{r})$ is a function with the same periodicity as the crystal lattice and so the full translational symmetry

$$u_{n\mathbf{k}}(\mathbf{r} + \mathbf{T}) = u_{n\mathbf{k}}(\mathbf{r}) \tag{3.3}$$

where $\mathbf{T}$ is a translation vector of the lattice. So the eigenstates of equation 3.1 take the form of a plane wave multiplied by $u_{n\mathbf{k}}(\mathbf{r})$. This is known as Bloch's theorem.

Even with all of the simplifications, a solution to the Schrödinger equation given in equation 3.1 over the whole Brillouin zone is still very complicated. Another simplifying approximation can be made, as often for transport or optical problems it is the energy dispersion close to the band extrema that is most important. These states are the first to become occupied and so play a pivotal role in many semiconductor properties. For a direct band gap semiconductor, where the CBM and VBM are coincident in $\mathbf{k}$-space (most often at the zone centre, the Γ-point), the energy dispersion can be approximated by a Taylor expansion about $|\mathbf{k} - \mathbf{k}_0|$, where $\mathbf{k}_0$ is the wave vector where the extrema occurs

$$E(\mathbf{k} - \mathbf{k}_0) = E(0) + A|\mathbf{k} - \mathbf{k}_0|^2 + B|\mathbf{k} - \mathbf{k}_0|^4 + \ldots \tag{3.4}$$

Setting $\mathbf{k}_0 = 0$ for an extrema at the Γ point and ignoring terms other than the first two (higher power terms contribute negligibly), we see we have a constant and a term dependent on the wave vector squared. This is analogous to the case of a free electron where $V(\mathbf{r}) = 0$ in equation 3.1, noting that the momentum operator is $\mathbf{p} = -i\hbar\nabla_{\mathbf{k}}$. Hence, the parabolic dispersion relation is give by

$$E_{e,h}(\mathbf{k}) = E_{e,h}(0) \pm \frac{\hbar^2 \mathbf{k}^2}{2\,m^*} \tag{3.5}$$

where m^* is the effective mass of the carrier, $E_{e,h}(0)$ is defines the zero of the energy scale, and the $\pm$ signifies the different direction the dispersion can take.

This approximation is quite useful for direct band gap semiconductors, where the

band edges are at the zone centre, and weak or negligible interaction occurs between the valence and conduction band, typically in wide gap semiconductors. However, despite the wide gaps of TCO materials, heavy degeneracy can cause strong interactions between the valence and conduction bands, so often non-parabolicity is considered a better choice for band structure modelling.

3.1.2 $\mathbf{k} \cdot \mathbf{p}$ PERTURBATION

A more accurate method of band structure calculation is the $\mathbf{k} \cdot \mathbf{p}$ perturbation method, which allows for the inclusion of non-parabolic bands. By inserting the Bloch function in equation 3.2 into the Schrödinger equation 3.1, utilising the momentum relation $\mathbf{p} = -i\hbar\nabla_{\mathbf{k}}$ acting upon the Bloch function, the Schrödinger equation becomes [168]

$$\left[H_0 + \frac{\hbar}{m}\mathbf{k} \cdot \mathbf{p} + V\right] u_{n\mathbf{k}}(\mathbf{r}) = E_{n\mathbf{k}} u_{n\mathbf{k}} \tag{3.6}$$

where $H_0 = \frac{\mathbf{p}^2}{2m} + V_0(\mathbf{r})$, magnetic effects may be included through V by the summation of terms such as the spin-orbit potential V_{SO} and crystal-field potential V_{Cr}, and $E_{n\mathbf{k}} = E_n(\mathbf{k}) - \frac{\hbar^2 k^2}{2m}$. At the Γ-point ($\mathbf{k} = (0, 0, 0)$) equation 3.6 reduces down to

$$[H_0 + V] u_{n\mathbf{k}}(\mathbf{r}) = E_{n0} u_{n\mathbf{k}} \tag{3.7}$$

whose solutions form an orthonormal set. Therefore, the wave functions at any value of $\mathbf{k}$ around the Γ-point and their energy eigenvalues can be expressed by treating the $\mathbf{k} \cdot \mathbf{p}$ interaction between the valence and conduction bands as small perturbations

$$\begin{aligned} u_{n\mathbf{k}}(\mathbf{r}) &= u_{n0}(\mathbf{r}) + \frac{\hbar}{m}\sum_{n' \neq n} \frac{\langle u_{n0}|\mathbf{k} \cdot \mathbf{p}|u_{n'0}\rangle}{E_{n0} - E_{n'0}} u_{n'0} \\ E_{n\mathbf{k}}(\mathbf{r}) &= E_{n0}(\mathbf{r}) + \frac{\hbar^2 k^2}{2m} + \frac{\hbar^2}{m^2}\sum_{n' \neq n} \frac{|\langle u_{n0}|\mathbf{k} \cdot \mathbf{p}|u_{n'0}\rangle|^2}{E_{n0} - E_{n'0}}. \end{aligned} \tag{3.8}$$

Information regarding the materials crystal symmetry is contained within the matrix element as knowledge of the eigenvalues of the Bloch function u_{n0} are required to determine its value. This is usually approached using group theory, much in the same way as we set out in appendix A. Also we may see more bands coming into play, and the degeneracy of some bands means the spin-orbit interaction may need to be considered[169].

It is far from trivial to solve equation 3.6 for a given point group as anisotropic effects often need to be considered, and there is still much experimental information lacking about the band structure for many materials. However, simple isotropic cubic systems have been investigated thoroughly over the years[26,54,155,161,169−172], and include many III-V materials and some II-VI including CdO. Many experiments such as cyclotron resonance and optical spectroscopes have informed band structure calculations for these materials, and the systems should hold a lot of relevance to many oxide systems as both possess degenerate valence bands of prominent p-orbital character, and a non-degenerate conduction band extrema of almost entirely s-character.

Kane[171] considered a 4-band model (3 degenerate valence bands and a conduction band) for Zincblende structures. One of the useful results to come from this analysis is that if spin-orbit splitting is neglected, and the system is simplified to only two bands, it is possible to arrive at

$$E_c(k) = \frac{1}{2}\left(-E_g + \sqrt{E_g^2 + 4k^2P^2}\right) + E_k \qquad (3.9)$$

where Kane's matrix element is given by

$$P^2 = \frac{\hbar^2}{2m^*}\left(1 - \frac{m^*}{m}\right)E_g. \qquad (3.10)$$

Whilst the theory proposed by Kane is for zincblende structures, it is often applied as

a more general model to account for conduction band non-parabolicity.

Finally, it is often of use to refer back to equation 3.4 of the expansion of $E(\mathbf{k})$ as the above discussion may inform to give a better approximation the coefficients required. To a good approximation

$$\frac{\hbar^2 k^2}{2m^*} \approx E(1 + \alpha E + \beta E^2) \tag{3.11}$$

where

$$\alpha = \frac{1}{E_g}\left(1 - \frac{m^*}{m}\right)^2 \qquad \text{and} \qquad \beta = -\frac{2}{E_g^2}\frac{m^*}{m}\left(1 - \frac{m^*}{m}\right). \tag{3.12}$$

These parameters can be found by expanding the square-root term in a Taylor expansion[161]. If the cube term is dropped then this equation is known as the α-approximation. If the α term is dropped then the parabolic band structure is retained. These approximate forms of the bands structure are often used when performing analysis requiring non-parabolic conduction bands.

Figure 3.1a shows the conduction band structure for a material with an effective mass of $0.22m_0$ and band gap of $E_g = 2.9$ eV (such as In_2O_3) calculated using the parabolic structure, two-band $\mathbf{k} \cdot \mathbf{p}$ method, and the expanded approximations to first and second order (α and β approximations). The non-parabolicity of the conduction band is evident at wave vectors past 0.1 Å⁻¹, where the parabolic curve diverges from the others. The effect of this can be seen on both the density of states, and carrier concentrations, seen in figure 3.1b and c. These can be calculated using the general expressions

$$g(k) = \frac{k^2}{\pi^2}\left(\frac{dE(k)}{dk}\right)^{-1} \tag{3.13}$$

and
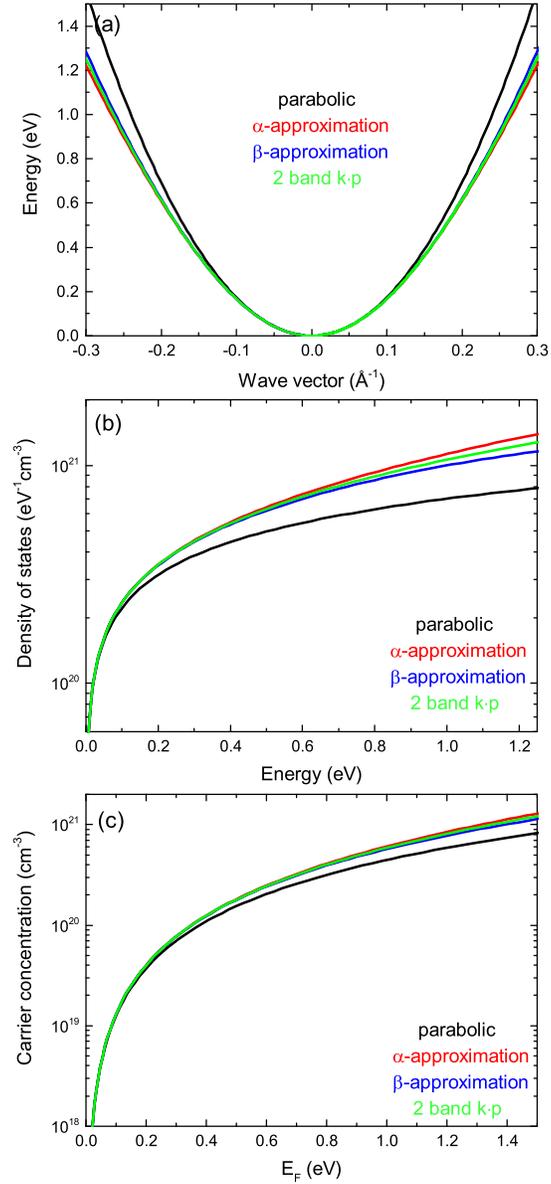
Figure 3.1: (a) conduction band dispersion using different approximations to account for the non-parabolicity of the band. In each case E_F is referenced to the CBM. Here a parabolic band was used, and the non-parabolicity is modelled using Kane's two-band model, the α-approximation, and the β-approximation. (b) calculated density of states from the approximations used above. (c) calculated carrier density for the approximations used above.

$$n = \int_0^\infty f(E)g(E)\,dE \tag{3.14}$$

where $f(E)$ is the Fermi-Dirac distribution function

$$f(E) = \frac{1}{1 + e^{\frac{E-\mu}{K_B T}}}. \tag{3.15}$$

Note that the carrier dependent electron effective mass can also be evaluated from these calculations using

$$m^*(E) = \hbar^2 k \left(\frac{dE(k)}{dk} \right)^{-1}. \tag{3.16}$$

3.2 Density Functional Theory

Whilst extremely powerful, band structure approximations such as those seen above are not always sufficient to provide an adequate description of the material under investigation. Sometimes information is required across the whole Brillouin zone, or occasionally we want to consider remote bands which would be ignored via other methods. To the electron spectroscopist, the ability to deconvolve the density of states into orbital contributions is extremely valuable, or it may be informative when considering doped materials to determine the effects of dopants on the host material structure. Many of these considerations can be accounted for using more sophisticated approaches, and likely the most successful of these to date is density functional theory (DFT). DFT is a large topic with a multitude of applications. Hence, only a brief discussion of the subject will be given here. Whilst DFT calculations have been invaluable to the work in this thesis, these were all done on a collaborative basis here, and so other sources should be sought for more information (see for example Martin[173]).

In DFT, the electron density can be represented by

$$n(\mathbf{r}) = \langle \Phi(\mathbf{r}) | \Phi(\mathbf{r}) \rangle \tag{3.17}$$

which is treated as the central variable when solving the Schrödinger equation, rather

than the many-body wave function itself. The first big strides toward modern DFT were taken by Hohenberg and Kohn[174], who showed that for a system of interacting particles under the influence of an external field (V_{ext}), the total energy of the system can be described by

$$E_{HK}[n] = F_{HK}[n] + \int V_{ext}(\mathbf{r_i})n(\mathbf{r_i})d^3r \qquad (3.18)$$

where $F_{HK}[n]$ is known as the Hohenberg-Kohn universal functional. The ground state energy is found by minimizing the functional, corresponding to the ground state electron density $E_{HK}[n_0]$. If we known the form of $F_{HK}[n]$ then we know the exact ground state energy of the system. The form of $F_{HK}[n]$ is actually very complicated and Hohenberg and Kohn gave no guidance on how to construct the correct functional.

The next step in the development of modern DFT was provided by Kohn and Sham[175] shortly after the Hohenberg-Kohn model was proposed, with the goal of providing an approach to compute an approximate functional $F[n(\mathbf{r})]$. They suggested that an interacting many-body system could be replaced by a different, simpler system. They assume the auxiliary system has the same ground state density $n_0(\mathbf{r})$ as the interacting many-body system, but the new system is made up of non-interacting electrons.

The Kohn-Sham model for the full interacting many body system can then be written as

$$E_{KS}[n(\mathbf{r})] = T_{KS}[n(\mathbf{r})] + V_{ee}[n(\mathbf{r})] + \int V_{ext}(\mathbf{r})n(\mathbf{r})d\mathbf{r} + E_{xc}[n(\mathbf{r})] \qquad (3.19)$$

where $V_{ee}[n(\mathbf{r})]$ is the classical Coulomb interaction energy of the electron density with itself, $T_{KS}[n(\mathbf{r})]$ is the kinetic energy of the non-interacting particles, and V_{ext} is the external potential due to the nuclei and any other external field. All of the difficulty in many-body calculations are contained within the exchange-correlation functional

$E_{xc}[n(\mathbf{r})]$, which is the sum of the non-classical electron-electron interactions and the correction to T_{KS} which accounts for the non-classical interactions (therefore giving the total T). The universal function is

$$F_{KS}[n(\mathbf{r})] = T_{KS}[n(\mathbf{r})] + V_{ee}[n(\mathbf{r})] + E_{xc}[n(\mathbf{r})]. \tag{3.20}$$

The exchange-correlation functional is the only unknown in the Kohn-Sham approach to DFT. If this function were known then the Kohn-Sham equation could be solved for independent particles using the self consistent variational theorem[173]. However, only approximate forms of the exchange-correlation functional have been proposed.

The electron exchange interaction is due to the wave function of indistinguishable particles being subject to exchange symmetry, either remaining unchanged (symmetric) or changing (antisymmetric) when two particles are exchanged, interchanging their spin and spatial coordinates. Fermions must have antisymmetirc wave functions under exchange due to the Pauli exclusion principle. Electron correlation was a term coined by P-O Löwdin[176] but may have been studied much earlier[177]. It has its physical origins in the effect that all of the electrons in an interacting many-body system have on the movement of a single electron. An exact wave function may not be able to be represented by single determinants, which may miss important physics such as the Coulomb correlation of the system. While Hatree-Fock theory (which attempts to solve the Schrödinger equation using many single electron wave functions[178,179]) provides an exact formalism for exchange as it uses the Slater determinant to describe the wave function, it does not include correlation effects and so is always higher in energy than the exact energy. For this reason, it usually preferable to make approximations to both the correlation and exchange functionals in the Kohn-Sham formalism, which is less computationally heavy than the Hartree-Fock method.
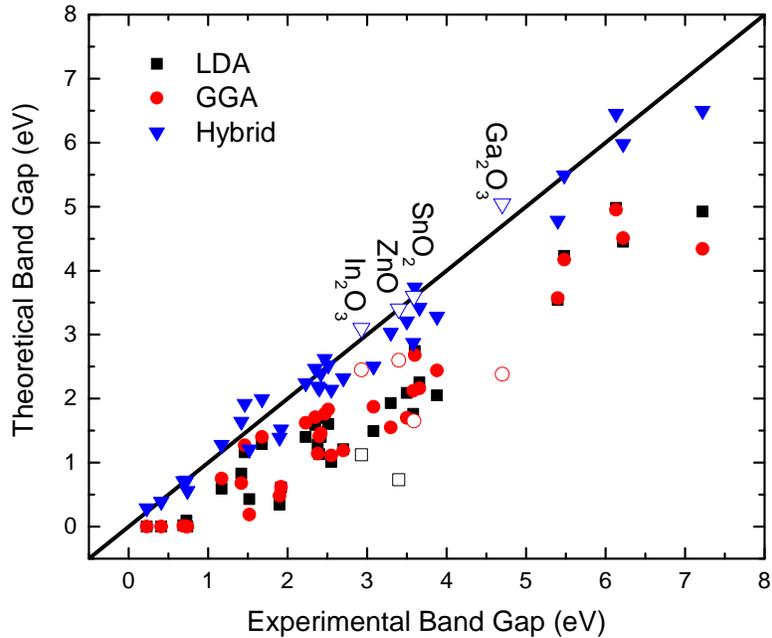
Figure 3.2: Theoretically calculated band gaps using different functional approaches versus experimentally measured gaps. Graph replicated from data in Heyd et al. with data relevant to TCO related semiconductors taken from other sources.[31,69,180-184]

Many attempts to describe the exchange-correlation functional have been suggested in the literature. However, most typically either the local density approximation[174,175] (essentially assuming a homogeneous electron gas system), the generalized gradient approximation[185,186] (using gradients of the density to gain more information around the electrons), or hybrid functionals[187-189] (combining the exact exchange formalism of Hartree-Fock theory with the approximate exchange and correlation calculated through DFT) are employed. Figure 3.2 shows a comparison of how well the different functionals perform at predicting the band gap of different semiconductors. TCO related semiconductors are highlighted as empty data points. Clearly hybrid functionals perform by far the best, in many cases matching the experimentally determined band gaps. Note that the particular functionals used are not specified here.

Finally, to show the difference in band structure calculation methods, figure 3.3 shows different approximations for the conduction band of In_2O_3. The non-parabolic
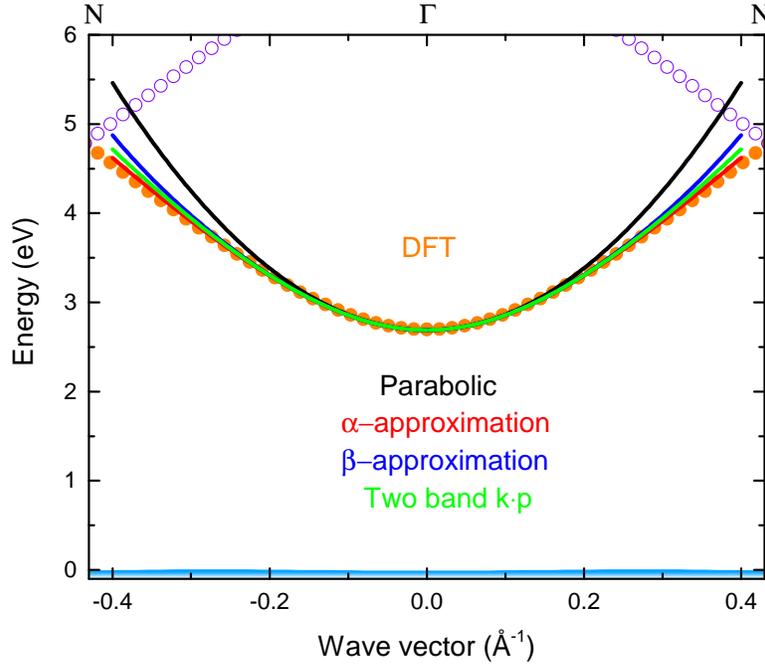
Figure 3.3: Comparison of parabolic, two-band k·p and approximate methods and a hybrid DFT method of band structure calculation for In_2O_3.

and indeed even the parabolic approximate bands reproduce the DFT calculated conduction band reasonably well around the Γ-point. However, away from this the approximations clearly struggle, especially after $\sim 0.2\text{Å}^{-1}$. Much of the time this will not be an issue in optical or transport calculations, as the Fermi level is likely to be much lower in energy than this point, ~ 1 eV above the CBM. The benefit of a much more simple calculation will often outweigh the inaccuracy. However, the true benefit of DFT cannot be realised by this simple comparison. Even, in figure 3.3 we see extra bands calculated through the DFT method which will have some interaction effect on the band of interest near the zone edges. This could be included in the k·p method but at greater cost. A much broader wealth of information is offered using the DFT approximations, while models for the band structure including the k·p method, are usually informed by experiment, *ab initio* or first principles calculations do not require such information, and are usually tested against experiment (as in figure 3.2).

3.3 Space Charge Calculations

3.3.1 POISSON'S EQUATION

Chapter 6 gives results on the band bending properties of β-Ga$_2$O$_3$, and how this is influenced by atoms adsorbed onto the surface. At the surface, translational symmetry of the crystal is broken, allowing evanescent gap states to exist within the semiconductor band gap. These have either donor or acceptor character dependent upon their position relative to the charge neutrality level of the semiconductor. Where ionized surface states exist, charge neutrality is attained by band bending. These ionized surface states set up the near surface space-charge. For n-type materials, ionized donor-like (acceptor-like) surface states will result in downward (upward) band bending at the surface, and the formation of an electron accumulation (depletion) layer (see figure 3.4). Hence, the space charge layer is a distribution of carriers close to the surface. The charged surface induces an electric field, which in turn is screened by the rearrangement of the charges (or carriers) at the surface over a distance approximately equal to the Thomas-Fermi screening length leading to either carrier accumulation or depletion at the surface.
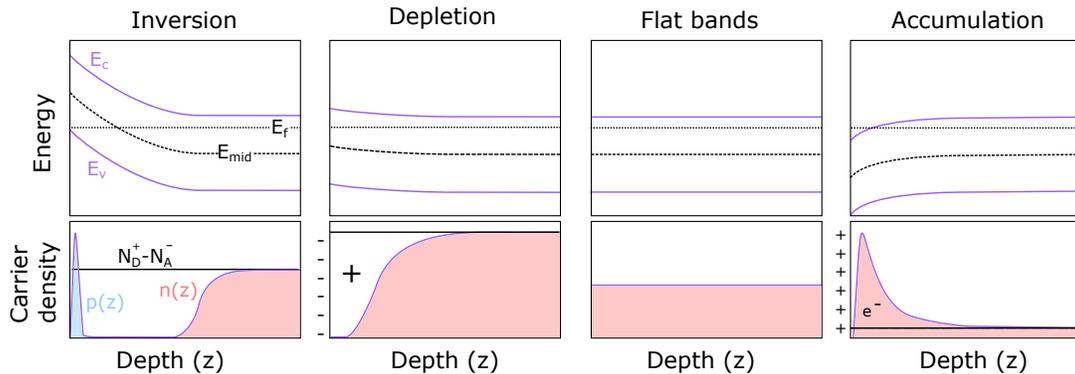
Figure 3.4: Schematic representation of the band bending and associated charge profiles for the case of inversion, depletion, flat bands, and accumulation space-charge layers in an n-type semiconductor. In the plots of the carrier concentration variation with depth, the net ionized donor density is shown ($N_D^+ - N_A^-$).

In order to investigate the space-charge layers in this material, a model is required that incorporates the space-charge potential ($V(z)$) and the electron density as a function of depth in a given sample. This can be described using Poisson's equation, which in the one-dimensional case is given by

$$\frac{d^2 V}{dz^2} = -\frac{e}{\varepsilon(0)\varepsilon_0} \left[N_D^+ - N_A^- - n(z) + p(z) \right] \qquad (3.21)$$

where $\varepsilon(0)$ is the static dielectric constant, $N_D^+(N_A^-)$ is the bulk ionised donor (acceptor) concentration, and $n(z)(p(z))$ is the depth dependent electron (hole) density. The potential $V(z)$ satisfies the boundary conditions $V(z) \to 0$ and $\frac{dV(z)}{dz} \to 0$ as $z \to \infty$ and there is no band bending in the bulk of the material, and $\frac{dV}{dz}\big|_{z=0} = \frac{e}{\varepsilon(0)\varepsilon_0} N_{ss}$, where N_{ss} is the surface state density.

In order to obtain the potential $V(z)$ and the carrier density, Poisson's equation has to be solved self-consistently with the one-electron Schrödinger equation. This is achieved by using a trial potential $V(z)$ with an iterative method to obtain a single solution[190,191]. However, this quickly becomes non-trivial or computationally expensive for non-parabolic bands.

3.3.2 Modified Thomas-Fermi Approximation

To simplify the problem, Paasch and Übensee[192] suggested using a modified Thomas-Fermi approximation (MTFA) where the potential is calculated by solving the Poisson equation subject to the boundary conditions discussed above. The carrier densities correspond to the conduction band and valence bands, which are calculated from

$$n(z) = \int_0^\infty g_c(E)f(E)f_{MTFA}(z)\,dE$$
$$p(z) = \sum_i \int_{E_{V_i}}^{-\infty} g_{v_i}(E)[1 - f(E)]f_{MTFA}(z)\,dE \qquad (3.22)$$

respectively, where the sum over i denotes the sum over a number of valence bands, and $g(E)$ is the density of states for a band (which may be non-parabolic for many semiconductors). $f(E)$ is the simple Fermi-Dirac distribution function including the potential dependence

$$f(E) = \frac{1}{1 + e^{\frac{E - E_F + V(z)}{K_B T}}} \tag{3.23}$$

and $f_{MTFA}(z)$ is the MTFA factor which accounts for the potential barrier at the surface [192,193]

$$f_{MTFA}(z) = 1 - \text{sinc} \left[\frac{2z}{L} \left(\frac{E}{K_B T} \right)^{\frac{1}{2}} \left(1 + \frac{E}{E_g} \right)^{\frac{1}{2}} \right] \tag{3.24}$$

where for non-degenerate semiconductors L is the thermal length $L = \frac{\hbar}{\sqrt{2m_0^* K_B T}}$, or in degenerate semiconductors it is the Fermi-length $L = \frac{1}{k_F}$ where k_F is the Fermi wave number, and $\text{sinc}(x) = \frac{\sin(x)}{x}$. This correction factor represents the interference of incident and reflected wave functions due to the potential barrier at the surface, and thus causes the carrier concentration to tend smoothly to zero at the surface.

Numerical solutions of equations 3.21 and 3.22 are found by using a trial potential, followed by an interval bisection method to converge to an approximate solution of the one-electron band-bending potential that satisfies the boundary conditions, and the carrier density profiles as a function of depth. Note that for doped materials it is usually assumed that only the majority carrier type is required for an accurate approximation of the potential in equation 3.21 (i.e. $n(z)$ and N_D^+ for an n-type semiconductor). Examples of solutions for n-type semiconductors are seen in figure 3.4. In figure 3.4 it is assumed that the band bending experienced by the valence and conduction bands is equal, an assumption that is carried throughout this thesis. However, this is not necessarily the case, as was shown previously for several materials including CdO [194]

and In_2O_3 [30]. The origin of this difference in the level of band bending experienced between the two band extrema for materials displaying surface electron accumulation, is attributed to many-body effects (band gap renormalization) in the accumulation layer, which is predominant in degenerate materials with high carrier densities in the accumulation layer relative to the bulk of the material. Due to the confining potential well that is set up from the extent of the downward band bending, the host conduction band states become quantized into subbands displaying dispersion in the surface plane. This can be further investigated using techniques such as angle-resolved photoemission spectroscopy (ARPES), in conjunction with coupled Poisson-Schrödinger type calculations [194]. This style of study would help to further inform the results of chapter 6, and could be suitable for future investigation. However, due to the low surface carrier density seen in β-Ga_2O_3, information regarding the quantized conduction band states may be very difficult to accurately determine.

4

Self-Compensation in F-Doped SnO$_2$

Chemical vapour deposited F-doped SnO$_2$ samples were obtained from NSG group. F ion implanted SIMS standards were prepared by Nianhua Peng at Surrey Ion Beam centre. Secondary ion mass spectrometry measurements were performed by Mark Farnworth at Pilkington Technology Centre and data analysed by the author. Density functional theory calculations were carried out and analysed by Benjamin A. D. Williamson in the group of David O. Scanlon at University College London. The results presented here were published as:

J. E. N. Swallow, B. A. D. Williamson, T. J. Whittles, M. Birkett, T. J. Featherstone, N. H. Peng, A. Abbott, M. Farnworth, K. J. Cheetham, P. Warren, D. O. Scanlon, V. R. Dhanak, and T. D. Veal, Self-Compensation in Transparent Conducting F-Doped SnO$_2$. Adv. Funct. Mater. 28, 1701900 (2018).

4.1 Introduction

FTO is a TCO consisting of a stannic oxide (SnO_2) framework with heavy donor incorporation of the fluorine dopant, maintaining a rutile structure[195,196]. A number of thin-film deposition methods are regularly used to prepare FTO, including spray pyrolysis[197], chemical vapour deposition[198], pulsed laser deposition[199], and magnetron sputtering[200]. Commercial FTO is most commonly deposited via atmospheric pressure chemical vapour deposition (APCVD) in an on-line coating process where the manufacturing of the glass substrate and deposition of the TCO films are performed in a continuous process. An example of these products is the TEC™ glass range produced by Nippon Sheet Glass (NSG)[201].

FTO possesses a fundamental electronic direct band gap of E_g=3.6 eV[202,203], and an optical band gap that can often exceed 4 eV depending on the level of fluorine incorporation[199,204]. This contributes to achieving optical transmission in the visible region commonly around 80%[205]. In these materials it is generally assumed that fluorine acts as a substitutional, singly-charged donor occupying an oxygen site. This is often assumed because oxygen and fluorine have nearly the same atomic radii and similar bond energies with tin which should assist in fluorine being easily incorporated into the material[196,206]. Fluorine doping of tin dioxide can result in very low resistivity FTO films, regularly less than 4×10^{-4} Ωcm[201,207,208].

Intuitively, it is expected that the more fluorine atoms that are incorporated into the tin dioxide matrix, the more free electrons become available for conduction[196]. If this is the case, the amount of fluorine incorporated is only limited by the trade-off between optical and electrical properties - as the carrier density is increased, there is a corresponding increase in conduction electron plasma frequency and associated

plasma reflectivity that limits the infrared transparency[209−211]. However, one interesting observation reported many times over the years is that the resistivity of FTO will initially decrease as carrier concentration increases and then begin to increase when carrier concentrations become sufficiently large[207,212−216]. While the initial decrease in resistivity is relatively simple to explain, being due to the extra free carriers contributing to conduction introduced into the material by the fluorine dopant, the origin of the eventual increase in resistivity at high doping levels is a much more debated issue. A number of phenomena have been suggested, with a general underlying theme of the fluorine interstitial playing a major role[195,196,214]. However, only very limited evidence is available supporting this claim, mainly based on x-ray diffraction studies[212−214,217,218].

This study uses a combination of Hall effect measurements and modelling, and theoretical calculations based on density functional theory (DFT) to determine the factors limiting the carrier mobility in APCVD-deposited FTO films on soda lime glass. This information is then related to quantitative chemical analysis using x-ray photoemission spectroscopy (XPS) backed by secondary ion mass spectrometry (SIMS). These results point to a self-compensation mechanism occurring in FTO at high doping levels. With the aid of DFT formation energy calculations and XPS results it is possible to infer the likely defect species.

4.2   Experimental and Theoretical Methods

Fluorine-doped tin dioxide (FTO) thin films deposited on glass by APCVD were obtained from NSG Group. Samples consisted of a multi-layer structure with an undoped SnO_2 layer $\sim$ 25 nm deposited directly on the glass substrate providing a rough surface for the subsequent layers to adhere to. A SiO_2 layer follows of thickness $\sim$ 25 nm acting as a sodium diffusion barrier, and finally the electrically active $F:SnO_2$ layer is deposited on top of this. The tetragonal rutile structure associated with the SnO_2 was

confirmed via x-ray diffraction (not shown here). The samples are polycrystalline in nature and no impurity phases were present. Samples were prepared for measurement by mechanically cleaning the surface with laboratory wipes and isopropyl alcohol to remove large particulates, as well as being treated in an ultrasonic bath submerged in diluted surface cleaner (decon 90 surface cleaning agent) and then isopropyl alcohol, and rinsed in deionized water.

Time-of-flight secondary ion mass spectrometry (SIMS) combined with profilometry allowed the thickness of the FTO films to be determined. Film thickness of the FTO layers ranged from 300 to 750 nm. Fluorine concentrations in the films were determined from XPS peak areas of the main core levels of the constituent elements, taking into account the atomic sensitivity factors (ASF) provided by Moulder *et al.*[122], which ensure measured peak areas are corrected to account for photoionization cross sections. It should be noted that the ASF also depend on some factors specific to the measurement instrument used such as the analyser transmission function. The ASF taken from Moulder *et al.* are not specific to our XPS system and so the uncertainty associated with the determination of compositions can be considered sizeable. The fluorine concentration of the samples varied from around 0.7% (atomic percentage) to 1.7%. Time of flight (ToF) SIMS measurements were made, using the IonTof TOF-SIMS 5 instrument. Pulsed Bi_3^+ analysis ions were used to bombard the sample surface while a 1 keV Cs source was used as the sputtering beam. Analysis was performed on a $50.8 \times 50.8 \mu m^2$ sputter area. The CsF_2^+ ion intensity versus erosion time were converted to fluorine concentration versus depth. To make standards for quantitative depth profiling using SIMS, undoped SnO_2 samples were deposited on glass via radio frequency magnetron sputtering, and fluorine ions were implanted into the materials at an energy of around $\sim$ 100 keV. This corresponded to an implantation depth of $\sim$ 114 nm. Three standards were deposited in total, with ion fluences of 1×10^{16},

5×10^{15}, and 1×10^{15} cm⁻². All ion implantation in this study was performed at Surrey Ion Beam centre.

Hall effect measurements were performed using the standard van der Pauw configuration at a field strength of 0.8 T to determine the free carrier concentration (n) and transport mobility (μ) of FTO samples. Measurements were performed at room temperature on the samples, all of which displayed n-type conductivity. The measured free carrier concentrations across the sample range varied from $(1.81 \pm 0.01) \times 10^{20}$ cm^{-3} to $(5.48 \pm 0.04) \times 10^{20}$ cm^{-3} and electron mobilities varied from 27.3 ± 0.2 cm^2/V·s to 38.2 ± 0.1 cm^2/V·s. Samples which displayed high carrier concentrations and low mobilities corresponded to those of high fluorine content. Temperature dependent Hall effect was also performed on some samples, with sample temperature being varied from 10 to 300 K (± 0.5 K) (see Figure 4.2).

High resolution XPS measurements were performed using a SPECS monochromatic Al Kα ($h\nu = 1486.6$ eV) X-ray source operated at 300 W. Photoelectrons were analysed using a PSP Vacuum Technology hemispherical electron-energy analyser, with mean-radius of 120 mm operated at a pass energy of 10 eV. The spectrometer was calibrated using a polycrystalline silver foil which had been Ar$^+$ sputtered to achieve a clean surface. The silver $3d_{5/2}$ and Fermi edge were measured for energy position and peak width calibration. For more information on the calibration process and estimated uncertainties of peak measurements made with this spectrometer see Ref. 219. All core level positions were referenced to the Fermi level of the FTO. While under ultra-high vacuum conditions, further surface treatment was performed to remove surface contaminants. This was done via Ar$^+$ ion bombardment while monitoring the C 1s peak and Sn 3d peaks at regular intervals. Sn $3d_{5/2}$ to C 1s peak ratios were compared between sputter cycles in order to assess the level of surface cleanliness achieved. It should be noted that due to the small quantity of fluorine present in the samples and

the relatively low cross section for photoionisation of F 1*s* photoelectrons, obtaining reasonable signal-to-noise on the fluorine 1s region takes 10-50 hours of data acquisition. This is consistent with what has been seen previously.[198,220,221]

Infrared (IR) reflectance measurements were performed using a Bruker Vertex 70v Fourier-transform infrared spectrometer at 11° angle of incidence (with respect to the normal of the sample surface). A CaF_2 beam splitter and DLaTGS detector were used. Spectra were recorded over an energy range of 0.10 to 0.75 eV in order to completely encompass the plasma resonance cut-off. FTIR measurements were performed under vacuum ($\sim$ 2 mbar) to minimise the effect of atmospheric water vapour and carbon dioxide vibrational modes on the recorded spectra. The reflectance was simulated using the transfer matrix method[148]. The simulation accounted for both *s* and *p*-polarized reflectance, considering a five layer (vacuum/F:SnO_2/SiO_2/SnO_2/soda-lime glass) stratified medium, assuming complete incoherence in the thick glass substrate. The simulation uses a two-oscillator expression for the dielectric function to allow for determination of optical parameters. Transmittance was measured with a Shimadzu UV-Vis-IR 3700 spectrophotometer over an energy range of 1 to 5 eV.

Ab initio calculations were performed using DFT implemented using the periodic code, VASP.[222–225] The projector-augmented wave method[226,227] was used to describe the interaction between the core electrons (Sn[Kr], O[He], and F[He]) and the valence electrons. The hybrid functional PBE0 developed by Adamo and Barone[189,228], was used in order to combat the self-interaction error and thus allowed for an accurate description of the band gap of SnO_2. Hybrid functionals were consistently shown to provide improved calculations of both geometry and electronic structure,[229–236] and PBE0 was shown to predict these properties for tin-based TCOs with a high degree of accuracy.[89,237–243] PBE0 incorporates 25% of exact Fock exchange to the PBE (Perdew, Burke, and Ernzerhoff)[186] formalism.

The intrinsic defects and extrinsic dopants were simulated using a $2 \times 2 \times 3$ supercell containing 72 atoms together with a Γ-centred $2 \times 2 \times 2$ k-point mesh and a 400 eV plane wave energy cutoff. All the defect calculations were spin-polarized. The individual systems were deemed to be converged when the forces on all the atoms were less than 0.01 eV per atom.

4.3 Results and discussion

4.3.1 TRANSPORT MOBILITY

Transport mobilities of the FTO samples as a function of carrier concentration, as measured primarily by Hall effect, are shown in figure 4.1. For degenerately doped semiconductors, such as transparent conducting oxides, the dominant carrier scattering/mobility reducing mechanism in the majority of cases is ionised impurity scattering[244,245]. To simulate this effect, the degenerate form of the Brooks-Herring formula[164,246] has been implemented (equation 2.93, as discussed in section 2.3.2), labelled ionized impurity (II) in figure 4.1. All donors are assumed to be ionized and have a charge state of $Z_D = 1$, corresponding to substitutional fluorine, F_O. Other scattering mechanisms have been taken into account and are shown in figure 4.1. These are acoustic deformation potential (ADP)[57,247], longitudinal polar-optic phonons (LPO) (using the formalism set out by Low and Pines[248] and adapted by Fonstad and Rediker for SnO_2[57], however a number of other approaches do exist[249,250]) and grain boundary scattering for both degenerate[251] (IG(Deg)) and non-degenerate[252] (IG) systems. The effects of neutral impurities, and other phonon effects were found to be negligible.

Individual carrier scattering mechanisms are modelled and displayed in Figure 4.1 along with the combined transport mobility calculated according to Matthiessen's rule. This approach assumes the scattering mechanisms are independent of each other. All
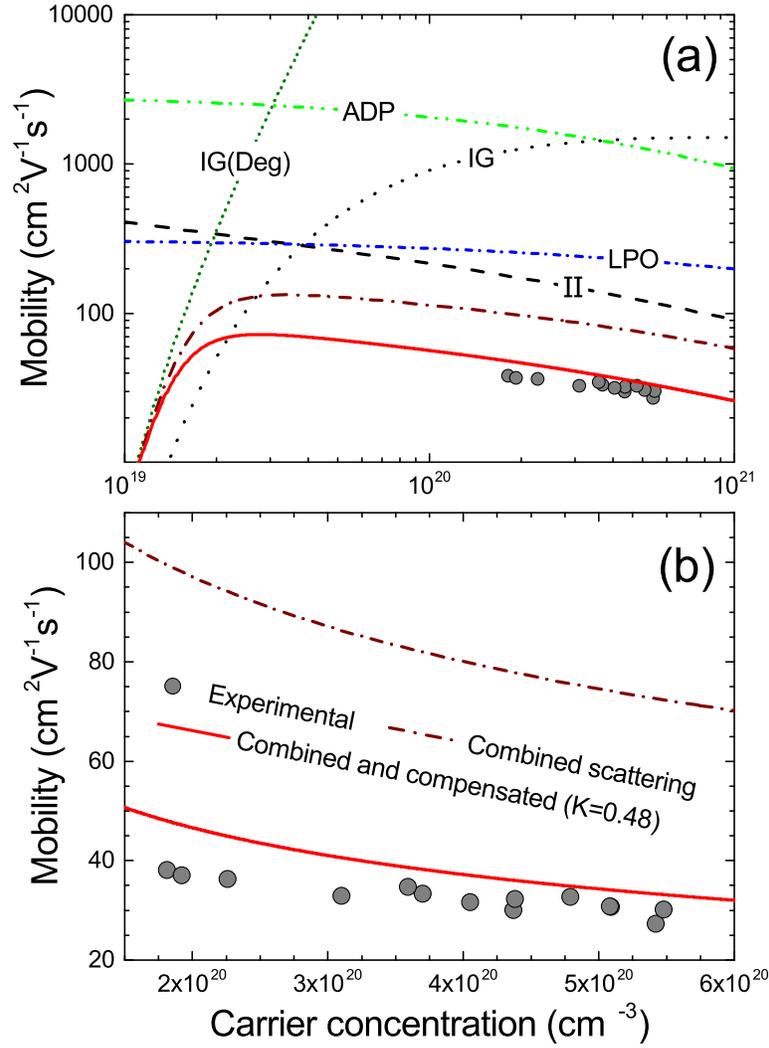
Figure 4.1: Transport data and simulation for Hall mobility versus carrier concentration of FTO. (a) displays all theoretical curves calculated over a wide carrier concentration range whilst (b) shows only the combined theoretical model, the effect of compensation and the experimental data over the carrier concentration range relevant to the data. Model curves of dominant scattering mechanisms result from successive addition via Matthiessen's rule. The scattering mechanisms displayed are longitudinal polar-optical (LPO), acoustic deformation potential (ADP), grain boundary for both the degenerate (IG(Deg)) and non-degenerate case (IG), ionized impurity (II), and the effect due to compensation in the system.

curves have been calculated using a band-edge effective mass of $m_r = 0.27m_0$[253] and a relative static dielectric constant of $\epsilon(0) = 12.2$[254]. The band-edge effective mass and dielectric constants are assumed to be isotropic for these polycrystalline films[57]. The non-parabolicity of the conduction band was also accounted for using a carrier

density-dependent effective mass[43].

As shown in Figure 4.1, ionized impurity scattering becomes the dominant mobility limiting mechanism in SnO_2 above a carrier density of $\sim 5 \times 10^{19}$ cm^{-3}. This holds well with the predictions made by Martinez *et al.*[255] Our data points reside well above this threshold, indicating the heavily limiting effects of grain boundaries are negligible for our films. However, even in this regime the theoretical mobilities calculated are roughly three times higher than those measured. This relationship is supported by the modelling of temperature dependent Hall effect which can be seen in Figure 4.2. The results agree well with the work of Haitjema *et al.*[256] who suggest their calculated theoretical mobility is four times greater than that found experimentally. We attribute this to the effects of self-compensation in these samples. Self-compensation arises by the formation of acceptor defects that counter the dopant impurity properties[245]. The effects of self-compensation can be included in the transport model utilising a factor termed the 'compensation ratio' ($K = \frac{N_A}{N_D}$), which is simply defined as the ratio of the number of acceptors to donors present in the material[257]. This modifies the mobility limit due to ionized impurities as $\mu_{II} = \mu_{II}(0)\frac{Z_D - |Z_A|K}{Z_D^2 + Z_A^2 K}$ where $\mu_{II}(0)$ is the unattenuated ionized impurity mobility limit and Z_D and Z_A are the charge state of donors and acceptors respectively[258]. If the charge state of the acceptor defect is assumed to be $Z_A = -1$, this equation reduces to $\mu_{II} = \mu_{II}(0)\frac{1-K}{1+K}$. Incorporating the compensation ratio into the model fit (the brown dash-dot curve labelled combined scattering in figure 4.1) the level of compensation is determined for a singly charged acceptor to be $K \sim 0.48$.

For degenerately doped systems, the carrier concentration is not expected to vary with temperature due to all carriers being ionized in the temperature range recorded. This is supported in figure 4.2(a), a straight line has been plotted with the data to act as a guide for the eye. The same transport model was applied as in figure 4.1, which
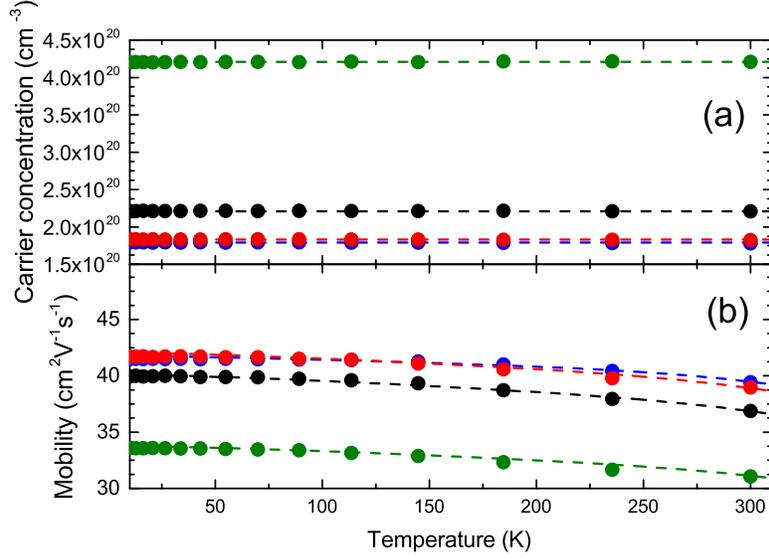
Figure 4.2: Temperature dependent carrier transport determined via Hall effect. (a) Carrier concentration as a function of temperature in the range of 10-300K. (b) Free carrier mobility as a function of temperature.

accounts for ionized impurity scattering, acoustic deformation potential, longitudinal polar-optic phonons and grain boundary scattering has been plotted in (b). Ionized impurity scattering is the dominant scattering mechanism as expected, and does not vary with temperature because all carriers are ionized at all temperatures. Above~125K the effect of longitudinal polar optical carrier scattering becomes much more important than at low temperatures which is evident in the downward curvature of the simulation lines.

Resistivity can be determined from carrier density (via either Hall effect) and mobility (obtained via Hall effect) measurements by $\rho = \frac{1}{ne\mu}$. Using both the data points and model curves from figure 4.1, resistivity is plotted as a function of carrier concentration seen in figure 4.3. The combined scattering curve (brown dashed curve) is determined via Matthiessen's rule, combining the main scattering mechanisms present for degenerate FTO, i.e. ionized impurities, grain boundaries and phonon effects. This represents the theoretical minimum resistivity possible for FTO. The solid red curve displays the compensated curve fitted to the data. The compensated curve accounts
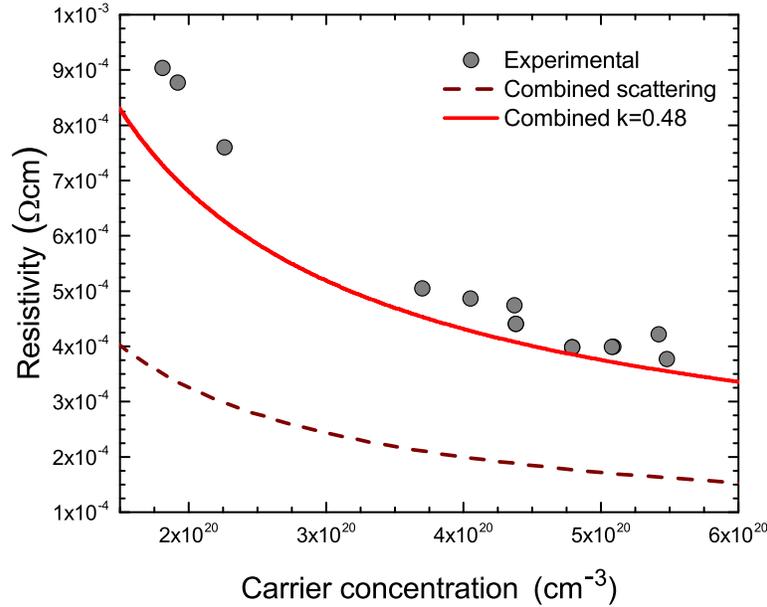
Figure 4.3: Transport data and simulation of resistivity and carrier density as determinedvia Hall effect and SIMS. The uncompensated model is shown as a dashed curved whilst the compensated (K=0.48) curve is solid.

for the effects of compensating acceptor defects determined to be due to the fluorine interstitial. These defects limit the mobility at a given carrier concentration, negating the intended effects of heavy dopant incorporation.

4.3.2 THEORETICAL PREDICTION OF DEFECTS

Hybrid density functional theory calculations have been performed to determine the formation energies of a range of likely intrinsic and extrinsic defects states in tin dioxide as a function of the Fermi level position. The defect species considered in this study are substitutional fluorine (F_O), interstitial fluorine (F_i), a fluorine substitutional-interstitial pair ($[F-F]_O$) together with the dominant intrinsic donor and acceptor defects in SnO_2 such as the oxygen vacancy (V_O) and the tin vacancy (V_{Sn}). All of these defects are displayed in Figure 4.4 for both *Sn-Rich/O-Poor* (left) and *Sn-Poor/O-Rich* (right) which are at the 'extremes' of the chemical potentials, i.e the formation of Sn metal under *Sn-Rich/O-Poor* conditions and O_2 gas under Sn-poor/O-rich conditions

and thus the experimental situation is expected to lie somewhere between these two regimes. In each plot the valence band maximum (VBM) is set to $E_F = 0$ eV and the onset of degeneracy occurs from the conduction band minimum ($E_F = 3.6$ eV) indicated by the graded orange area.
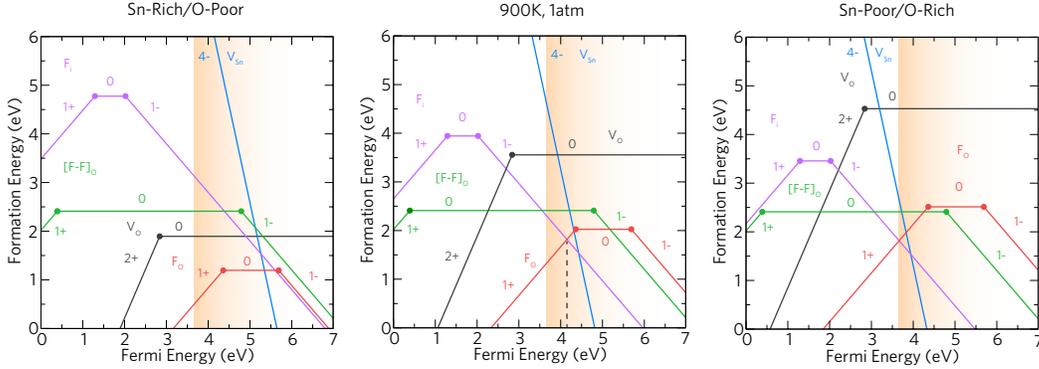
Figure 4.4: The calculated formation energies as a function of Fermi level position for *Sn-Rich/O-Poor* (left), at 900 K and 1 atm (middle) and *Sn-Poor/O-Rich* (right) growth conditions. In each regime the VBM is set at 0 eV and the conduction band is denoted by the orange area with the CBM at 3.6 eV. The dashed black line shown in the middle plot (900 K, 1 atm) represents the Fermi energy at the point where the F_O^+ and F_i^- lines cross. The solid dots indicate the transition levels from charge state q to q', $\varepsilon(q/q')$.

Under both growth regimes the oxygen vacancy (V_O) acts as the lowest formation energy intrinsic donor, behaving as a 'deep' defect with a negative-U behavior (the 2+/0 transition level occurs ~ 0.76 eV below the conduction band minimum (CBM)) meaning that V_O is unlikely to be a source of conductivity in SnO_2 which has been seen in previous theory[62,237,238] and experimental[71] studies alike. Oxygen vacancies have been identified as the intrinsic defect present in undoped TCOs such as In_2O_3, ZnO and SnO_2 via positron annihilation spectroscopy.[259] The neutral charge state for the tin vacancy (V_{Sn}) in each of the growth regimes has a very high formation energy and thus will not form or will form in negligible quantities. Under conditions which favour *p*-type defects (*Sn-Poor/O-Rich*) where the formation energy is ~ 8.37 eV. This defect lies *ultra* deep in the band gap where the 0/1- lies ~ 1.75 eV above the VBM.

Under *Sn-Rich/O-poor* conditions substitutional fluorine (F_O) is the lowest forma-
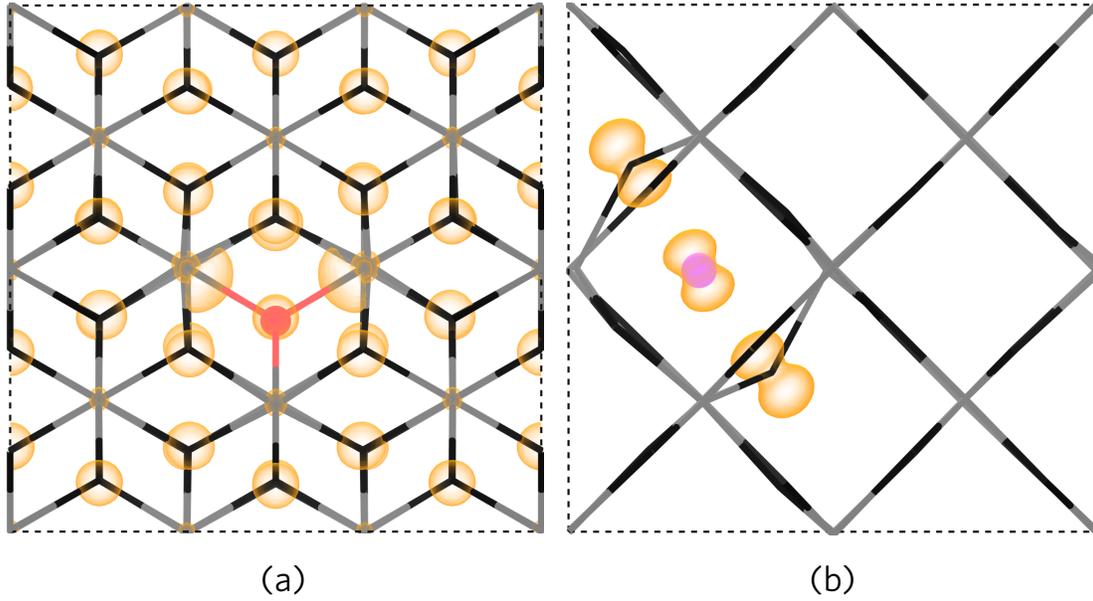
Figure 4.5: The calculated partial charge densities of (a) F_O^+ and (b) F_i^- down the {010} and {001} directions respectively. The Sn (grey) and O (black) atoms are depicted using a stick model for clarity, whilst the F atoms are coloured red (F_O) and pink (F_i) corresponding to the defect colour used in Figure 4.4. Charge densities of 0.001 $eV\text{\AA}^{-1}$ and 0.02 $eV\text{\AA}^{-1}$ were used for (a) and (b) respectively.

tion energy donor and is *shallow* with the 1+/0 transition occuring ~0.76 eV above the CBM and the 0/1- level occuring ~2.09 eV above the CBM. Figure 4.5a displays the partial charge density of F_O in the neutral charge state (F_O^0) showing the delocalisation of electron density in the conduction band, consistent with the resonant nature of substitutional F. There is also negligible distortion to the SnO_2 lattice as shown in Figure 4.5a giving rise to the low formation energy of F_O. The 1- charge state in this case does not act as an acceptor but the extra electron is instead donated to the conduction band. Interstitial fluorine (F_i) was found in the calculations to distort from the 'perfect' interstitial site towards a lattice oxygen site causing a displacement of the oxygen (Figure 4.5b). Figure 4.5b also shows that the electron density is highly localised in a *p*-orbital on the F_i and on the two opposing O *p*-orbitals, thereby trapping charge. This defect was found to be a *very* deep donor state as the 1+/0 transition occurs ~2.30 eV below the CBM. At Fermi energies above the CBM, V_{Sn} begins to compensate F_O (~1.7 eV

above the CBM) thus negating the extra electrons and trapping the Fermi level at this point. Another species, the fluorine substitutional-interstitial pair which have been postulated theoretically at high concentrations and seen experimentally via simple changes in lattice parameters has also been calculated alongside F_O and F_i.[214,217,260] The [F-F]$_O$ defect possesses a relatively high formation energy and acts as an ultra deep donor with the 1+/0 charge state lying 3.2 eV below the CBM.

Under *Sn-Poor/O-Rich* conditions, the formation energy of F_O is raised and those of F_i and V_{Sn} are lowered. Under these conditions, the F_i^- and F_O^+ defect states cross at $\sim$ 0.03 eV above the CBM trapping the Fermi energy at this point; this 'self-compensation' mechanism has been seen to occur in anatase TiO$_2$ also[229]. At higher Fermi energies, V_{Sn}^{4-} crosses the F_O^+ line at $\sim$ 0.23 eV above the CBM potentially causing further compensation. The formation energies of the [F-F]$_O$ defect charge states remain the same over the chemical potential range and the neutral charge state occurs at a lower formation energy than F_O under *Sn-poor/O-rich* conditions.

The middle plot in Figure 4.4 represents the realistic growth conditions under APCVD at a temperature of $\sim$900 K and a pressure of 1 atm. These conditions lie somewhere between the extremes of the chemical potentials discussed previously and as such, the defect landscape transitions accordingly. The crossing point of the F_O^+ and F_i^- defect states now occurs at $\sim$ 0.55 eV above the CBM (shown by the dashed black line at E_F=4.15 eV) and above this point, compensation occurs.

In addition to the DFT calculations, we have calculated the partial charge densities for both the F_O^+ and the F_O^- defect charge states. Here it is shown that the electron density is delocalised when the F substitutes an oxygen and is localised when F is in an interstitial position. F_i^- also displays the relatively sizeable lattice distortion caused by the localisation of two electrons on the F atom and two adjacent O atoms. This depicts interstitial F as a deep acceptor (F_i^-), leading to the decrease in mobility

seen when self-compensation occurs for a Fermi level of $\sim$4.1eV above the valence band maximum.

The inset of Figure 4.8 (see section 4.3.3) shows the absorption coefficient, α, which is calculated from the transmission data and film thickness. The optical gap is determined to be 4.2 eV from linear extrapolation of α^2 versus photon energy. Accounting for the valence band dispersion, the Fermi level is found to be E_F = 4.1 eV above the valence band maximum (VBM) corresponding to the point where F_O^+ begins to be compensated by F_i^- which is in excellent agreement with the theoretically calculated value of 4.15 eV at 900 K, 1 atm. The charge state of F_i in this regime is $Z_A = -1$, justifying the initial assumption applied in the transport model and the compensation ratio of $K = 0.48$.
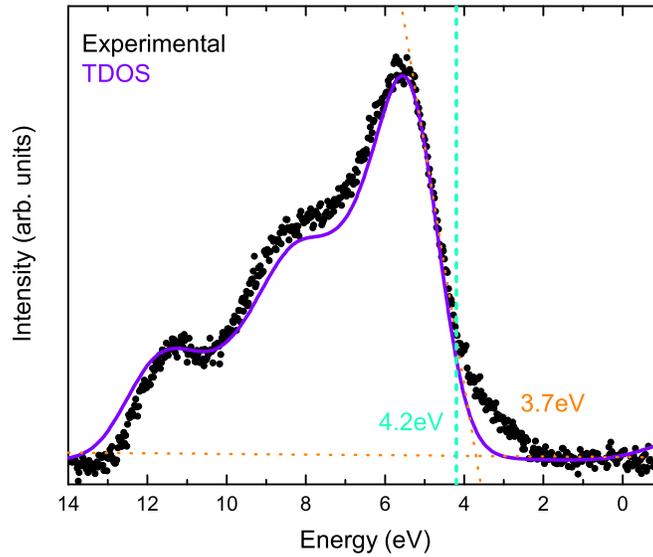
Figure 4.6: Experimental and theoretical total density of states (TDOS) for the valence band spectrum of FTO. The feature at $\sim$ 3 eV is associated with Sn(II) lone-pair distortion[261], which forces splitting of the energy levels according to the revised lone pair-model[262]. Because this represents a sub-stoichiometric coordination of the Sn, this is not predicted in the calculation.

Because the substrate is made from soda lime glass, the optical transmission spectra may be affected by absorption from the substrate around 4 eV, very close to the absorption edge of SnO_2. To give confidence in the measured optical gap, the valence band

maximum to Fermi-level position was measured using XPS also (Figure 4.6). Despite the material being a degenerate semiconductor, there are not enough occupied conduction band states to see conduction band emission (hard XPS may assist in this as the *s* states of the CBM will be relatively stronger. However, no HAXPES valence band spectrum was recorded for FTO in this thesis). Cross-section effects may contribute to the reduction of the intensity of this state. However, due to its highly conducting nature, we expect no charging issue and so the zero of the binding energy scale should align with the Fermi-energy. A linear extrapolation to the leading valence band edge shows a valence band onset at 3.7 eV. However, this does not account for the resolution of the analyser, or the lifetime of the states. Instead, the theoretical total density of states (DOS) after broadening and cross-section correction is superimposed. Because the DOS is calibrated for the VBM at 0 eV, a shift of 4.2 eV was required to match it to the data. This indicates a VBM to Fermi-level gap of 4.2 eV (XPS samples all occupied electronic states so no dispersion correction was required as was applied to the optical data). This is in excellent agreement with the gap determined from the optical data, and the theoretically determined Fermi-level pinning point.

4.3.3 CORE-LEVEL XPS AND OPTICAL REFLECTIVITY

In order to probe experimentally for evidence of F_i, high-resolution core-level XPS spectra of the Sn $3d_{5/2}$, O 1s and F 1s regions were recorded for degenerately doped FTO ($n = 4.27 \times 10^{20}$ cm^{-3}). All of these spectra were recorded after a low energy Ar$^+$ sputter to remove surface contamination. The level of contamination was monitored by taking survey spectra between sputter cycles. A noticeable reduction of a high binding energy component of the F 1s was also observed after sputtering. We attribute this to surface contamination associated with fluorine bonded to carbon, consistent with the large shift to higher binding energy seen for fluorocarbon species elsewhere[122,263,264].

Heavily-doped wide band-gap semiconductors such as TCOs display almost metallic-like properties in the sense that they possess a large density of free carriers. This needs to be considered in the fitting of the core-level spectra, but is often ignored. At the high doping levels ($n > 10^{20}$cm^{-3}) present in our samples, plasmon loss features are commonly observed which manifest as a high binding energy component, representing the fraction of photoelectrons that have lost energy to the collective excitations of the free electron gas during photoemission from the material. However, interpreting these loss peaks in XPS is notoriously difficult due to their complex nature, with many different approaches having been employed[265].
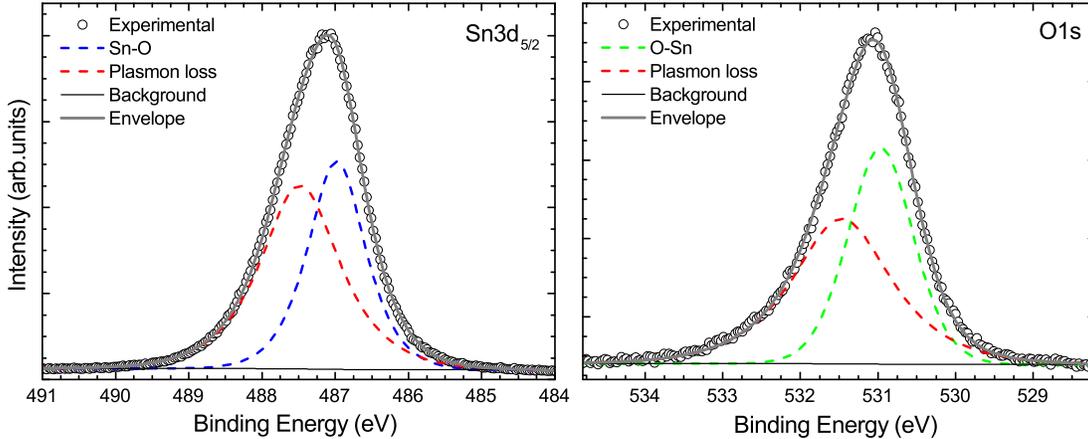
Figure 4.7: XPS spectra of the Sn3d$_{5/2}$ and O1s core levels of FTO (n= 4.27×10^{20}cm^{-3}) measured after argon plasma sputtering to remove surface contaminants. Two peaks are used to fit the data in each spectrum, a low binding energy component representing the un-attenuated peak and a high binding energy component representing the energy loss of the core-level to the collective free carrier gas.

Here we employ the fitting procedure of Egdell *et al.*[105,106] who base their analysis on the Kotani-Toyazawa screening model[107]. A comprehensive discussion regarding the merits and drawbacks surrounding this approach can be found in section 2.1.3, and in the literature[109,266]. Using this approach, both the Sn 3d$_{5/2}$ and O 1s core-level peaks (figure 4.7) are fitted using two symmetric Voigt functions, one component at lower binding energy for the photoelectrons with no energy loss to plasmons, and plasmon-loss component at higher binding energy. A Shirley background is also used

in the fitting[141].

In order to achieve a meaningful fit to the data, the energy separation between the plasmon loss peak and the no loss component is required. This separation is determined by the free carrier plasmon frequency. The surface plasmon frequency can be determined from high resolution electron energy loss spectroscopy (HREELS) and dielectric theory simulations[267,268], or the bulk plasmon frequency from infrared reflectivity measurements and modelling[269]. IR reflectivity spectra of the FTO ($n = 4.27 \times 10^{20}$ cm^{-3}) is displayed in figure 4.8.
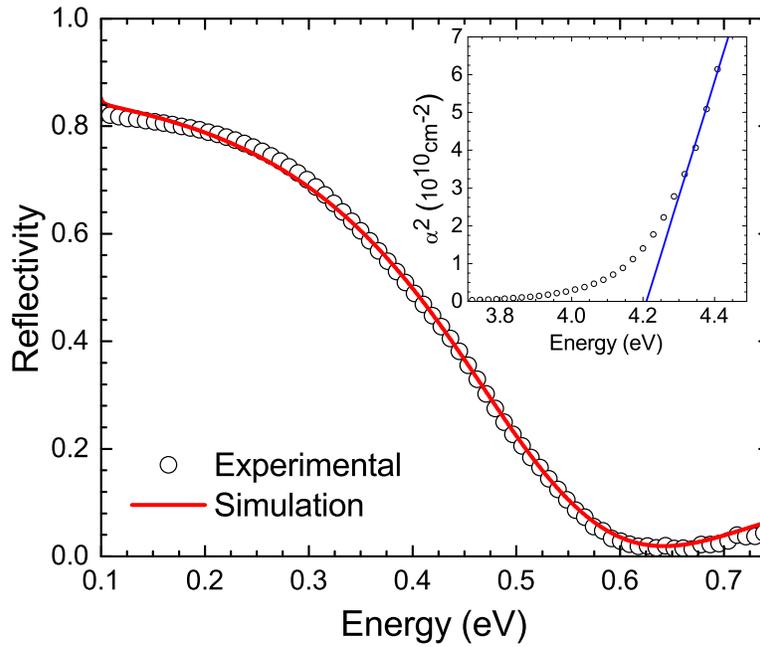
Figure 4.8: IR reflectivity data (circles) and optical model simulation (solid line) of FTO ($n = 4.27 \times 10^{20}$cm^{-3}) deposited on a glass/SnO$_2$/SiO$_2$ substrate. The inset plot displays optical absorption data with a linear extrapolation estimating the absorption onset. This shows an optical gap of 4.21 eV, corresponding to a Fermi level position 4.10 eV above the VBM.

The transfer matrix method-simulated reflectivity spectrum seen in figure 4.8 allows for the determination of the plasmon energy ω_p. The extracted plasmon energy is $\omega_p = 0.50$ eV, with a corresponding damping parameter of $\gamma_p = 77.0$ meV. The equation for the plasmon energy is given as $\omega_p = \sqrt{\frac{ne^2}{m^* \epsilon_0 \epsilon(\infty)}}$ with the surface plas-

mon energy varying only by a factor of $\sqrt{\frac{\epsilon(\infty)}{\epsilon(\infty)+1}}$. Using a relative high frequency dielectric constant of 3.9, this results in a surface plasmon energy of $\omega_{sp} = 0.45$ eV as determined from the simulated plasmon energy. As emitted photoelectrons originate from up to 10 nm from the surface, the separation between screened and unscreened components is expected to lie in the region of 0.45 eV $\leq \Delta E_p \leq$ 0.50 eV.

Utilising this information, the fitting procedure for the Sn $3d_{5/2}$ and O $1s$ core levels in figure 4.7 involved simply constraining the no loss peak to plasmon loss peak energy separation to the determined plasmon energy range and allowing parameters to be optimised in the fitting procedure in order to achieve the best least squares fit. As can be seen from figure 4.7 there is excellent agreement between the fit and the experimental data. Both the Sn $3d_{5/2}$ and O1s peaks display sizeable plasmon loss components at 0.5 eV higher than the no loss peak. Attenuated peaks are situated at 486.9eV and 530.9eV for Sn $3d_{5/2}$ and O $1s$ respectively, in good aggreement with other reported binding energy values[255,270,271].

The full width at half maximum (FWHM) of the core-line components for the Sn and O are both under 1.0 eV. The full-widths of the high-binding energy components are broader than this owing to the finite plasmon lifetime broadening which has to be considered in addition to the natural line width and instrumental broadening of the core level peak. The plasmon loss peaks display greater Lorentzian character than the low binding energy peaks. It is evident when the plasmon loss mechanism is accounted for in the XPS fitting that only a single Sn-O chemical environment can be discerned in the Sn $3d_{5/2}$ and O $1s$ spectra. This is expected given the bonding structure of SnO_2. Whilst a Sn-F bond peak could be expected in the Sn $3d_{5/2}$, the fluorine content in these samples is low and so it is not possible to distinguish it here in the presence of the strong Sn-O signal. The asymmetric peak shape of the tin and oxygen regions is commonly seen in reports on FTO[205,207,220,221,272], although they are

very rarely associated with plasmon loss events taking place.

Turning attention now to the F $1s$ core level region, the same constraints are applied as with the Sn $3d_{5/2}$ and O $1s$ core levels, the data for which is shown in figure 4.9. A much reduced fluorocarbon species is still present, which is likely a result of the low sputtering energy used (as not to damage the structure of the FTO) not fully removing the surface carbon, although carbon incorporation into the bulk cannot be ruled out. Taking this into account, a single no loss core-level peak and associated plasmon loss peak pair (and additional high binding energy contaminant peak) provided an extremely poor fit to the data. The shoulder component could not be fitted well under these constraints. This suggests there is another species of fluorine present within the material, in addition to the expected substitutional fluorine. In order to fit the F $1s$ spectra with two fluorine species present, a similar fitting procedure was used to that of the Sn $3d$ and O $1s$ core-levels. However, in the case of the fluorine it is acknowledged that both fluorine core-lines will produce plasmon loss features and in fact it is possible to further constrain the area ratios of the loss feature to its corresponding core-line because electrons originating from these two fluorine species will experience the same screening from the free electron gas. Carbon at the surface will not have an associated plasmon loss as electrons originating from the surface will not interact with the free electron gas in the bulk.

The loss features are again constrained to $\omega_{sp} \leq \Delta E_p \leq \omega_p$ from the respective core-line peaks. The substitutional peak and the peak labelled interstitial F in figure 4.9 are constrained to have the same FWHM as each other. The two loss peaks are constrained to have the same FWHM as each other, but it is allowed to differ from that of the no loss components. The binding energy positions of the substitutional and interstitial peaks are determined to be 684.9eV and 685.7eV respectively. As expected, the loss peaks take on a more Lorentzian character with a larger FWHM than the other components.
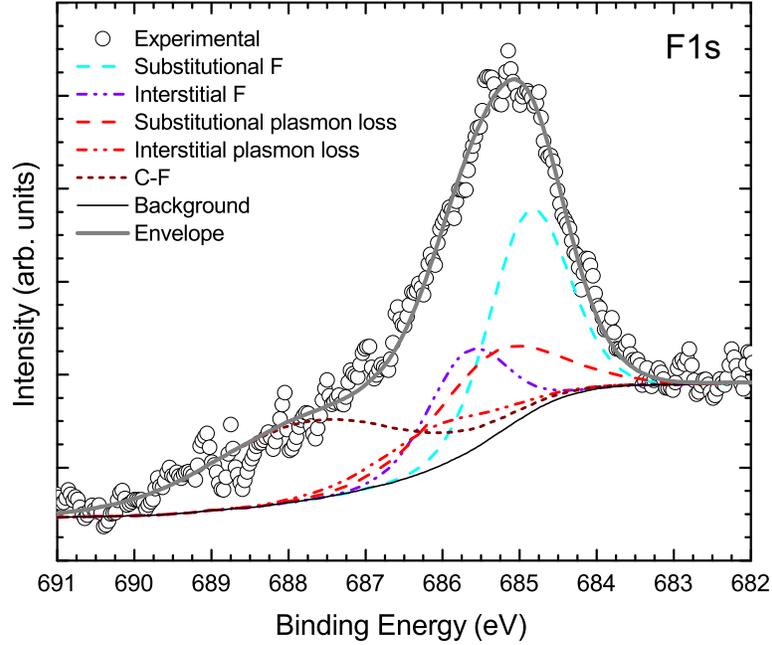
Figure 4.9: XPS spectra for the F 1s core level of FTO (n= 4.27×10^{20}cm^{-3}) measured after argon plasma sputtering to remove surface contaminants. Four peaks were used to fit the data representing substitutional fluorine (F$_O$), interstitial fluorine (F$_i$), and two further symmetric peaks representing the energy loss of each of the core levels to the collective free carrier gas. Additionally a small fluorocarbon species is seen at higher energy.

The fitting is again in good agreement with the data. Multiple data sets from a range of FTO coatings with varying fluorine content have been fitted using the same fitting parameters and procedure. The fitting parameters for the XPS data from these other coatings are consistent with the ones from the spectra shown (see Figure 4.10).

The peak areas of the substitutional and interstitial F 1s peaks are extracted from the fit and the ratio of the two calculated. This ratio for this particular sample is found to be $\frac{A_{int}}{A_{sub}} = 0.47$. For six samples analysed with 1.8×10^{20} cm$^{-3} < n < 5.5 \times 10^{20}$ cm^{-3}, this peak area ratio is found to lie in the range 0.47 to 0.53 (see Figure 4.10). The peak areas are representative of the concentration of the given species of the material present in the sample, and hence it is possible to deduce the fraction of the fluorine interstitial is ∼0.5 of the substitutional fluorine incorporated into the material. This ratio is remarkably close to the compensation ratio of $K = \frac{N_A}{N_D} = 0.48$ found from transport
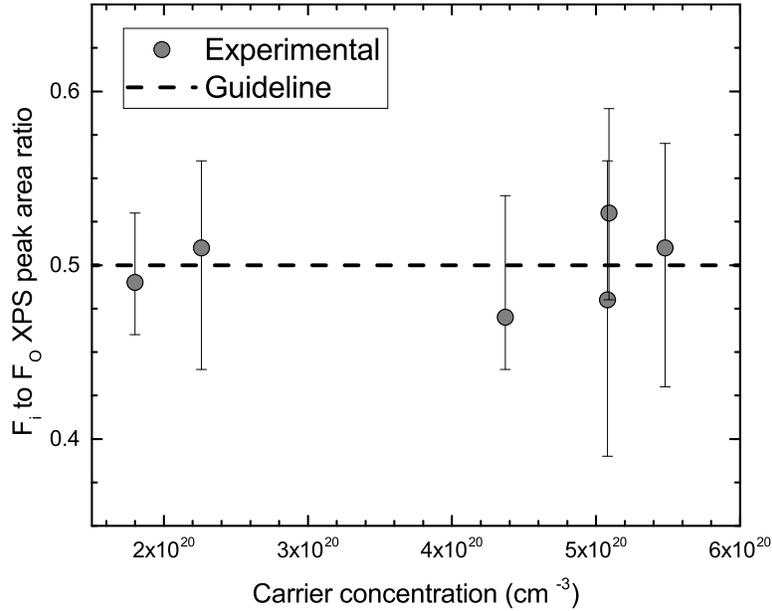
Figure 4.10: Ratio of interstitial to substitutional F as a function of free carrier concentration for a number of FTO samples as determined using XPS and Hall effect.

modelling in Fig. 4.1. Therefore, the experimental XPS evidence strongly supports the hypothesis of a compensating defect, and is consistent with the proposed defect of the fluorine interstitial determined as the lowest formation energy compensating defect by the DFT. This finding is also supported by higher energy photoemission measurements (6450 eV), although contributions from Na-F contamination, likely originating from the glass, also contribute heavily to the line shape, see Figure 4.11. Despite this, the narrower line shape demonstrates the asymmetry caused by multiple peaks better in the HAXPES data.

The fluorine interstitial acting as a singly charged compensating acceptor has been a well established hypothesis [196,212,213,217,260] for FTO, with researchers even providing similar evidence from XPS measurements such as Suffner *et al.*[214] on FTO nanoparticles. However, Suffner *et al.* did not include the effects of plasmon losses in their fitting and no transport properties of the material were reported. To our knowledge, nobody has yet made the connection, quantitatively or otherwise, between the trans-
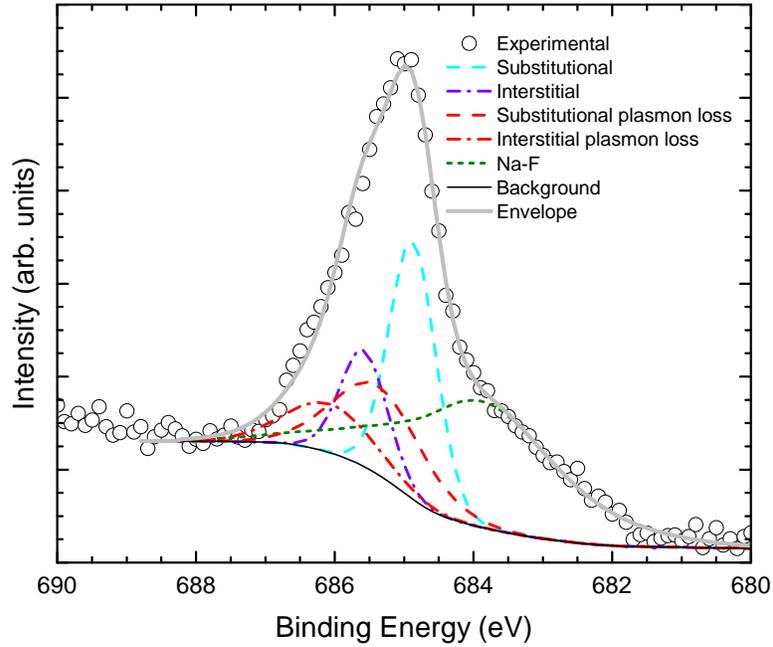
Figure 4.11: HAXPES spectra of F 1s core level region for FTO (n=4.27$\times 10^{20}$ cm^{-3}).

port compensation behaviour and the XPS chemical analysis that we have performed for FTO in this work. In addition to this, the DFT defect chemistry analysis has been clearly instructive in both the transport modelling and XPS fitting, proving to be a powerful tool in the determination of the likely defect species and in understanding the charge state of the defect which directly relates to the transport modelling. It is also worth reiterating that the effects of plasmon losses in FTO as seen in XPS analysis have been scarcely touched upon in the literature. This could easily lead to the misassignment of spectral features.

From the above, it can be inferred that the carrier density of these FTO films is heavily compensated, with the measured free carrier concentration being about one third of the total fluorine incorporation - for every two substitutional F donors, there is approximately one interstitial F acceptor, resulting in roughly one free electron for every three F atoms. Or, more precisely, for our range of substitutional to interstitial F 1s area ratios of 0.47 to 0.53, there are on average between 2.9 and 3.1 F atoms per free

electron. Indeed, this is supported by comparing the F contents estimated from XPS with the free electron densities from Hall effect. From XPS, the F content in the FTO films is estimated to be in the range 0.7-1.7 atomic % (but with considerable uncertainty in the absolute atomic % values). Considering the atomic density of SnO_2 of 8.4×10^{22} cm^{-3}, this corresponds to F concentrations in the range of around 5.9×10^{20} to 1.4×10^{21} cm^{-3}. Comparing this to the Hall carrier concentration range of $1.8-5.5 \times 10^{20}$ cm^{-3} reveals between 2.5 and 3.3 F atoms per free electron, in agreement with the expectation of the F concentration being about three times the carrier density.

While the XPS provides a reasonably strong indication of the presence of interstitial fluorine, the physical interpretation is muddied somewhat due to the presence of plasmon losses etc. To provide additional evidence of this finding with greater accuracy than XPS and without the limitations of the surface sensitivity of XPS, ToF SIMS was used. F-ion implanted standards were used in combination with profilometry (for depth calibration) to obtain the atomic concentration of F. SIMS was performed on typical FTO samples with a Hall carrier concentration of $n = 4.05 \times 10^{20} cm^{-3}$ and $n = 4.27 \times 10^{20}$ cm^{-3} (corresponding to the sample from which the data was recorded in figure 4.9). Figure 4.12 shows the SIMS data of the samples and a F-implanted standard. The y-axis in Figure 4.12 was scaled so that the area under the curve corresponding to the fluorine ion implanted film is $1 \times 10^{16} cm^{-2}$. This enabled the fluorine concentration as a function of depth to be determined. The depth-averaged total concentration of F determined by calibrated SIMS was found to be [F] = $(1.08\pm0.11) \times 10^{21}$ cm^{-3} and [F] = $(1.06\pm0.11) \times 10^{21}$ cm^{-3} respectively. The uncertainty is estimated by considering the uncertainties in the thicknesses obtained from profilometry (±25 nm), the implantation doses ($\pm2\%$) and the possible matrix effects due to F concentrations in excess of 1 atomic %. This indicates about 2.5-2.7 F atoms per free electron. This confirms that the total F content is approximately 3 times greater than the free electron density,

n. Note that the FTO spectra in figure 4.12 display a wavy nature, which is due to the non-uniformity in the doping method used during deposition. Uniform doping is assumed in the transport model for simplicity, and because the analysis of the SIMS is based on an integration over depth, the non-uniformity will not have a bearing on the total fluorine content.
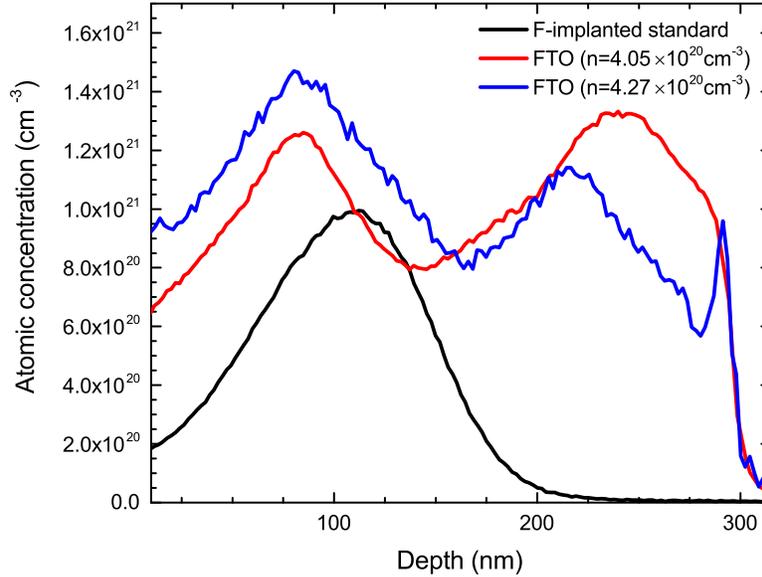
Figure 4.12: Calibrated secondary ion mass spectrometry (SIMS) data of the atomic fluorine signal from a typical FTO sample (n_{Hall}= 4.05×10^{20}cm^{-3} and n_{Hall}=4.27×10^{20}cm^{-3} corresponding to the XPS data) and a fluorine-implanted standard (10^{16}cm^{-2}). The x and y-axes have been calibrated to depth (nm) and concentration(cm^{-3}) respectively.

This additional F content has the effect of reducing the achievable mobility from a maximum of around $\mu_{theory} = 120$ cm$^2/V \cdot s$ to under $\mu_{experiment} = 40$ cm$^2/V \cdot s$ for a carrier density of $n \sim 4 \times 10^{20}$ cm^{-3}. This has strong implications for the material performance. With this in mind, it is demonstrated that FTO possesses intrinsic limitations on its mobility and carrier density due to this self compensation. Although these materials display excellent transparency and conductivity properties, in order to improve industrial-scale TCOs, alternative dopants need to be identified and their defect chemistry explored in order to find dopants which do not exhibit this

self-compensating phenomenon inherent to fluorine doping of SnO_2. One example of such a novel dopant (which is the subject for chapter 5), is Mo in In_2O_3 which gives higher mobilities than Sn in In_2O_3 (see Ref. 88). Another example is Ta doping of SnO_2, where for films grown by pulsed laser deposition, mobilities as high as 83 $cm^2/V \cdot s$ have been reported for carrier densities of around 3 $\times 10^{20}$ cm^{-3}[273]. Such values are consistent with the transport modelling for the uncompensated case shown in Fig. 4.1, but have yet to be realised using a scalable deposition method. It is also noted that the mobility of molecular-beam epitaxy-grown Sb-doped SnO_2 with free electron density of 2.6×10^{20} cm^{-3} is limited to 35 $cm^2/V \cdot s$ even though calibrated SIMS indicates an Sb concentration of 2.8×10^{20} cm^{-3}, suggesting negligible compensation from Sb-related defects[274]. For the Sb-doped case, other mobility-limiting mechanisms may be present, such as hybridization of Sb-orbitals with the Sn-dominated conduction band states, leading to increased electron effective mass and reduced mobility.

4.4 Conclusion

Heavily *n*-type fluorine-doped tin dioxide ($n > 1 \times 10^{20}$ cm^{-3}) deposited on soda-lime glass via APCVD has been shown to exhibit inherent self-compensation, limiting the achievable free electron density, mobility and resulting conductivity. Modelling of mobility versus carrier density data from Hall effect measurements indicates ionized impurity scattering dominates and the mobility is limited to $<$40 cm^2/V·s by the presence of acceptors, with a compensation ratio of $K = 0.48$. Density functional theory formation energy calculations determined interstitial fluorine in the -1 charge state to be the lowest formation energy acceptor defect for degenerately doped FTO. Core-level XPS measurements and analysis were performed on FTO, including paying particular attention to the modelling of plasmon loss components of core level lines which result from energy loss to the collective excitations of free carriers in degenerately-doped

semiconductors. A high binding-energy shoulder component was found in the F $1s$ core level-region and attributed to interstitial fluorine, F_i. This component has half the intensity of that due to substitutional donor fluorine, F_O, consistent with the determined compensation ratio. This quantitative connection between fluorine chemical analysis and transport modelling has not previously been made. Hence, new evidence of the fluorine interstitial being the defect responsible for FTO falling well short of the theoretical ionized impurity scattering-limited mobility of >100 cm^2/V·s has been provided. Quantitative analysis on F concentrations of samples made through Hall effect, XPS and SIMS provides further evidence of compensation in FTO. With this in mind, it is clear that novel dopants are required to avoid these inherent limitations to conductivity seen in FTO.

5

Replacing Sn:In$_2$O$_3$ with Mo:In$_2$O$_3$ - Resonant doping in oxide semiconductors

Films were grown via AACVD by S. Sathasivam in the group of I. Parkin at University College London. S. Sathasivam also performed SEM measurements. Density functional theory calculations were carried out and analysed by B. A. D. Williamson in the group of D. O. Scanlon at University College London. Undoped-In$_2$O$_3$ films were sputtered by T. J. Featherstone (University of Liverpool) and ion implanted at Surrey Ion Beam Centre by N. Peng. Secondary ion mass spectrometry measurements were performed by M. Farnworth at Pilkington Technology Centre and data analysed by the author. HAXPES measurements were performed by the author, T. J. Featherstone, H.

J. Edwards (University of Liverpool) and A. Regoutz (University College London), at IO9 beamline, Diamond light source with the expertise of beamline scientists D. Duncan and T-L. Lee. The HAXPES data was analysed by the author. XAS measurements were performed at beamline 23-ID of the National Synchrotron Light Source II, by Z. W. Lebens Higgins in the group of L. F. J. Piper at Binghamton University, who also performed the analysis of this data set. The results presented here are published as:

J. E. N. Swallow, B. A. D. Williamson, S. Sathasivam, M. Birkett, T. J. Featherstone, P. A. E. Murgatroyd, H. J. Edwards, Z. W. Lebens-Higgins, D. A. Duncan, M. Farnworth, P. Warren, N. Peng, T-L. Lee, L. F. J. Piper, A. Regoutz, C. J. Carmalt, I. P. Parkin, V. R. Dhanak, D. O. Scanlon and T. D. Veal, Resonant doping for high mobility transparent conductors: the case of Mo-doped In_2O_3. Mater. Horiz., 2019, DOI: 10.1039/C9MH01014A.

5.1 Introduction

The global TCO market is predicted to have an annual growth rate exceeding 17%, reaching \$8.04 billion by 2022[23]. The demand for higher conductivities and the volatile and high price of indium has led to the search for new TCO materials.[275,276] However, as shown in Fig. 5.1, alternative indium-free transparent conductors such as graphene, carbon nanotubes and oxide/metal/oxide layers have failed to surpass or even reach ITO's performance[276-278]; ITO continues to account for 60% of the transparent conductor market and 60% of global indium use.

Given this failure of the alternative transparent conductors, it is remarkable that In_2O_3 doped instead with molybdenum, giving double the conductivity of ITO, has
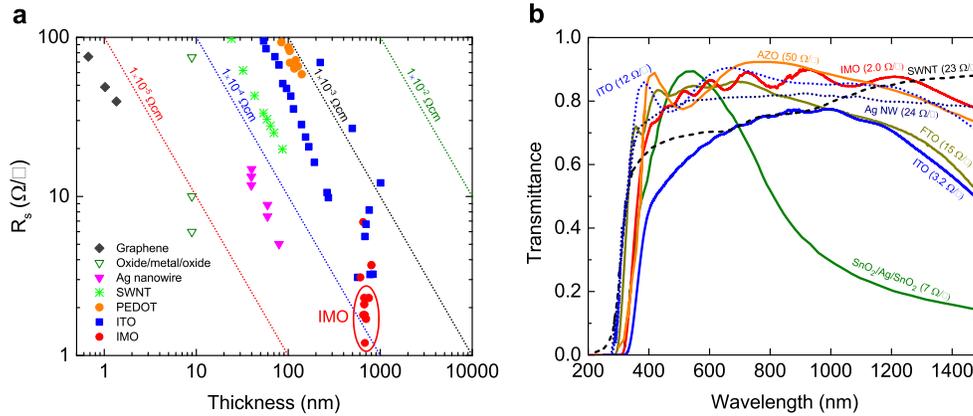
Figure 5.1: Sheet resistances and optical transmittances of transport conducting electrodes. a Sheet resistance R_s as a function of film thickness for different transparent conducting films [276]: PEDOT-PSS [279]; ITO films [280,281]; SWNTs [280]; Ag nanogrid [280]; oxide/Ag/oxide films [282]; and graphene [280]. This Figure is based on Ellmer's [276] with data points for IMO added and highlighted by the red ellipse. The IMO films have the lowest sheet resistances. The dotted diagonal lines denote constant resistivity values. **b** Spectral transmission of different transparent electrode films: single walled carbon nanotubes (SWNT) ($23\,\Omega/\square$), Ag nanowires ($24\,\Omega/\square$), aluminium-doped ZnO (AZO, $50\,\Omega/\square$), ITO ($12\,\Omega/\square$),[282] $SnO_2/Ag/SnO_2$ ($7\,\Omega/\square$),[283] compared with our measurements of IMO ($2\,\Omega/\square$), ITO ($3.2\,\Omega/\square$), and fluorine-doped SnO_2 (FTO, $15\,\Omega/\square$). All materials which approach IMO's sheet resistance have much inferior infrared transparency.

not been adopted by industry in spite of first being suggested over 15 years ago. Reports of Mo-doped In_2O_3 have been sporadic and have contained little insight into the origins of its superior performance. In fact, reports of IMO showing promising performance seem to have been treated merely as a curiosity, and thus IMO was never identified as a research direction deserving of further detailed investigation. Perhaps this lack of research and dearth of understanding has prevented industry overcoming its inertia to substitute an established high performing material for a significantly better one. Here, a comprehensive explanation is presented of why IMO outperforms ITO. This work opens the pathway for immediate commercial adoption of IMO, and for the development of other high performing TCOs.

5.2 Experimental and Theoretical Methods

Material characterisation

Films were deposited on SiO_2 coated glass soda lime glass (NSG) via AACVD. This methodology is described elsewhere.[88] Film durability was tested and compared to that of AACVD ITO and commercial ITO coated glass which proved comparable. IMO film thicknesses ranged from 670–800 nm while ITO films ranged from 700—1600 nm. Film thickness was determined by cross sectional scanning electron microscopy (SEM) using a JEOL JSM-6301F field emission SEM at an accelerating voltage of 5 kV. Cross-sectional images were taken from unmasked samples. X-ray diffraction (XRD) measurements were performed using a Rigaku SmartLab instrument, and patterns collected over $20° < 2\theta < 65°$ with a step size of $0.02°$. The patterns shown in figure 5.2 confirm the presence of In_2O_3 of the cubic bixbyite phase. The films grew with a (100) preferred orientation, as indicated by the strong peak from the fourth order reflection from the (100) planes seen in figure 5.2.

Hall effect measurements (for the determination of resistivity ρ, charge carrier mobility μ and free carrier concentration n) were performed at room temperature with a Keithley 2182A nanovoltmeter and 6220 current source at a maximum field strength of 0.8T in the van der Pauw geometry. Time of flight secondary ion mass spectrometry (ToF-SIMS) was performed on a IonTof TOF-SIMS 5 instrument to determine the dopant concentration as a function of depth. Pulsed Bi_3^+ analysis ions were used to bombard the sample surface while a 1 keV Cs source was used as the sputtering beam. Analysis was performed on a $50.8 \times 50.8 \mu m^2$ sputter area. The SIMS data was calibrated using ion implanted standards, implanted at Surrey Ion Beam centre. Undoped sputtered In_2O_3 was ion implanted with Sn at 1.8 MeV (corresponding to a depth of $\sim$406 nm), and Mo at 1.5 MeV (corresponding to a depth of $\sim$402 nm). Doses
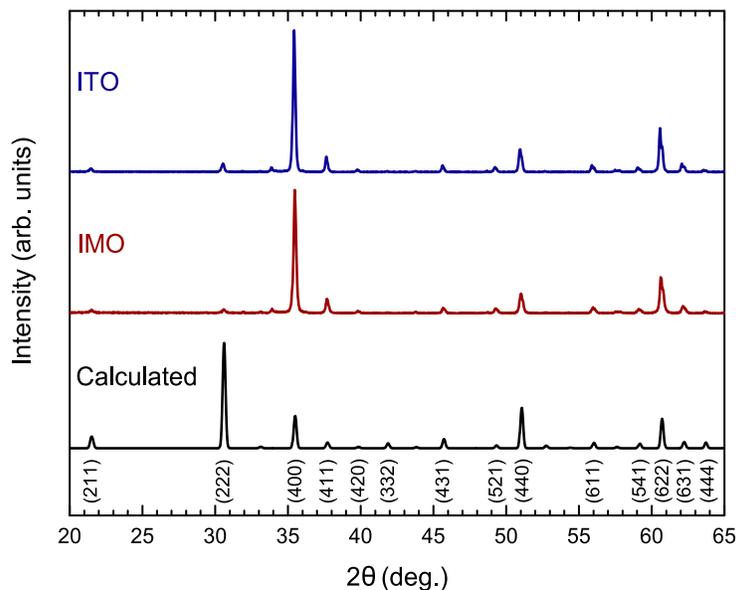
Figure 5.2: X-ray diffraction patterns of ITO and IMO.

of 3.57×10^{14}cm^{-2}, 3.57×10^{15}cm^{-2}, and 3.57×10^{16}cm^{-2} were selected here. Mo was relatively difficult to detect at the dopant concentrations in the films in SIMS, with negligible signal being detected from our IMO films for Mo$^+$ and CsMo$^+$ ions. However, a clear but noisey signal could be obtained from CsMoO$^+$ and CsSnO$^+$ ions, and so these were used to determine the elemental contents which were averaged over the depth of each film.

Infrared (IR) reflectance measurements were performed using a Bruker Vertex 70v Fourier transform infrared spectrometer. All measurements were performed at 45° angle of incidence, utilising the Pike Veemax II accessory, and S-polarising filter. All FTIR measurements were performed with a near-IR light source, CaF$_2$ beam splitter and DLaTGS detector. Transmission measurements were obtained with use of a Shimadzu UV-Vis-IR 3700 spectrophotometer with an integrating sphere detector. Hard x-ray photoelectron spectroscopy was performed at the I09 beamline at Diamond Light

source. Measurements were performed at both 6.45 keV and 2.15 keV for different depth dependence, and to probe the different orbitals through the variation of their photoionization cross sections. Photoelectrons were collected and analysed using a SCIENTA EW-4000 electron energy analyser. The sample was irradiated in grazing incidence geometry, that is an angle of $\sim 3°$ between x-ray beam and sample surface. The energy calibration and resolution of the system were determined from a polycrystalline Au foil. The energy resolution was determined by fitting a Gaussian-Fermi function convolution to the Fermi edge, and was found to be 300 meV. Inverse photoemission spectroscopy measurements were performed using a PSP Vacuum Technology BaO cathode dispenser electron source and an isochromat NaCl photon detector, both at $45°$ to the sample normal. Energy calibration was performed using the known energy position of the lowest occupied molecular orbital of multilayer C_{60} deposited in situ and the spectrometer resolution determined from a polycrystalline Au foil was 1.27 eV. The IMO and ITO samples were annealed to 300°C for 2 hours in preparation for IPES measurement.

Theoretical and modelling methods

Hybrid density functional theory (DFT) calculations were carried out on In_2O_3 using a plane-wave basis within the periodic code, VASP.[222,223,284,285] Screened non-local exchange-correlation density functional (HSE06) is used.[286] The projector augmented wave method (PAW)[287] was used to describe the core and valence electrons interaction. The band unfolding code BandUp was used to obtain a primitive cell representation of the band structure[288,289].

Sample plasma energies were extracted from IR reflectivity measurements utilizing a two-oscillator dielectric model,[290] (see equation 1.6) and transfer matrix method for optical simulation[148]. This methodology is described in detail in section 2.2.4. A four layer (vacuum/In_2O_3/SiO_2/soda-lime glass) stratified medium was used in the model,

with complete incoherence being assumed in the glass substrate and partial incoherence in the vacuum/In_2O_3 interface. The absorption spectra were calculated from the transmission spectra of the samples using the Beer-Lambert relation. Optical gaps were extracted from the absorption spectra using a method proposed by Hamberg[33], illustrated by Dolgonos *et al.*[147] which better accounts for degeneracy in materials compared to the more common Tauc analysis (see section 2.2.3). A carrier statistics model was used in order to relate optical gap and plasma energy through the density-of-states average effective mass, see equation 1.11.

5.3 Results and Discussion

5.3.1 TRANSPORT AND MATERIAL PROPERTIES
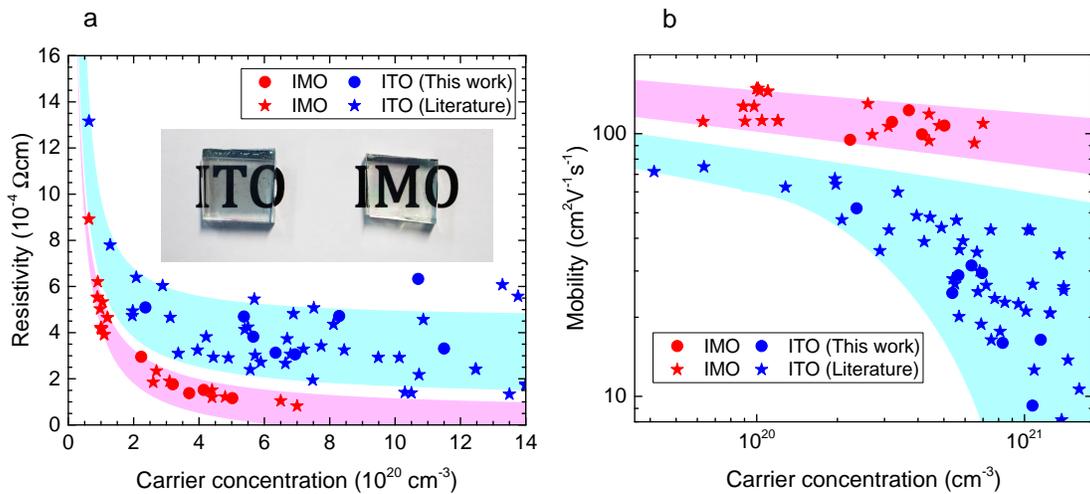
Figure 5.3: Transport properties. a Resistivity versus free carrier concentration for polycrystalline ITO and IMO films. Inset shows a photograph of ITO and IMO films. **b** Hall mobility versus free carrier concentration. Data from this work are circles. Data plotted as stars are taken from the literature for IMO[88,291–295] and ITO[88,296–301] films. In this carrier concentration regime, the electron mobility is limited by ionized impurity scattering.

Comparing Mo and Sn dopants in In_2O_3 (Fig. 5.3), it is found that higher mobilities

and lower resistivities are universally obtained with Mo. Fig. 5.3, which shows the Van der Pauw resistivity and mobility as a function of carrier concentration, includes data from chemical vapour-deposited films from this work as well as those from other studies of polycrystalline IMO and ITO. Film thickness varied from 670–800 nm for IMO and 700–1600 nm for ITO as determined by cross-sectional scanning electron microscopy (Fig. 5.4).
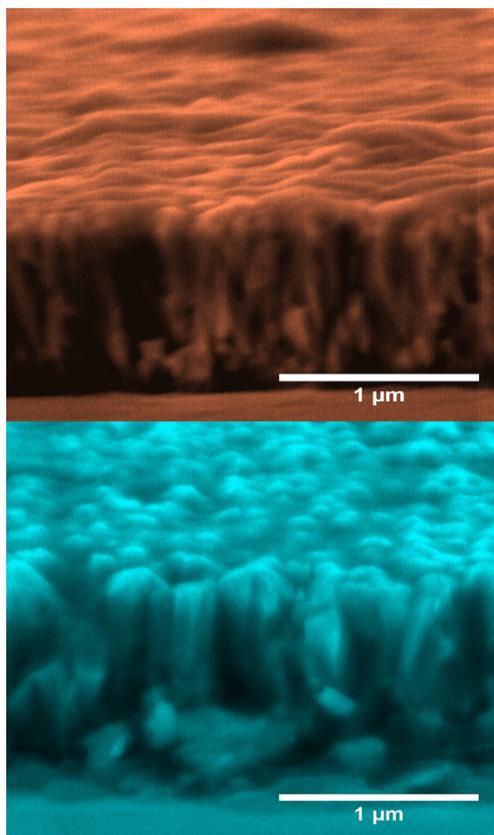
Figure 5.4: Typical cross-sectional SEM image for both IMO (top) and ITO (bottom), used to determine film thickness for the range of films in this study.

Calibrated secondary ion mass spectrometry (Fig. 5.5) indicates that the dopant concentrations are the same as the free electron densities, implying that within the uncertainty there is one free electron per Mo atom in IMO and Sn atom in ITO. The dopant concentrations averaged over the depth of each film are compared with the Hall effect

electron concentrations in the inset of the figure. Previously published Sn concentrations versus electron densities for ITO films are also shown for comparison. While many of the data points fall close to the dashed line corresponding to equal dopant and electron concentration, many of the ITO data points at high dopant concentration lie to the right of the line, indicating significantly higher dopant concentration than electron density. Therefore, while compensation can be present in ITO films particularly at very high dopant densities (over 10^{21}cm^{-3}), our SIMS and Hall effect results indicate that significant compensation due to excess density of dopant atoms is not present in our films.
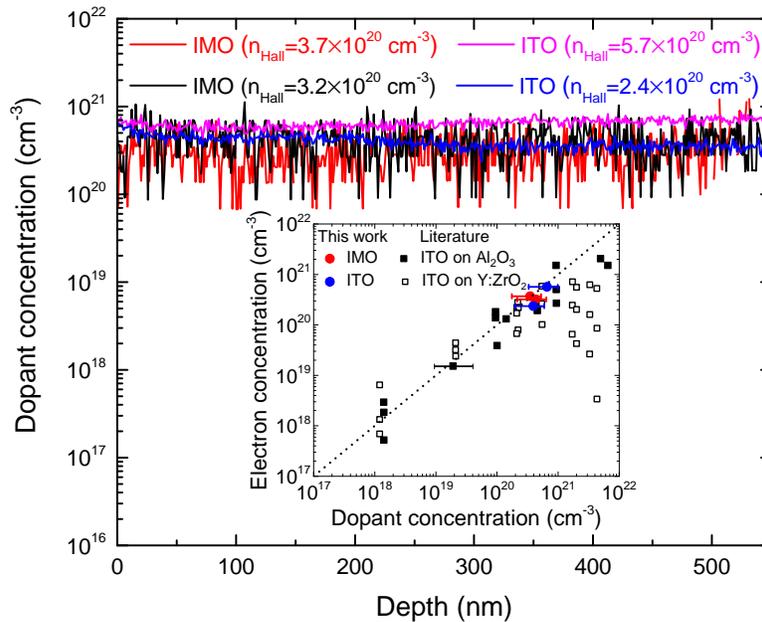
Figure 5.5: Dopant concentration versus depth for IMO and ITO films determined by secondary ion mass spectrometry. The Mo and Sn concentrations are calibrated using Mo and Sn ion implanted standards. The inset shows the dopant concentration from SIMS versus electron concentration from Hall effect measurement for our IMO (red) and ITO films (blue). This indicates that, to within the measurement uncertainty, there is one electron per dopant atom. Previously published data on ITO films is also shown for comparison.[302]

For IMO, a resistivity as low as 0.8×10^{-4} Ωcm has been achieved, almost half of the 1.4×10^{-4} Ωcm for ITO. This is enabled by IMO's extremely high mobilities displayed in Fig. 5.3b. Mo clearly outperforms Sn as a dopant, with IMO films having mobilities

of up to 150 cm²V⁻¹s⁻¹ compared with the 80 cm²V⁻¹s⁻¹ for ITO. Table 5.1 provides further details of electronic and optical properties. From all available results, a trend of twice the conductivity for IMO over ITO of the same carrier concentration is apparent. Therefore only half the amount of indium is required for IMO films with both better optical properties and improved conductivity over ITO, giving the potentially huge savings in materials and costs.

Table 5.1: Important physical parameters for the series of IMO and ITO films used in this study. These are carrier density (n), free carrier mobility (μ), resistivity (ρ) and sheet resistance (R_S) as determined by four point probe and Hall effect measurements, thickness (d) as determined by cross sectional SEM and confirmed by IR reflectivity modelling, optical transmission at 550nm (T_{550}), the Haacke figure of merit (Φ_H) determined from R_S and the optical transmission at $\lambda = 550$nm, and the plasma frequency from infrared reflection measurements and modelling.

	n (10^{20}cm^{-3})	μ $(\text{cm}^2\text{V}^{-1}\text{s}^{-1})$	ρ $(10^{-4}\Omega\text{ cm})$	R_S $(\Omega/\square)$	d (nm)	T_{550} (%)	Φ_H $(10^{-3}\Omega^{-1})$	ω_p (eV)
IMO	2.2	94.8	3.0	3.7	800	73.6	12.6	0.51
	3.2	110.8	1.8	2.3	750	72.2	16.7	0.63
	3.7	122.7	1.4	2.1	670	69.9	13.3	0.65
	5.0	107.3	1.2	1.7	700	67.7	12.0	0.69
	4.1	99.4	1.5	2.3	670	52.4	0.8	0.71
ITO	2.4	52.0	5.1	3.2	1600	57.7	1.3	0.61
	7.0	29.5	3.1	2.8	1100	54.6	0.8	0.82
	5.7	28.9	3.8	2.4	1600	51.9	0.6	0.70
	6.3	31.5	3.1	2.4	1300	59.5	2.3	0.82
	8.3	16.0	4.7	3.9	1200	62.8	2.5	0.87
	10.7	9.2	6.3	7.0	900	66.0	2.2	0.93
	11.5	16.4	3.3	3.7	900	68.8	6.4	0.93
	5.4	24.7	4.7	6.7	700	75.6	9.1	0.68

5.3.2 OPTICAL PROPERTIES

The main factor that limits the properties of TCOs is the trade-off between doping for increased conductivity and consequential decrease in optical transparency due to band tailing and the raised conduction electron plasma energy. Optical transmission spectra of IMO and ITO (Fig. 5.6a) show that IMO generally has greater transmission

at 550 nm (2.25eV), where the human eye is most sensitive (see also Table 5.1). Optical reflection and absorption spectra of IMO and ITO (Fig. 5.7) show similar optical band gaps (via the method of Dolgonos et al. [147]), yet the conduction electron plasma edge from infrared reflectivity lies at strikingly lower energies in IMO, greatly enhancing transparency (see also Fig. 5.6a). Plasma frequencies were extracted using a two oscillator dielectric model and the transfer matrix method for optical simulation. A four layer (vacuum/In$_2$O$_3$/SiO$_2$/soda-lime glass) stratified medium was used in the model, with complete incoherence being assumed in the glass substrate and partial incoherence in the vacuum/In$_2$O$_3$ interface. Due to the nature of specular reflection by non-flat as-grown films, there is a dip in the intensity of the reflection spectra. This phenomena has been observed previously and does not shift the plasma edge [303,304].
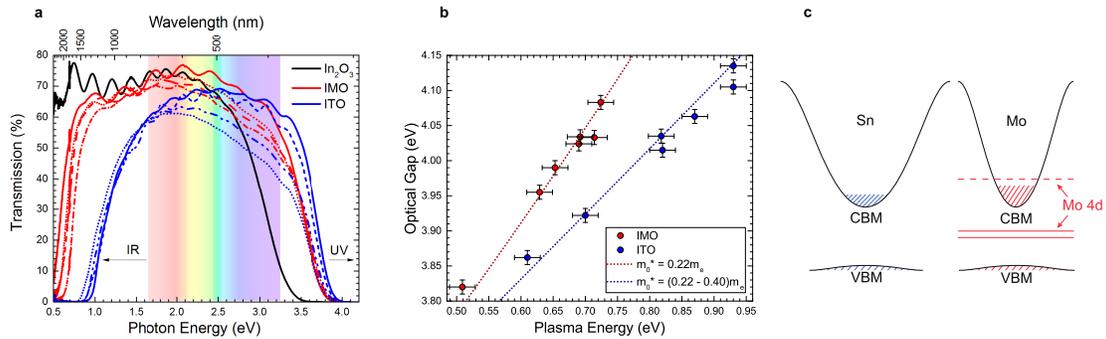
Figure 5.6: Optical transmission and associated properties. a Optical transmission of IMO and ITO films (on 3.6 mm-thick soda lime glass) with a range of carrier concentrations, and a nominally undoped In$_2$O$_3$ film, revealing greatly enhanced transparency in IMO due to plasma edges lying at lower energies. **b** Optical gap versus plasma edge energy of IMO and ITO films, the steeper gradient for IMO indicates a (highly desirable) reduced conduction band minimum effective mass with respect to that of ITO. Model curve lines are calculated using Fermi statistics with a density-of-states averaged effective mass. **c** Schematic representation of the band structure of ITO (left) and IMO (right) with the same carrier concentration. The IMO has a narrower conduction band dispersion with lower band edge effective mass than ITO and so a higher Fermi level is required to obtain the same number of states as in ITO. In IMO, the resonant and midgap Mo $4d$ bands are shown, respectively, by the dashed and solid red lines. No Sn state is shown in the ITO band structure schematic as the Sn $5s$ states mix with the states of the conduction band.

Plasma frequency relates to carrier density and effective mass as $\omega_p^2 \propto n/\langle m^* \rangle$ and so these data indicate an inherent difference in the effective masses of conduction electrons between In$_2$O$_3$ doped with Sn and Mo, diverging from the commonly accepted

rigid band model of doping.
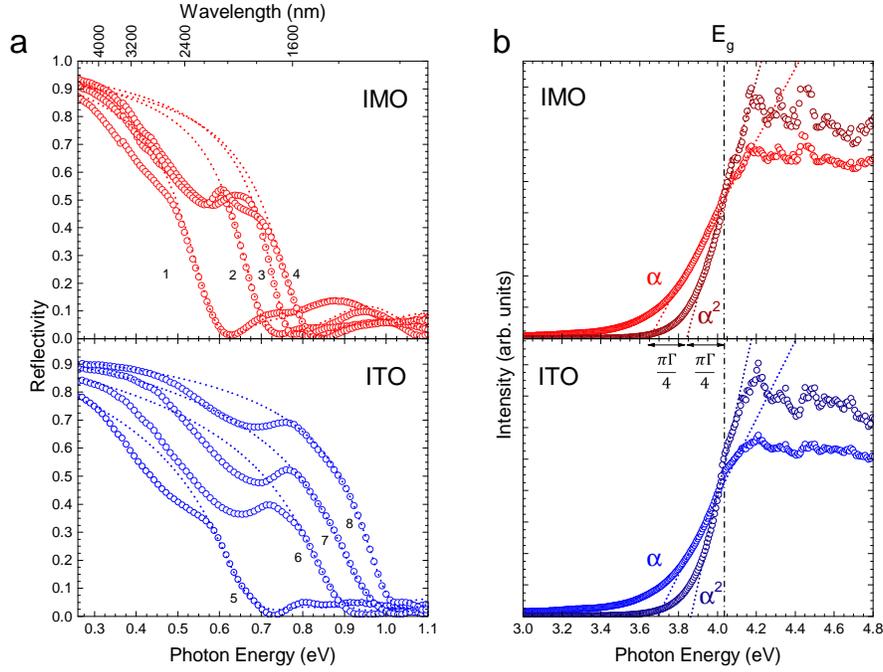
Figure 5.7: Optical measurements and data extraction methods. **a.** Infrared-reflectivity measurements as a function of photon energy for both IMO and ITO films with a range of carrier densities: (1) 2.23×10^{20} cm^{-3}, (2) 3.20×10^{20} cm^{-3}, (3) 5.02×10^{20} cm^{-3}, (4) 4.14×10^{20} cm^{-3}, (5) 2.36×10^{20} cm^{-3}, (6) 6.94×10^{20} cm^{-3}, (7) 8.29×10^{20} cm^{-3}, (8) 1.07×10^{21} cm^{-3}. Dotted lines represent model fit curves used to extract the film plasma frequency. A dip in each of the spectra compared with the model curves is associated with surface roughness. **b.** Typical absorption spectra for IMO (n=4.14×10^{20}cm^{-3}) and ITO (n=6.94×10^{20}cm^{-3}) films. The displayed data sets exhibit similar energy gaps but very different carrier densities. Both the absorption coefficient α and α^2 are plotted and the optical gaps extracted using the methods originally derived by Hamberg et al.[33], illustrated by Dolgonos et al.[147]

The red-shifted IMO plasma edges are understood after plotting optical gap versus plasma energy in Fig. 5.6b. Both the IMO and ITO optical gaps increase monotonically with plasma energy due to band filling; however, the steeper IMO curve indicates a desirably reduced electron effective mass (a similar dependence holds between optical gap and carrier density in Fig. 5.8a). This effect is modelled in Fig. 5.6b using carrier statistics. The plasma frequency is given by

$$\omega_p^2 = \frac{ne^2}{\varepsilon_0\varepsilon_\infty\langle m^*(E)\rangle} \tag{5.1}$$

where $\langle m^*(E) \rangle$ is the density of states averaged effective mass accounting for oscillations of free electrons in the conduction band below the Fermi level E_F,

$$\langle m^*(E) \rangle = \frac{\int_0^\infty g(E)\, m^*(E) f(E, E_F)\, dE}{\int_0^\infty g(E) f(E, E_F)\, dE} \qquad (5.2)$$

where the denominator is equal to the free electron density, n, and $f(E, E_F)$ is the Fermi-Dirac function and $g(E)$ is the non-parabolic density of states derived from the non-parabolic dispersion relation

$$E \left(1 + \frac{E}{E_g} \right) = \frac{\hbar^2 k^2}{2 m^*} \qquad (5.3)$$

The optical gap of a degenerately doped semiconductor is usually determined by the lowest energy direct transitions between the valence band maximum (VBM) and the Fermi level in order to account for the conduction band filling. This is complicated for In_2O_3 by the transitions from the highest lying valence band states being dipole forbidden.[25] So to model the observed optical transitions, it was necessary to included an additional 0.8 eV energy separation between the VBM and the highest valence bands for which transitions are dipole allowed.

The conduction band minimum (CBM) or *band edge* effective mass used is $0.22m_e$ as determined previously from angle resolved photoemission and density functional theory (DFT) calculations.[30,31,38] Modelling of the ITO data required a deviation from the rigid band model by increasing the band edge effective mass incrementally with Sn-content from $0.22m_e$ to $0.40m_e$ (in excellent agreement with previous theoretical predictions in Ref. 38). The carrier density-dependence of the effective mass for IMO and ITO is shown in Fig. 5.8b. The correspondingly different conduction band dispersions in heavily doped IMO and ITO are shown schematically in Fig. 5.6c. The increasing band edge effective mass in ITO is consistent with the additional decrease
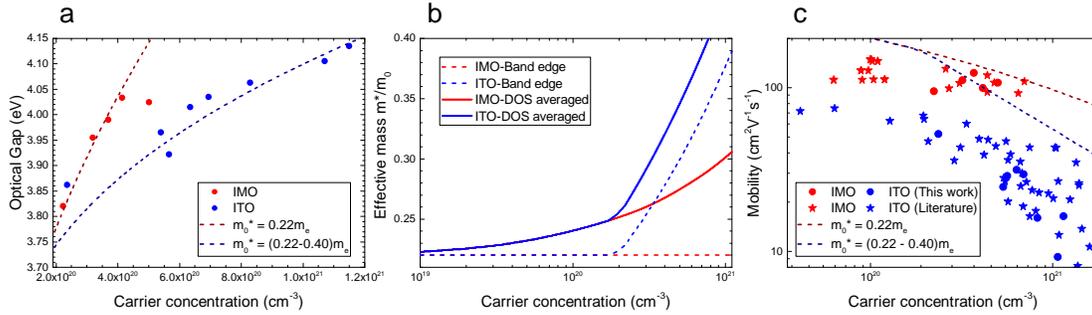
Figure 5.8: a Optical gap as a function of carrier density for IMO and ITO films. The model curves are shown as dashed lines, and are calculated in the same way as that for Fig. 5.6 using equation 5.2. **b** Calculated band edge and density-of-states averaged effective masses for IMO and ITO. The ITO values deviate from those of IMO at 2 $\times 10^{20}$ cm^{-3} as this is the lowest carrier density of ITO studied here. We expect that a smooth transition to occur between IMO and ITO values will occur in the 10^{19} cm^{-3} carrier density range as the Sn concentration increases and the mixing between Sn 5s and CBM states increases. **c** Mobility as a function of carrier density. Measurements presented from this work are displayed as circles whilst data taken from the literature for IMO [88,291–295] and ITO [88,296–301] samples deposited via various methods are stars. Model curves are calculated using the degenerate form of the Brooks-Herring model for ionized impurity scattering, taking into account conduction band non-parabolicity [44,305].

in mobility as carrier concentration increases seen in Fig. 5.3b and modelled in figure 5.8c. In contrast to ITO, the IMO data in Fig. 5.6b is well reproduced using a fixed band edge effective mass of 0.22m$_e$ for all Mo concentrations.

Figure. 5.8b shows the calculated band edge and DOS-averaged effective masses for IMO and ITO using equation 5.2. Under Mo doping there is negligible hybridisation between the CBM and Mo 4d dopant states and this means that the band edge effective mass is constant over the whole carrier concentration range. In contrast, as Sn dopants are incorporated into In$_2$O$_3$, the hybridisation between Sn 5s states and the host conduction band around the CBM increases. For ITO, as doping increases, a sharp change is seen in the band edge effective mass, consistent with calculations by Walsh *et al.*[38] Here, a carrier density for the onset of hybridisation of $n = 2 \times 10^{20}$ cm^{-3} is arbitrarily assigned as we have no Sn doped samples below this carrier density.

Above the Mott criterion ($n \sim 2 \times 10^{18}$ cm^{-3}), the DOS averaged mass increases for both IMO and ITO. Again they diverge at $n = 2 \times 10^{20}$ cm^{-3} from which the band edge effective mass of ITO is rapidly increased in order to model the data. Free carrier

mobilities as a function of carrier concentration are plotted in figure 5.8c. Theoretical scattering limits are calculated using the degenerate Brooks-Herring model, modified by Zawadzki[162,163,305] to account for conduction band non-parabolicity, and simplified by Pisarkiewicz *et al.*[44] (see section 2.3.2). The band edge effective mass variation determined for ITO is applied. From this it is possible to calculate the non-parabolic Fermi level effective mass[43] (as it is the carriers at the Fermi level that are important in scattering events). The theoretical curves reproduce well the divergence of ITO and IMO mobility data at high carrier densities.
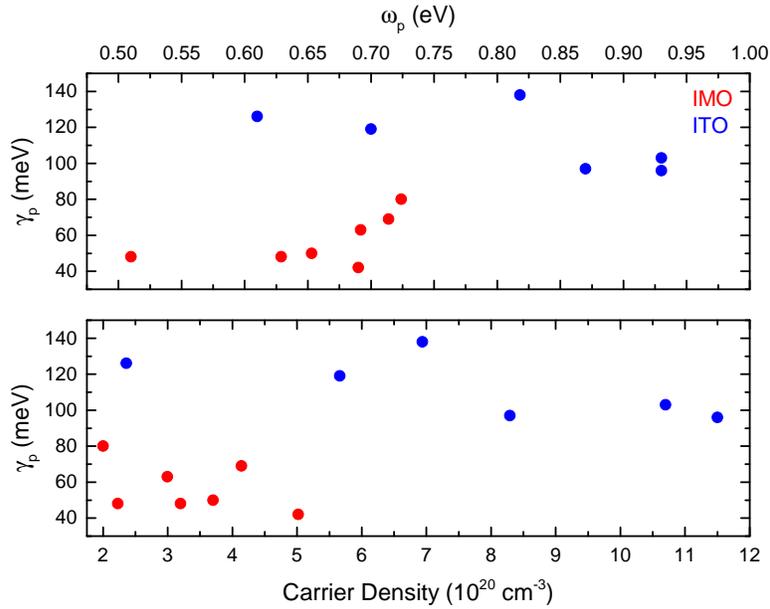
Figure 5.9: Plasmon damping coefficient γ_p as a function of plasma frequency ω_p (top) and carrier density (bottom).

It is worth at this point discussing the damping parameters γ_p extracted from the model fits, which are displayed as a function of plasma frequency and carrier density in figure 5.9. Plasmon damping affects the width of the reflection spectra seen in figure 5.7a, and is related to the lifetime of the plasma electron as $\gamma_p = \frac{1}{\tau}$, and electron mobility as $\mu = \frac{e\tau}{m^*} = \frac{e}{m^* \gamma_p}$. Figure 5.9 indicates that ITO displays higher damping parameters that IMO over the carrier density range available, which in turn
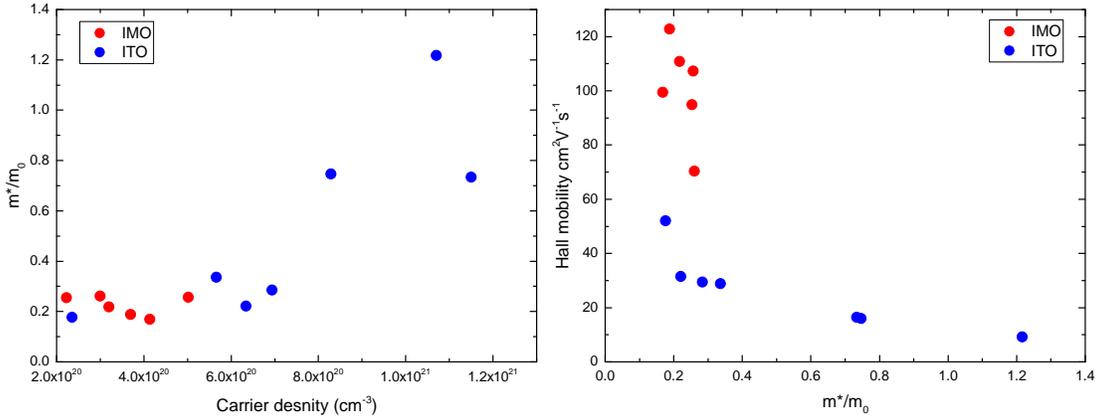
Figure 5.10: Effective mass extracted from the plasmon damping parameters in figure 5.9 and equation 1.15 using the hall mobilities as a function of carrier density, and Hall mobility versus this effective mass.

means a lower plasmon lifetime and so lower electron mobility. This may well relate to the idea of modulation doping discussed in section 1. Because the effective mass and scattering lifetime are related parameters, it is very difficult to disentangle the extent to which both of these parameters impact electron mobility in the films. Using the electron mobilities determined from Hall effect measurements and these damping parameters, a carrier effective mass can be extract from $\mu = \frac{e\tau}{m^*} = \frac{e}{m^* \gamma_p}$. This is plotted in figure 5.10. Clearly, at extremely high carrier densities, the effective mass from ITO samples diverge greatly from $m^* = 0.22 m_0$, and is even far above the effective mass predicted by the carrier dependent model including non-parapolicity, modelled by equation 1.10 shown in figure 1.9. This supports the assignment of strong hybridization of the dopant level in ITO. At lower doping densities the effective masses extracted for ITO are more similar to those for IMO. This would indicate that the difference in mobilities (also shown in figure 5.10) is not entirely due to hybridization, and has some relation to the longer carrier lifetimes seen in IMO. However, due to the complication of the dip in the reflection spectra giving rise to a large uncertainty in the fitting, the rest of the focus of this chapter will be on the impact of the effective mass variation on the mobility. Even so, it must be acknowledged that this variation in

plasmon damping seen in figure 5.9 will also have an impact on the electron mobilities in the films. This should be investigated further with materials with more comparable carrier densities, and using films with much lower surface roughness.

5.3.3 ELECTRONIC STRUCTURE
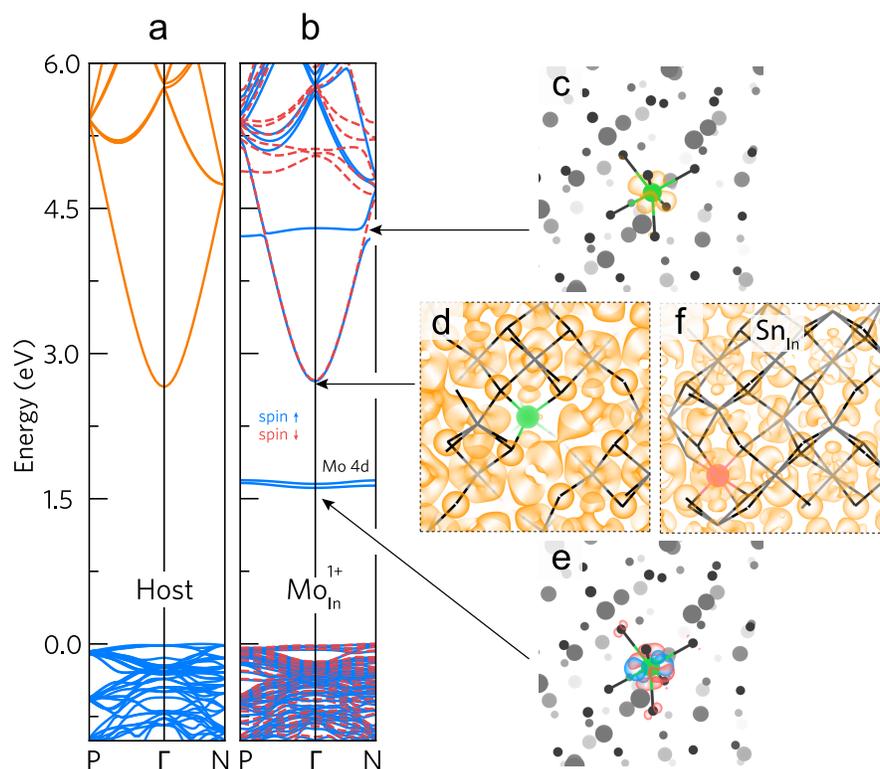
Figure 5.11: Band structures and partial charge densities. **a** and **b** Band structures for undoped In_2O_3 and for IMO, including both spin components, with the Mo donor in the ionized 1+ charge state. The zero of the energy scale is the VBM. **c**, **d** and **e** show the partial charge densities for IMO corresponding to the resonant state in the conduction band, at the CBM and for the two states in the band gap respectively. **f** The partial charge density at the CBM is also shown for ITO. The Mo_{In}^{24d} atom is green, the Sn_{In}^{8b} atom is red and the In_2O_3 lattice is portrayed using a wireframe. For the resonant state and the midgap states, the O and In atoms are shown by black and grey spheres and the 'bonds' on the In_2O_3 lattice are removed in order to clearly visualise the MoO_6 distorted octahedra.

The undoped In_2O_3 and IMO band structures calculated using density functional theory (DFT) are shown in Fig. 5.11a and 5.11b. Mo $4d$ levels in IMO are energetically separated from the CBM and do not mix readily with the In $5s$-O $2p$-dominated CBM. The Mo $4d$ levels can be seen as mid-gap states and as a donor state resonant in the

conduction band in the calculated band structure (Fig. 5.11b). Mo states mix minimally with those of the In_2O_3 CBM; in contrast Sn states strongly perturb the CBM. This is supported by figures 5.11c–f, which depict the DFT-calculated partial charge densities. At the CBM, whilst the IMO electron density is delocalised across the lattice, minimal density exists near the Mo (green) atom (figure 5.11d); in contrast, considerable density exists across the ITO lattice, including around the Sn (red) atom (Fig. 5.11f). The Sn 5s states hybridize with the In 5s/O 2p character CBM.
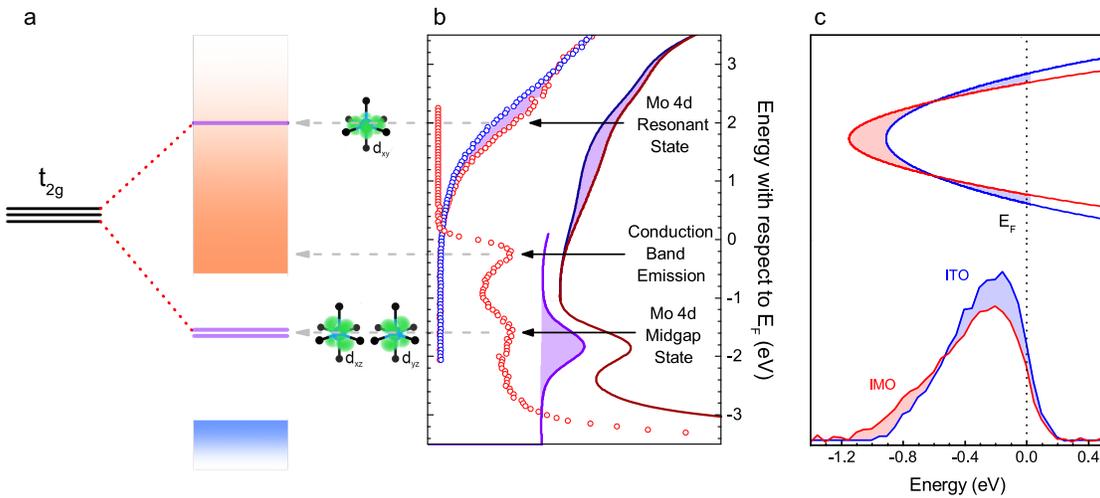
Figure 5.12: Spectroscopy of electronic structure. a Schematic diagram of Mo $4d\,t_{2g}$ states and their corresponding orbitals (depicted in green), splitting into the upper state, d_{xy}, resonant with the conduction band (orange shading), the two lower states, d_{xz} and d_{yz}, in the band gap and the valence band (blue shading). The splitting resulting from tetrahedral elongation Jahn-Teller distortion of the MoO_6 octahedron is shown for comparison with the highly distorted Mo_{In}^{24d} depicted in Fig. 5.11. **b** Inverse photoemission data from ITO (blue) and IMO (red), showing extra intensity in the IMO data, corresponding to an unoccupied Mo $4d$ state resonant with the conduction band. Also shown is high energy photoemission data of the conduction band emission, Mo $4d$ midgap states and valence band onset of IMO. The features in part **a** of the figure are aligned to the corresponding peaks in the spectra. **c** Schematic band diagram, showing wavevector (vertical axis) versus energy (horizontal axis) representing the conduction bands of similarly doped IMO (red) and ITO (blue) aligned to the Fermi level. Below are background subtracted HAXPES spectra displaying conduction band emission from IMO (red, $n = 5.0 \times 10^{20}$ cm^{-3}) and ITO (blue, $n = 6.3 \times 10^{20}$ cm^{-3}). The spectra are normalised to the In 5s-dominated valence band peak at -9 eV. The shaded regions emphasise the additional intensity of the ITO peak and additional width of the IMO peak.

The Mo 4d states are depicted schematically in Fig. 5.12a and are observed in the photoemission and inverse photoemission spectra shown in Fig. 5.12b. Conduction band photoemission shown in Fig. 5.12c directly evidences the different conduction

band dispersion for IMO and ITO discussed above (see also figures 5.13). While the midgap states might be thought to impair transparency, this is shown not to be the case from the transmission data in Fig. 5.6. This is because optical transitions to and from the midgap Mo $4d$ states have only minimal dipole intensity.
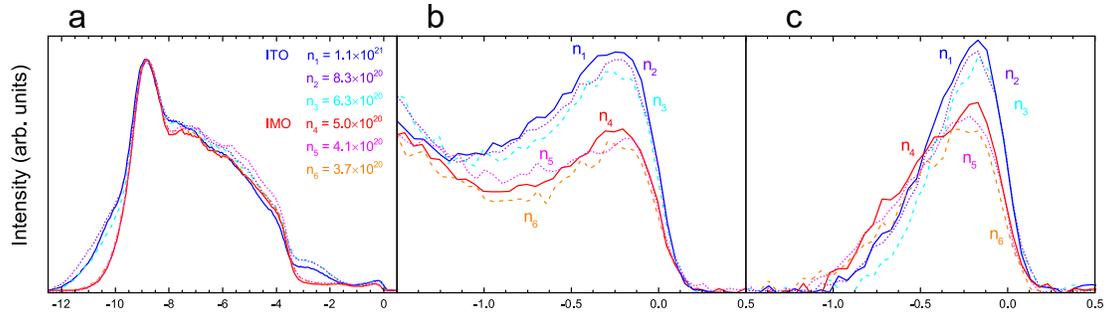
Figure 5.13: a HAXPES VB spectra for several ITO and IMO films. **b** Conduction band spectra for IMO and ITO. **c** Background subtracted conduction band spectra. Spectra are normalised to the In $5s$ feature at -9 eV.

Figure 5.13 shows hard x-ray photoemission spectra for a range of different dopant concentrations in ITO and IMO films. These spectra can be used to further investigate the dependence of the conduction band density-of-states on the carrier concentration. The spectra are normalised to the In 5s-dominated peak at a binding energy of -9 eV. Immediately it can be seen that samples with higher carrier concentration have a greater intensity (figure 5.13b) just below the Fermi level (zero of the binding energy scale). The spectra from ITO films have more intense peaks than those from IMO, even for films of similar carrier density. By subtracting the background from the conduction band states (due to the Mo 4d states in IMO and Sn-Sn disorder-related states in ITO), it is possible to directly compare the energy widths of the density of states, showing that IMO has lower band edge effective mass, giving lower density of states per energy interval and so greater Burstein-Moss shift than for similar carrier density ITO.

High energy x-ray photoemission spectra of the Sn $3d$ and Mo $3d$ core levels are displayed in figure 5.14. Both plots display data taken using photon energies of 6450
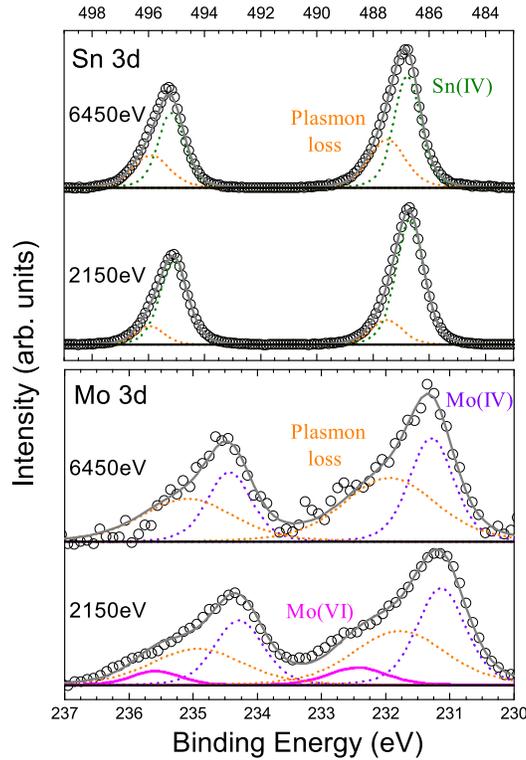
Figure 5.14: HAXPES spectra of the Sn $3d$ and Mo $3d$ core levels in ITO ($n = 6.3 \times 10^{20}$ cm^{-3}) and IMO ($n = 5.0 \times 10^{20}$ cm^{-3}) respectively recorded using photon energies of 6450 eV (top) and 2150 eV (bottom).

eV and 2150 eV, enabling discrimination between bulk and surface features. Spectra have had a Shirley background subtracted and are fitted with Voigt functions. All doublet peak areas have been constrained to a 3:2 ratio consistent with d-orbitals, and doublet separations have been constrained (8.4 eV for Sn 3d and 3.15 eV for Mo 3d). The full-width at half maximum of $3d_{5/2}$ and $3d_{3/2}$ peaks from the same chemical species or plasmon loss feature have been constrained to be the same. The binding energy scale is referenced to the Fermi level.

Sn $3d_{5/2}$ and $3d_{3/2}$ core-level peaks were fitted with two components each. The component at 486.7 eV is due to Sn in the Sn(IV) oxidation state, consistent with other measurements[219]. There is no lower binding energy peak and so there is no Sn present in the Sn(II) oxidation state. Hence, the HAXPES data are consistent with Sn acting

172

as a one electron donor in ITO. The component at 487.5 eV is a plasmon loss feature which arises from energy loss of photoelectrons to the free electron gas in conductive materials.[137,306] The positions of the plasmon loss features are constrained to the plasma frequency, ω_p, determined for each film by infrared reflectivity. The plasmon loss component has marginally larger relative intensity when 6450 eV photons are used. It is a general observation that the lower energy component is larger relative to the plasmon peak taken with HAXPES compared to XPS[137]. However, the difference is usually only negligible at such high carrier concentrations and any difference seen could easily be within the uncertainty of the fitting. Hence, no claims are made here regarding the relative intensities of the plasmon loss and no loss peaks under illumination from different photon energies.

Mo $3d_{5/2}$ and $3d_{3/2}$ core-level spectra are also displayed in figure 5.14. The lower 2150 eV photon energy data was fitted with three peaks. The first is at 231.2 eV, corresponding to Mo in the Mo(IV) oxidation state. The second is a plasmon loss peak associated with the Mo(IV) component, lying at higher binding energy by the plasma frequency, $\omega_p \approx 0.6$ eV, measured by infrared reflectivity. The third component is at 232.5 eV. The substitutional Mo has a binding energy very close to that of MoO_2 of 231.0 eV,[307] confirming the Mo(IV) oxidation state of the Mo dopant in IMO. The highest energy component is attributed to Mo in the Mo(VI) oxidation state of MoO_3 at the surface.[307] This peak is much smaller than the Mo(IV) and is negligibly small in the high photon energy data, confirming the surface oxide assignment. This interpretation is consistent with that presented previously for Mo $3d$ data from IMO, but with the improved signal-to-noise ratio and resolution here enabling the identification of the plasmon loss feature associated with the Mo(IV) component. Photoelectrons from the surface MoO_3 exit the sample without passing through the IMO and so do not lose energy to the free carrier plasma and so have no associated plasmon loss feature.
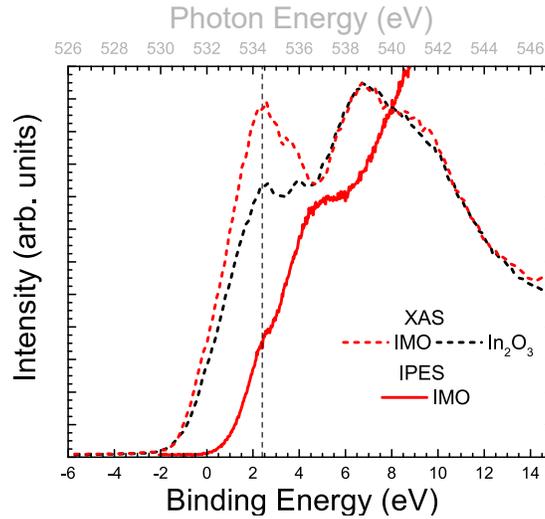
Figure 5.15: O K-edge XAS of IMO and In_2O_3 and IPES spectra of IMO on a common energy scale. The resonant Mo $4d$ state is seen in both spectra at $\sim$ 2.3 eV above the Fermi level.

X-ray absorption spectroscopy (XAS) can be used to probe the unoccupied states of a material above the Fermi level. This technique is employed here in this study to provide support for the assignment of the Mo $4d$ states in our IPES spectra. O K-edge absorption spectra of IMO and nominally undoped In_2O_3 were recorded in partial electron yield mode. The photon energy was calibrated using a TiO_2 reference, which was then aligned with the Fermi level using the O $1s$ peak from our HAXPES measurements with the inclusion of a 1 eV core hole correction.[308] This is plotted in figure 5.15 along with the IPES spectra for IMO and undoped In_2O_3 films (from figure 5.12) which is energy referenced to the known energy position of the lowest occupied molecular orbital of multilayer C_{60} deposited in situ, giving the Fermi energy. The oxygen partial density of states in the XAS plot has a much higher intensity at around $\sim$2.3 eV, in excellent agreement with the Mo $4d$ feature seen in the IPES. This is indicative of some hybridization of the Mo $4d$ states with the oxygen states. The alignment of these features after independent charge referencing supports the identification of the unoccupied Mo $4d$ state as the donor in IMO films (note that differences in intensity, resolution and indeed the states that are being measured are among the things

that may mean spectral alignment and interpretation may be more complicated than first thought. Indeed XAS probes a convolution of initial and final states making the understanding of the spectra a rather complicated task even in simple systems). While supportive, the complex nature of the bonding in Mo in In_2O_3 means it is not obvious how the XAS spectra should differ between the two samples. Indeed, XAS and IPES being excited state measurements means simple ground state DFT cannot compute the electronic density of states very accurately and so some of the more nuanced features such as the modification of the O 2p DOS are hard to predict. We do not show this here therefore. However, a clear difference is evidenced in figure 5.15, and so whilst correlation does not imply causation, it is satisfying to see this trend in the two spectra.

5.3.4 Sn_{In} and Mo_{In}

Sn_{In} acts as a low formation energy *resonant* donor under both *In-rich/O-poor* and *In-poor/O-rich* conditions with formation energies for the neutral charge state of 0.42 eV and 0.99 eV respectively (see figure 5.16). Sn preferentially sits on the 8b Wyckoff site in correlation with previous theory results.[137,309] The 1+/0 transition level occurs around 0.85 eV above the CBM which allows for very high carrier concentrations as seen in experiment, and through Burstein-Moss shifts of ∼0.85 eV.[137] Under *In-poor/O-rich* conditions, Sn_{In} is compensated for by V_{In}^{3-} around 0.55 eV above the CBM trapping the Fermi level just below this point.

Due to Mo possessing three additional valence electrons to In, two spin states were modelled: a *high-spin* (denoted 'HS' and involves three electrons in the spin-up (α) component) and a *low-spin* (denoted 'LS' and involves two electrons in the spin-up and one in the spin-down (β) component). The two inequivalent In sites (corresponding to 24d and 8b Wyckoff positions) were also assessed in this study for each spin
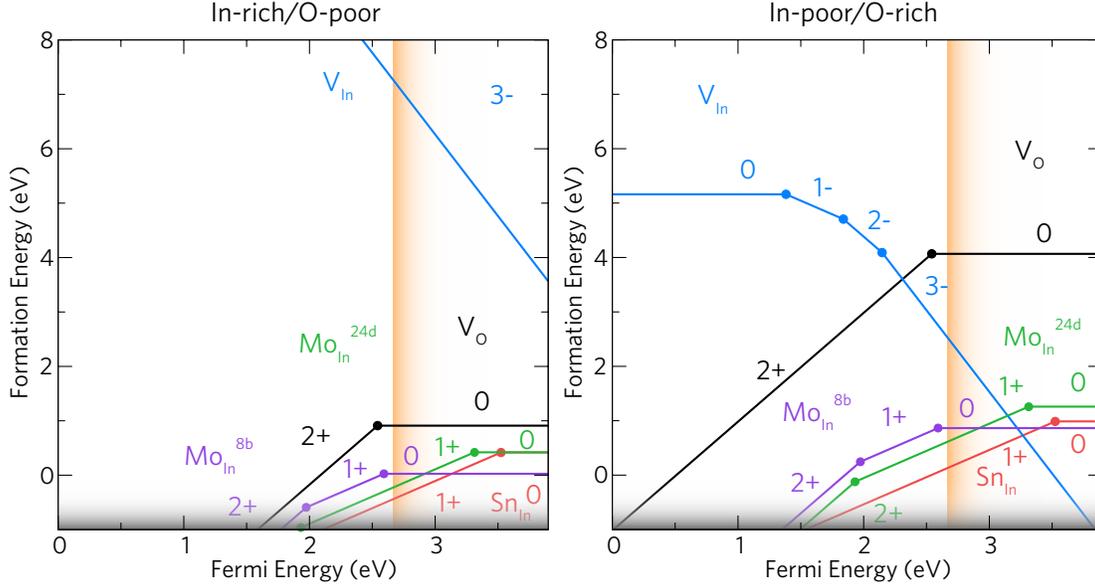
Figure 5.16: The thermodynamic transition levels for In_2O_3 under *In-rich/O-poor* and *In-poor/O-rich* conditions. The Fermi level ranges from the VBM (0 eV) to the CBM (2.63 eV). The substitutional Mo (Mo_{In}) defects are split into the 24d site (green) and the 8b site (purple) in the high-spin configuration (HS) which was found to be the lowest energy configuration for both sites.

configuration.

From the HSE06 relaxations it was shown that the high-spin configurations for the neutral charge state possessed the lowest formation energies for both the 8b and 24d sites by ~ 0.87 eV and ~ 0.31 eV respectively. These results are in keeping with a previous study by Xu et al. who predict this as the ground state configuration of substitutional Mo in the In_2O_3 lattice. Mo_{In}^{8b} acts as a *deep* defect with a 1+/0 transition level occurring around 0.08 eV below the CBM.[310] The neutral charge state possesses a formation energy of 0.02 eV and 0.86 eV under *In-rich/O-poor* and *In-poor/O-rich* conditions respectively, making Mo on this site a highly soluble dopant. Two further transition levels occur within the band gap: a 2+/1+ and 3+/2+ transitions around 0.70 eV and 1.59 eV below the CBM respectively. Mo_{In}^{24d} on the other hand acts as a *resonant* one-electron donor with the 1+/0 transition level occurring around 0.64 eV above the CBM. Despite the neutral charge state of Mo_{In}^{24d} possessing a formation energy of

0.42 eV and 1.25 eV under *In-rich/O-poor* and *In-poor/O-rich* conditions respectively, this charge state is only reached at Fermi levels far above the CBM. From figure 5.16, the ionised charge state (1+) possesses a lower formation energy at the CBM and thus it can be rationalised that Mo_{In}^{24d} is the *dominant* dopant species to be present in the In_2O_3 lattice. As with the 8b site, the 2+/1+ and 3+/2+ transition levels occur deep within the band gap around 0.74 eV and 1.35 eV below the CBM.

5.3.5 MOLECULAR ORBITAL DESCRIPTION OF IMO

To explain why Mo in In_2O_3 has two occupied $4d$ states within the band gap and one donor state resonant with the conduction band, consider the electronic structure and crystal field environment of Mo in In_2O_3. Mo and Sn have respective electron configurations of $[Kr]4d^5 5s^1$ and $[Kr]4d^{10}5s^2 5p^2$, while bixbyite cubic In_2O_3 has two distinct indium sites: the 8b and 24d Wyckoff sites, with respective *regular* and *tetragonally-distorted* octahedral crystal environments. At an indium site, three electrons participate in bonding, leaving Sn or Mo respectively in a $[Kr]4d^{10}5s^1$ or $[Kr]4d^3$ configuration, that is one Sn $5s$ or three Mo $4d$ as the outermost electrons. At Fermi levels matching experiments (around VBM+3.3 eV) Sn or Mo preferentially substitute on the respective 8b or 24d sites (each donating one conduction electron), whilst Mo_{In}^{8b} has a neutral charge state (donating no carriers), see Fig. 5.16. The Jahn-Teller-like distortion at the 24d site splits the three-fold degenerate t_{2g} levels (valid for octahedral symmetry), shifting the d_{xy} orbital up in energy to the vicinity of the CBM, and d_{xz} and d_{yz} down into the gap, as shown in the unfolded band structures for the neutral charge state (Mo_{In}^{24d}, q=0) in Fig. 5.17. If all Mo $4d$ states lie below the CBM (as reported elsewhere from calculations),[310] then no effective doping would occur at room temperature; the experimental results shown in Fig. 5.3 clearly rule this out. If the highest occupied Mo $4d$ state lies just above the CBM (reported elsewhere),[88] it may merge

with the occupied CBM (although this will be dictated by the orbital wave function overlap), increasing the electron effective mass and reducing mobility, and the Fermi level would be pinned by the partially occupied $4d$ state.

Both previous works fail to consider Mo atoms in a +1 charge state,[88,310] which increases the distortion of the octahedra and the splitting of the t_{2g}-derived $4d$ energy levels, making the d_{xy} orbital resonant with the conduction band. This contributes a conduction electron without detrimental hybridization with the CBM states; whilst the singly-occupied d_{xz} and d_{yz} orbitals are stabilized by the bond length contraction and pushed deeper into the band gap (see Mo_{In}^{24d} q=+1 in Fig. 5.17). A schematic representation of this t_{2g} level splitting is shown in Fig. 5.12a along with partial electron densities for the d_{xy}, d_{xz}, and d_{yz} orbitals (full schematic molecular orbital diagrams for MoO_6 octahedra are shown in Fig. 5.18). Indeed, the calculated partial electron charge densities around the Mo atom at the resonant Mo level (Fig. 5.11c) and two mid-gap levels (Fig. 5.11e) show clear similarities respectively to the schematic d_{xy}, and a combination of the d_{xz} and d_{yz} orbitals of Fig. 5.12a. This picture of the Mo $4d$ state energies in IMO is further corroborated by additional confidence in the calculated results coming from comparisons of the bond lengths from calculations and those determined from previous extended x-ray absorption fine structure measurements (see Fig. 5.19 and related discussion).

The unfolded bandstructures relating to the undoped host, Mo_{In} on the 24d and 8b sites (HS) and Sn_{In} were calculated and are shown in figure5.17. Both the neutral (q=0) and ionised (q=1+) charge state for each dopant was calculated to compare the electronic structure arising from the structural changes upon relaxation. In the high-spin configurations, the dopant states occur in the spin-up (α) component (in the energy range -1−6 eV) and as such the spin-down (β) component is not shown in this comparison. From the HSE06 calculations, the density of states at the CBM is found to
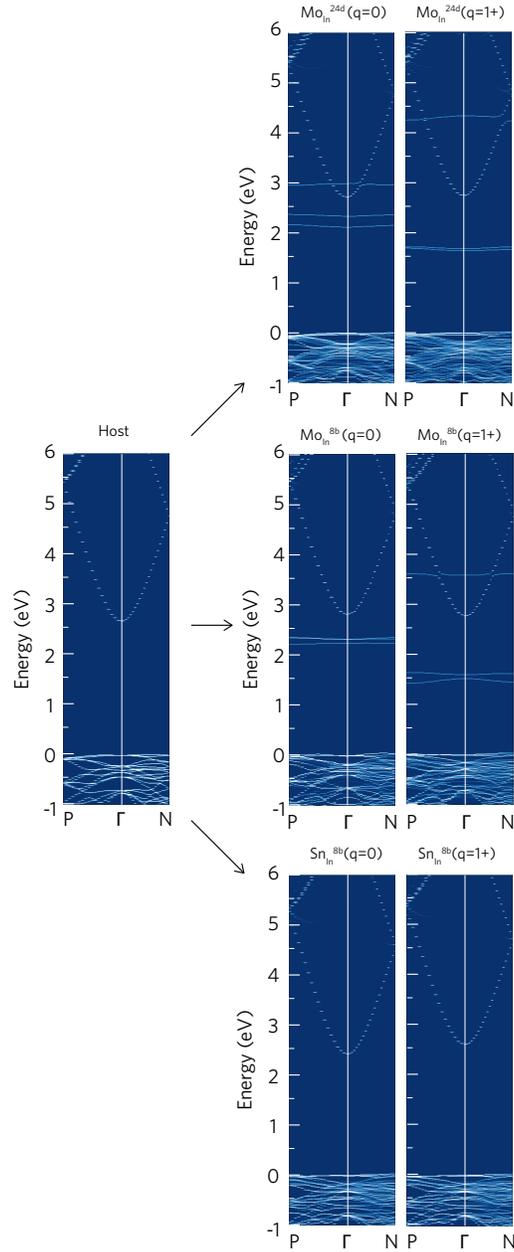
Figure 5.17: The unfolded band structure comparison for the undoped host, Mo_{In}^{24d}, Mo_{In}^{8b}, Sn_{In}^{8b} defects. Each dopant state is represented in the neutral (q=0) and ionised (q=1+) charge state and is represented by the spin-up (α) component. The spectral weight of each state is presented by the brightness of the curve.

be around 3% Sn *s* character. This hybridisation has the effect of raising the electron effective mass by reducing the dispersion. This effect is seen in previous band structure calculations containing corrections to enable a more reliable description of the

effective mass changes for In_2O_3 with a change in band edge effective mass upon Sn doping from $0.22\,m_e$ to $0.40\,m_e$ (Ref. 38). The Sn $5s$ atomic orbital level lies directly between the O $2p$ and In $5s$ levels, resulting in strong hybridization with the In $5s$-O $2p$-dominated CBM. The partial charge density of the CBM in figure 5.11f displays the electron density delocalised over the In s, O p and Sn s states. This hybridisation explains the perturbation of the CBM dispersion upon Sn-doping which leads to the increased band edge effective mass.

In Mo-doped In_2O_3, the two different In Wyckoff sites give rise to different electronic structures. In the neutral charge state, Mo_{In}^{24d} displays negligible Mo hybridisation at the CBM and instead two *filled* bands of Mo $4d$ character exist in the band gap around $2.11\,eV$ and $2.33\,eV$ above the VBM respectively. A further empty band occurs around $2.98\,eV$ above the VBM ($\sim 0.31\,eV$ above the CBM) also of $\sim 73\%$ Mo $4d$ character.

In its ionised form (q=1+), Mo(IV) oxidation state, Mo_{In}^{24d} displays an enhanced splitting of the Mo bands with the two filled bands lowering in energy in the band gap to around $1.66\,eV$ and $1.67\,eV$ above the VBM. These are depicted in a partial charge density plot in figure 5.11. The Mo $4d$ band above the CBM is raised to $4.25\,eV$ above the VBM ($1.58\,eV$ above the CBM) and is depicted in figure 5.11b. These results correspond well to the inverse photoemission spectra, with the unoccupied Mo $4d$ being observed only about 0.5 eV higher than calculated by DFT. This also highlights the short comings of the findings of Xu *et al.*[310] who do not take into account the ionised dopant state for any of their transition metal-doped TCO calculations. Without considering the ionised dopant state, the calculation of the dopant d energy levels is erroneous and so it is impossible to determine reliably whether a particular dopant will perform well. The fact that their neutral charge state calculations give a d level close to the CBM explains why they suggest Mo will not produce good electrical properties in In_2O_3. This is despite previous experimental evidence to the contrary which is now rationalised

by our theoretical results giving the donor d level far above the CBM for the 1+ Mo charge state.
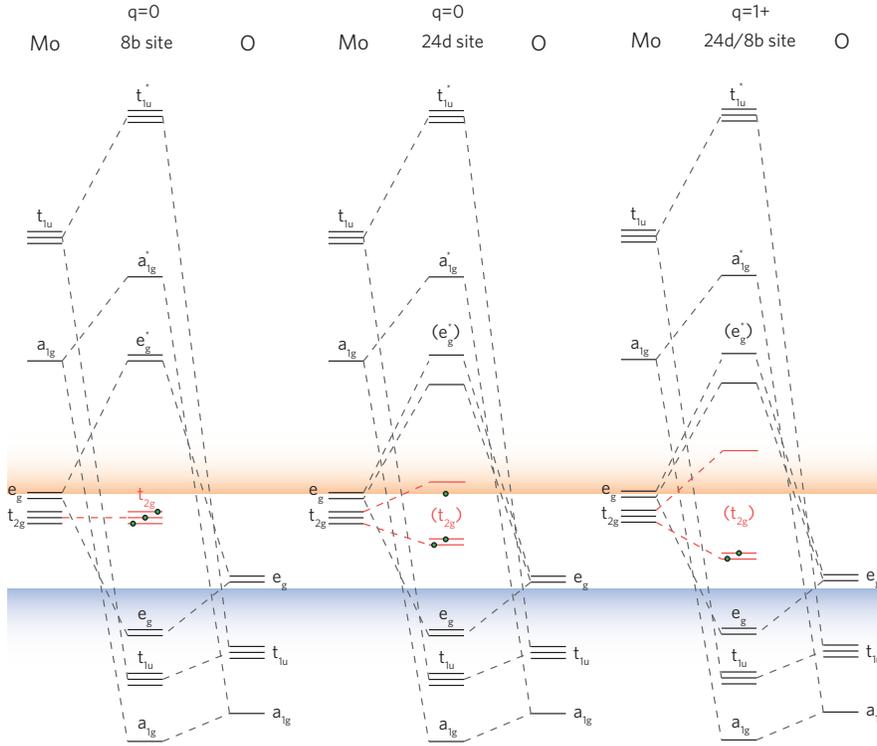
Figure 5.18: The schematic molecular orbital diagrams for the MoO$_6$ octahedra on the 8b site, the 24d site in both the q=0 and 1+ charge states. This diagram depicts the Jahn Teller-like distortion that occurs within the In$_2$O$_3$ lattice, giving rise to the splitting of the degenerate non-bonding t$_{2g}$ state on the Mo. The valence band and conduction bands are represented by the blue and orange regions respectively. Only the electrons corresponding to the t$_{2g}$ site are shown, depicted by the small green circles.

As there is negligible Mo d state contribution at the CBM, Mo does not perturb the CBM dispersion, allowing IMO to retain the host In$_2$O$_3$ band edge effective mass, resulting in the high electron mobilities seen in this work and previous literature.[88] Figure 5.11d depicts the CBM for the Mo (24d)-doped In$_2$O$_3$ cell showing the absence of any electron density contribution from Mo. The cause of the Mo $4d$ splitting can be attributed to the Jahn-Teller-like distortion Mo$_{In}^{24d}$ in the neutral charge state which, in the 1+ charge state increases due to the further contraction of the bond lengths further splitting the levels. This can be rationalised using crystal-field splitting seen

in the simplified molecular orbital diagram for MoO_6 octahedra in figure. 5.18.

When doped on the 8b site, Mo_{In} possesses three *filled* Mo $4d$ bands in the band gap at 2.23 eV, 2.31 eV and 2.31 eV respectively. As the 8b In site is only slightly distorted from the perfect octahedral (O_h) symmetry, a slight splitting of the degenerate non-bonding t_{2g} levels is expected (see figure 5.18). As the Mo–O bond lengths are uniform with little distortion around the Mo_{In} in the neutral charge state (see figure. 5.18), three occupied non-bonding Mo $4d$ bands are seen in the band gap. Mo can therefore be seen to exist as Mo(III) in this configuration. As with the 24d site, Mo_{In}^{8b} undergoes a further splitting of the t_{2g} bands in its ionised form due to a similar Jahn-Teller distortion. Here, the two $4d$ states in the band gap appear at 1.51 eV and 1.59 eV above the VBM respectively whilst the third unoccupied band exists around 3.57 eV above the VBM (~ 0.89 eV above the CBM).
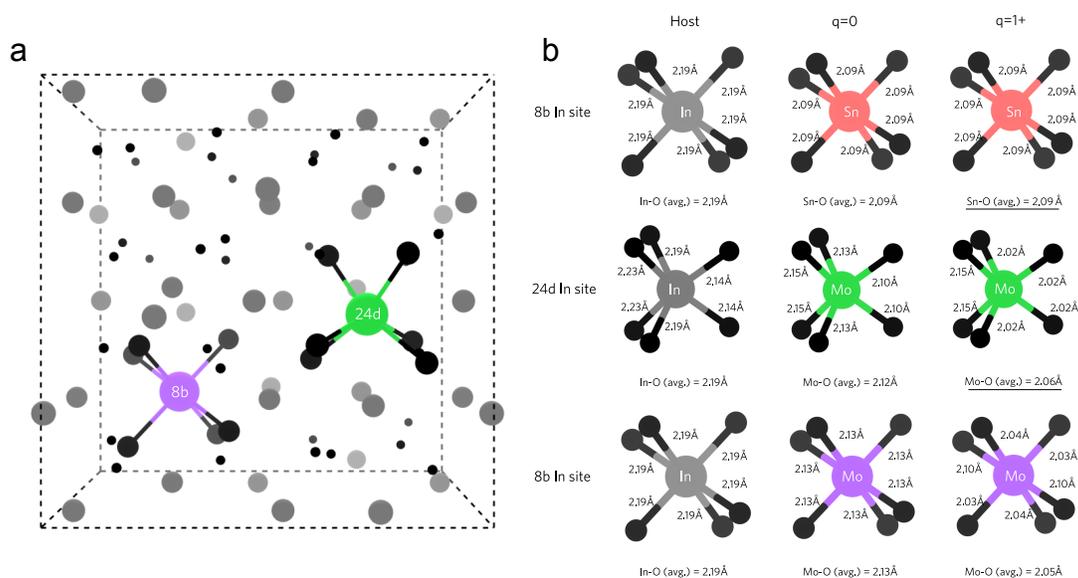
Figure 5.19: a The two different Wyckoff indium sites within the In_2O_3 supercell. The 8b and 24d sites are depicted in purple and green respectively, whilst the oxygens and indium atoms are shown by the black and grey spheres. The 'bonds' are only shown on the highlighted sites for clarity. **b** The M–O octahedra (where M=In, Mo, Sn) for the different In-Wyckoff positions (24d and 8b) and different charge states (q=0 and 1+) showing the different calculated bond lengths. Each dopant site is colour coordinated to the thermodynamic transition levels in Extend Data Fig. 5.16 whilst the black spheres correspond to oxygen. The expected configuration based on the defect thermodynamics (for Mo and Sn) is shown underlined.

Figure 5.19 displays the calculated bond lengths present in the In_2O_3 lattice as a set of MO_6 (where M=In, Mo or Sn) octahedra relating to the 8b and 24d In sites. For undoped In_2O_3 the average In–O bond lengths (for both 8b and 24d sites) are 2.19 Å which correspond well to EXAFS studies on In_2O_3 which quote values in the range of 2.16–2.18 Å [88,311,312] which are within 1.5% of the HSE06 calculated values. The Sn–O bond lengths contract uniformly to around 2.09 Å in both the neutral and 1+ charge state (ionised). These values are in keeping with experimental EXAFS values of $\sim$ 2.05–2.08 Å for ITO. [311]

In Mo-doped In_2O_3 the 24d site shows a contraction of the bond lengths from the neutral (q=0) to the ionised (q=1+) charge states (figure 5.19b with an average bond length of 2.12 Å and 2.06 Å respectively. For the 24d site, Mo will be ionised within the In_2O_3 lattice, therefore the relaxed q=1+ structure is likely to be seen and matches well with EXAFS measurements of 2.05 Å for Mo–O. [88] Further analysis of the Mo–In bond lengths gives rise to two (average) values of 3.38 Å and 3.84 Å which are in excellent agreement with the EXAFS values of 3.36 Å and 3.82 Å. [88] When incorporated on the 8b In site, the Mo–O bonds contract uniformly compared to the In–O bond lengths, giving rise to an average bond length of 2.13 Å. In the 1+ charge state, Mo_{In}^{8b} undergoes a further contraction of the bond lengths, however *non-uniformly*, with a Jahn-Teller distortion occurring (equatorial Mo–O:2.03–2.04 Å, axial Mo–O:2.10 Å). Although the average bond lengths correspond to 2.05 Å which matches the EXAFS literature values, as Mo_{In}^{8b} is a thermodynamically *deep* defect. It is expected to exist in the neutral charge state at Fermi levels at and above the CBM.

5.3.6 DURABILITY OF IMO

This research shows that IMO has fundamentally better optoelectronic properties than ITO. As it is also a cost effective replacement for ITO, from these considerations, it

should take its place in the transparent conductors market. However, any replacement for ITO must also be at least as durable - it must withstand the wear and tear and environmental conditions it is likely to be exposed to in typical uses. ITO is predominantly used as an electrode in modern electronic devices, such as touch screens and televisions. Hence, these films are nearly always encapsulated in a device which offers the film a lot of protection. The conditions ITO is exposed to our arguably less harsh than those experienced by FTO in low emissivity window coatings and in solar cells. Nevertheless, here simple lab-based tests are performed in order to simulate some different environmental stresses on the IMO and ITO films, and on a commercially available product (Ossila ITO film on glass substrate, $d = 100nm$, $R_s = 20\Omega/\square$) for comparison.
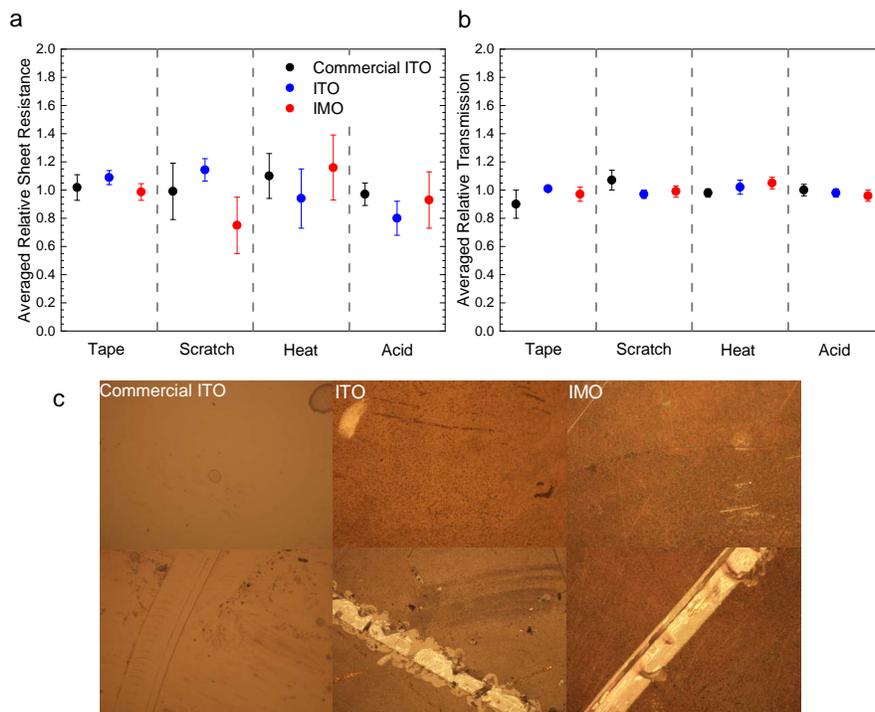
Figure 5.20: Durability testing: a) Relative change in sheet resistance between pre and post-test samples. b) Relative change in optical transmission between pre- and post-test samples. c) Optical micrograph of pre-test (top) and post-scratched (bottom) samples.

All durability tests were repeated several times to enable the uncertainty in the

change of sheet resistance and transmission to be given as the standard deviation in the results. Sheet resistance of films was determined using a CMT-SR2000 4-point probe system and transmission was measured using a Shimadzu UV-Vis-IR 3700 spectrophotometer. Sheet resistance values were averaged over a number of measurements across each sample. Transmission values were taken at wavelengths of $\lambda =$500nm, 700nm, and 1250nm to give a reasonable representation of the whole spectrum. Finally, the relative values of both the sheet resistance and transmission were found by dividing the measured value post-test by that of a pre-test sample (i.e. a relative sheet resistance >1 indicates the test increased the sheet resistance).

Optical Microscopy was performed using a Nikon Eclipse LV100 microscope in bright field setting at $\times$20 optical zoom. Optical measurements were performed after each test, although little difference was noticeable except subsequent to scratch testing when the obvious imprint of the scratch could be seen.

The scotch tape test was performed using Intertape LA-26 Polyester Laminate Tape. Adhesive tape was pressed onto films and left to set. Subsequent removal of the tape allows a measure of film de-lamination. None of the films showed any signs of de-lamination as indicated in figure 5.20a and b.

The scratch test was performed using an Elcometer 3092 sclerometer which has a tungsten carbide tip and a maximum 20N force spring which was utilised for all measurements. Fig. 5.20a shows some changes to the sheet resistance following the scratch test, with a slight increase for ITO and small decrease for IMO. However, the statistical uncertainty on these measurements was greater than after the tape test, indicating that the differences are not significant. The optical transmission post scratch test was unchanged, as shown in figure 5.20b. The optical micrographs shown in Figure 5.20c clearly demonstrate the difference is morphology of the three films (top row) and the effect that the tungsten carbide tip has on the films. The commercial ITO is

thinner than our films and has a lower degree of surface roughness which is apparent in the micrograph, appearing less opaque and less granular (less light scattering from the surface). The scratch applied to the commercial ITO does not seem to penetrate deep, likely due to the film being much thinner. Our films seem to be penetrated more deeply with a noticeable difference in depth of field. However, this again does not appear to have greatly affected the electronic or transmissive properties and so is of little concern.

Films were heated in a furnace in ambient at 250℃ for three hours to simulate the effects of heat expelled from components in an electronic device. Figure 5.20a and b again demonstrate all films performed well in the tests in terms of electronic and optical properties.

Finally films were immersed in dilute sulfuric acid for three hours. The acid had a pH of around 3.5, slightly more acidic than acid rain or sweat which are the most likely acids to come into contact with the films in applications. No appreciable sheet resistance or optical losses were noticeable post treatment, as can be seen in Figure 5.20a and b.

These tests demonstrate that AACVD IMO can withstand general wear and tear and environmental conditions as well as both a commercially produced ITO film and ITO deposited using the same technique. This indicates that, subject to more extensive industrial durability testing, there are no barriers to the widespread adoption of IMO as a superior replacement for ITO.

5.4 Conclusion

Beyond having a transformative impact by replacing the ubiquitous ITO with better and cheaper IMO, this work enables design criteria for other high mobility TCOs to be established: (1) transition metal dopants are required so that their donor d states

hybridize minimally with the host cation *s* states of the CBM, avoiding problems inherent with conventional *s* level dopants; (2) the transition metal dopant atoms must possess *d* orbital energy levels such that the dopant is a singly charged donor to minimize mobility-reducing ionized impurity scattering; and (3) the donor level must be resonant within the conduction band to avoid a high effective mass, low mobility *d* state at or close to the CBM.

IMO is an optically and electronically superior material to the commercially-dominant transparent electrode, ITO. This makes IMO not only more suited for the many diverse applications that ITO is currently used for, but also implies enormous potential for industrial cost savings as much thinner films of IMO can be produced with properties equal to or better than ITO of the same carrier concentration. A combination of theory and experiment have been used to explore the reasons for the drastic differences between these two In_2O_3-based TCOs. The rigid band model fails to describe the variation of optoelectronic properties of ITO with doping concentration. For heavy Sn-doping, the role of the Sn $5s$ states in perturbing the host In_2O_3 conduction band and increasing the band edge effective mass must be included to describe both the evolution of optical gap and mobility with doping. Having identified this shortcoming of ITO, it has been shown that this is avoided by instead using Mo as the dopant. The different energetic position and atomic orbital character of Mo states compared with those of Sn means that negligible hybridization between Mo states and the CBM occurs. These findings provide a new immediately viable and durable alternative to ITO, the industry standard TCO. While Mo and Sn are similarly priced, IMO films with half the thickness offer huge cost savings over ITO films of the same sheet resistance. Moreover, the understanding presented will enable the design of other TCO materials and dopants with improved transparent conducting properties.

6

Surface properties of novel oxide semiconductor: Ga_2O_3

Bulk single crystals of ($\bar{2}01$) β-Ga_2O_3 were purchased from Novel Crystals Technology, Inc., Tamura Corporation. Density functional theory calculations were carried out and analysed by J. B. Varley at Lawrence Livermore National Laboratory. Band bending calculations were performed by T. Veal at the University of Liverpool. The results presented here were published as:

J. E. N. Swallow, J. B. Varley, L. A. H. Jones, J. T. Gibbon, L. F. J. Piper, V. R. Dhanak, and T. D. Veal, Transition from electron accumulation to depletion at β-Ga_2O_3 surfaces: The role of hydrogen and the charge neutrality level. APL Materials 7, 022528 (2019).

6.1 Introduction

Beta phase gallium oxide (β-Ga$_2$O$_3$) is a transparent, oxide semiconductor material that has attracted a large amount of interest in recent years due in part to its wide band gap ($\sim$ 4.7 eV at room temperature),[313,314] gas adsorption-tunable conductivity, and high breakdown field (8 MVcm^{-1}). These properties mean that β-Ga$_2$O$_3$ is a promising material for a large range of applications, such as solar-blind ultraviolet photodetectors[315], gas sensing devices,[316] and high power electronic devices.[317] Despite the suitability of this material for such a large range of applications, there is a distinct lack of spectroscopic information on its surface electronic behaviour. Unlike other oxide semiconductors, such as In$_2$O$_3$ [Ref. 59], SnO$_2$ [Ref. 318], CdO [Ref. 319] and ZnO [Ref. 320], the surface electronic behaviour has not been well understood in the case of Ga$_2$O$_3$, which is an important factor for gas sensing devices and electronic contacts.

Semiconductors with the greatest mismatch between their cation and anion in terms of atomic size and electronegativity, such as CdO [Ref. 59,319] and In$_2$O$_3$ [Refs 59, 321] and InN [Ref. 322,323], appear to only exhibit surface electron accumulation, as opposed to the electron depletion present at the surface of the majority of n-type semiconductors.[71] The slightly less mismatched materials such as ZnO [Refs 324–327] and SnO$_2$ [Refs 328,329] can have either of these two types of surface space charge, depending on how the surface is treated. Indeed, it has been proposed recently that electron depletion is the inherent space charge property of the ZnO surface.[326,327]

As the Ga cation is smaller and has higher electronegativity than Zn, such a trend suggests that Ga$_2$O$_3$ may behave, in terms of surface space charge properties, more like ZnO than In$_2$O$_3$. Upward surface band bending and electron depletion is assumed in Ga$_2$O$_3$ due to the difficulty in achieving Ohmic contacts with low conductivity mate-

rial,[330] consistent with similar observations for ZnO. In this context, recent x-ray photoemission spectroscopy (XPS)[331−333] and angle resolved photoemission spectroscopy (ARPES)[334−336] has been interpreted as indicating surface electron depletion and upward band bending. However, previous analysis of XPS valence band spectra employed the method of linear extrapolation of the leading valence band edge which underestimates the surface valence band maximum (VBM) to Fermi level separation. This is because, just as in the case of In_2O_3,[59] the top of the valence band has very little dispersion and a high effective mass, leading to a very rapid onset of the density of states. Instrumental broadening in XPS introduces a significant slope to the measured onset of the valence-band photoemission, leading to the aforementioned underestimation. In the case of the previous ARPES data where the resolution is higher, the broadening is lower and so a higher surface VBM to Fermi level separation has been reported of about 4.9–5.0 eV. However, the presence of electron depletion was still incorrectly inferred due to a band gap of about 4.9 eV being assumed, higher than the measured room temperature value, some 0.2 eV or so lower.[334−336]

Further, the space charge layer of the ZnO surface has been shown to be highly dependent on surface properties,[324,325] especially the effects of H adsorption.[337] Whilst H doping in the bulk has been considered in recent years for Ga_2O_3 both experimentally[338] and theoretically,[339,340] its influence on the surface has mainly been considered from a gas sensing perspective in terms of changes to the measured conductivity,[341,342] rather than in photoemission studies. Here we report the surface electronic behaviour and band bending of β-Ga_2O_3, and contrast our interpretation of XPS data with that contained in other recent reports.

6.2   Experimental and Theoretical Methods

Bulk single crystalline ($\bar{2}$01) β-Ga$_2$O$_3$ (from Novel Crystal Technology, Inc., Tamura Corporation) grown using the edge-defined film-fed growth method[166] was used in this study. Sn-doped samples were used to make them sufficiently conducting to avoid sample charging effects during photoemission measurements. The net donor density ($N_D^+ - N_A^-$) was determined by the manufacturer via capacitance-voltage measurements to be 6×10^{18} cm^{-3} and the free electron density was confirmed to be n$\sim$ 6×10^{18} cm^{-3} using Hall effect measurements. Optical transmittance was performed using a Shimadzu UV-Vis-IR 3700 spectrophotometer which employs a photomultiplier detector to reach energies up to 6.5 eV, enabling the absorption onset of β-Ga$_2$O$_3$ to be determined. XPS was performed using a SPECS monochromatic Al Kα ($h\nu$ =1486.7 eV) x-ray source operated at 250 W. A PSP Vacuum Technology hemispherical electron-energy analyzer with a 120 mm mean radius was employed to detect photoelectrons, operated at a pass energy of 10 eV (50 eV for survey scans). The XPS system is described in detail elsewhere.[343] Binding energies are stated throughout with respect to the Fermi level. Binding energies are calibrated using the Fermi edge of an Ar$^+$-ion bombarded polycrystalline silver sample, also enabling the resolution of 0.4 eV to be determined. This resolution includes broadening due to the x-ray source, the electron analyser and thermal effects. The uncertainty in stated binding energies is $\pm$0.05 eV. All measurements were performed at room temperature. Prior to in-situ cleaning, the Ga$_2$O$_3$ was investigated in its as-received form. In-situ surface cleaning was performed via sample annealing at between 200 and 800°C for 30 min cycles[332,344,345] and a final 2 hour cycle at 800°C. All annealing steps were performed consecutively on the same crystal.

The DFT uses the HSE06 screened hybrid functional[346,347] and projector-augmented wave (PAW) approach[348] as implemented in the VASP code.[224,225] The semi-core Ga 3d

states are explicitly included as valence electrons and adopted a fraction of 32% exact-exchange in the hybrid functional. These choices lead to an excellent description of the lattice and electronic structure, with a 0 K direct gap of 4.87 eV (indirect of 4.86 eV).[349] Temperature effects lead to a decrease of $\sim$ 0.2 eV between 0 K and room temperature according to recent measurements.[350] The treatment of the Ga d states leads to slightly improved lattice parameters as compared to experiment (12.21, 3.03, and 5.79 Å for the a, b, and c lattice parameters, and a β angle of 103.8 degrees), and most importantly for this study, an improved description of orbital interactions that can influence the valence band features observable via photoemission spectroscopy. While the primitive unit cell was used for the bulk electronic structure calculations, defect calculations of H_i defects were performed in 120-atom supercells for the Ga_2O_3 and 216-atom supercells for SnO_2 and corrected for finite-size effects following the approach of Freysoldt *et al.* as described elsewhere.[339,351,352]

6.3   Results and Discussion

6.3.1   X-ray Photoemission Spectroscopy

In order to determine the band bending at the surface of β-Ga_2O_3 and its evolution with annealing, XPS has been used to measure the valence band and core level spectra at each annealing stage. The valence band spectra give the surface VBM to Fermi level separation. The bulk CBM to Fermi level separation can be determined using semiconductor statistics from the bulk carrier density and, with knowledge of the band gap, this gives the bulk VBM to Fermi level separation. The difference between the bulk and surface VBM energies with respect to the Fermi level thus provides the sign and magnitude of the surface band bending. Changes in the band bending as a function of annealing result in changes to the VBM binding energy with respect to

the Fermi level and can be confirmed by observing the corresponding shifts of the core level peaks. The core levels provide additional information about contamination via both peaks due to adsorbed species and chemical shifts of the elements of the semiconductor.

The valence band spectra for β-Ga$_2$O$_3$ are shown in Fig. 6.1 for the sample (a) before surface cleaning, (b) after an 800°C anneal for 2 hours, and (c) the evolution of the valence band over the whole cleaning procedure. The calculated hybrid DFT valence band density of states (VBDOS) are compared to the XPS valence band spectra. Photoionization cross-section corrections taken from Ref. 123 have been applied to the calculated VBDOS with the result shown in Fig. 6.1 with and without instrumental (0.4 eV full width at half maximum (FWHM) Gaussian) and lifetime broadening (0.45 eV FWHM Lorentzian). The calculated and thus corrected VBDOS is in extremely good agreement with the experimental data.

The methodology of extracting the valence band position is demonstrated in Fig. 6.1(a) and (b) which involves shifting the broadened calculated VBDOS until it aligns with the spectral features of the XPS data. The leading edge of the unbroadened VBDOS is then taken as the VB-edge position (relative to the Fermi level aligned to zero). This method of determining the VBM energy with respect to the Fermi level is well established[353] but not widely used and the much simpler method of linear extrapolating the leading edge is generally preferred and has been widely used for the analysis of Ga$_2$O$_3$ XPS data.[332,354,355] However, as discussed above, for materials with very flat valence bands, with a correspondingly sharp onset of the VBDOS, the linear extrapolation method underestimates the VBM to Fermi level separation due to the instrumental broadening. Comparing the green dotted linear extrapolations in Fig. 6.1(a) and (b) to the red lines of the aligned unbroadened DOS onset, enables us to determine that linear extrapolation underestimates the VBM position by ~ 0.5 eV. All the calculated
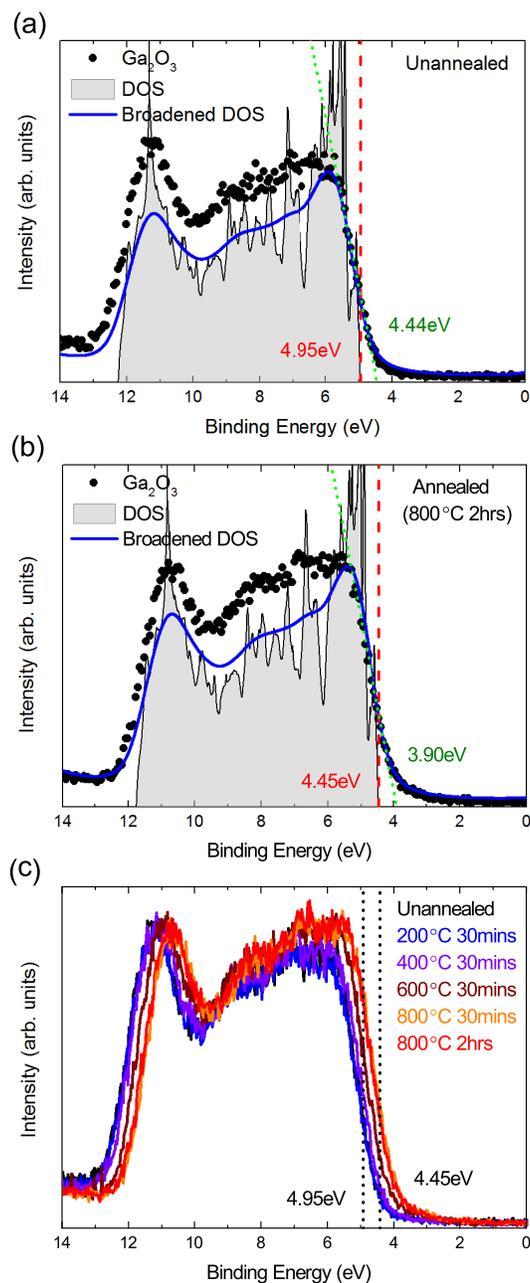
Figure 6.1: Valence band photoemission spectra and DFT valence band total DOS with and without instrumental and lifetime broadening applied for β-Ga_2O_3 (a) as entered, and (b) after annealing at 800°C for 2 hours. (c) the evolution of valence band spectra under subsequent annealing cycles at increasing temperatures.

VBDOS features match very well with the Ga_2O_3 data, giving good confidence that the VB position is well represented. For materials, such as ZnO, where the valence

band onset is more gradual, the two methods for determining the VBM position are in much closer agreement with each other.

Focusing now on Fig. 6.1(c), it is apparent that after each annealing step the VB edge shifts to lower binding energy. From the valence band spectra, the surface VBM to Fermi level separation is found to vary from 4.95±0.10 eV after no surface preparation to 4.45±0.10 eV after surface cleaning using the highest temperature annealing (800°C).

6.3.2 SURFACE BAND BENDING
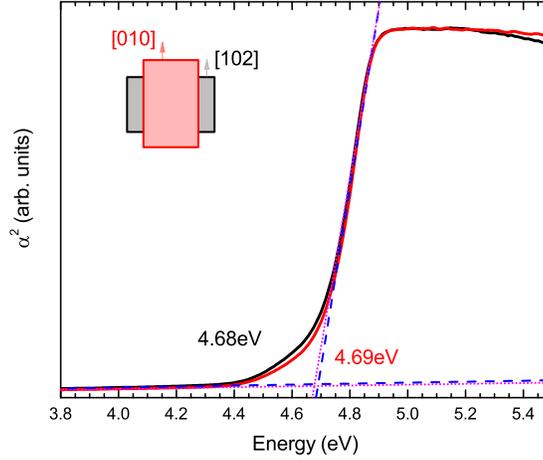
Figure 6.2: Optical absorption as a function of photon energy for β-Ga$_2$O$_3$ with the [010] plane upward (red) and alternately the [102] plane facing the incoming light (black).

As mentioned above, to use photoemission to determine the sign and amount of band bending at the polar β-Ga$_2$O$_3$ ($\bar{2}$01), it is important to have a measure of the band gap of the material investigated. Fig. 6.2 shows two absorption spectra of β-Ga$_2$O$_3$ derived from transmission data, with the second spectrum taken with the crystal oriented perpendicular to its own long axis in the first measurement (schematically shown in inset). A linear extrapolation indicates the onset of optical absorption is at 4.68 eV. This is consistent with other optical measurements[350], energy loss spectroscopy[355] and the calculation results mentioned above. This allows a conservative

estimate of the band gap of ~4.7±0.2 eV when excitonic effects and optical anisotropy are considered.[313,356] It also is worth noting the relative independence of band gap on carrier concentration up to $\sim 10^{19}$ cm^{-3} displayed in Ref. 350. The Mott criterion given by $n_c^{1/3} a_B \approx 0.25$ where a_B is the effective Bohr radius (given by $a_0 \frac{\epsilon(0)}{m^*/m_0}$) and n_c the critical carrier density[357] gives n_c around 3×10^{18} cm^{-3}, meaning our absorption onset only minimally deviates from the fundamental band gap (from band filling effects). Indeed, our semiconductor statistics calculations, assuming a band edge effective mass of 0.28m$_0$ [Refs 358–360] and a nonparabolic conduction, band[191] indicate that for Ga$_2$O$_3$ with a carrier density of 6×10^{18} cm^{-3} the Fermi level is about 30 meV (0.03 eV) above the conduction band minimum (CBM). That is, the bulk CBM is essentially at the Fermi level. Therefore, given the measured band gap of about 4.68 eV, this indicates β-Ga$_2$O$_3$ ($\bar{2}$01) goes from downward to upward surface band bending as a function of in situ annealing. That is, the as received β-Ga$_2$O$_3$ ($\bar{2}$01) surface exhibits an electron accumulation layer which transforms into an electron depletion layer as the annealing treatment progresses. In order to illustrate the band bending and charge density as a function of depth quantitatively, the Poisson equation has been solved within the modified Thomas-Fermi approximation (MTFA),[361,362] with a nonparabolic conduction band, as described in chapter 3.[190,191]

Fig. 6.3(a) and (b) shows the calculated band bending before and after annealing for the β-Ga$_2$O$_3$ ($\bar{2}$01) surface, and the associated electron density plots as a function of depth respectively. For the unannealed β-Ga$_2$O$_3$, the 0.24 eV downward band bending is associated with a corresponding accumulation of electrons in the near-surface region with a sheet electron density of 5×10^{12} cm^{-2}. For In$_2$O$_3$, weak conduction band emission was observed below the Fermi level in photoemission spectra.[30,59] However, in that case the surface sheet density was an order of magnitude higher than observed here for Ga$_2$O$_3$, explaining the absence of this feature in Fig. 6.1. This is also supported
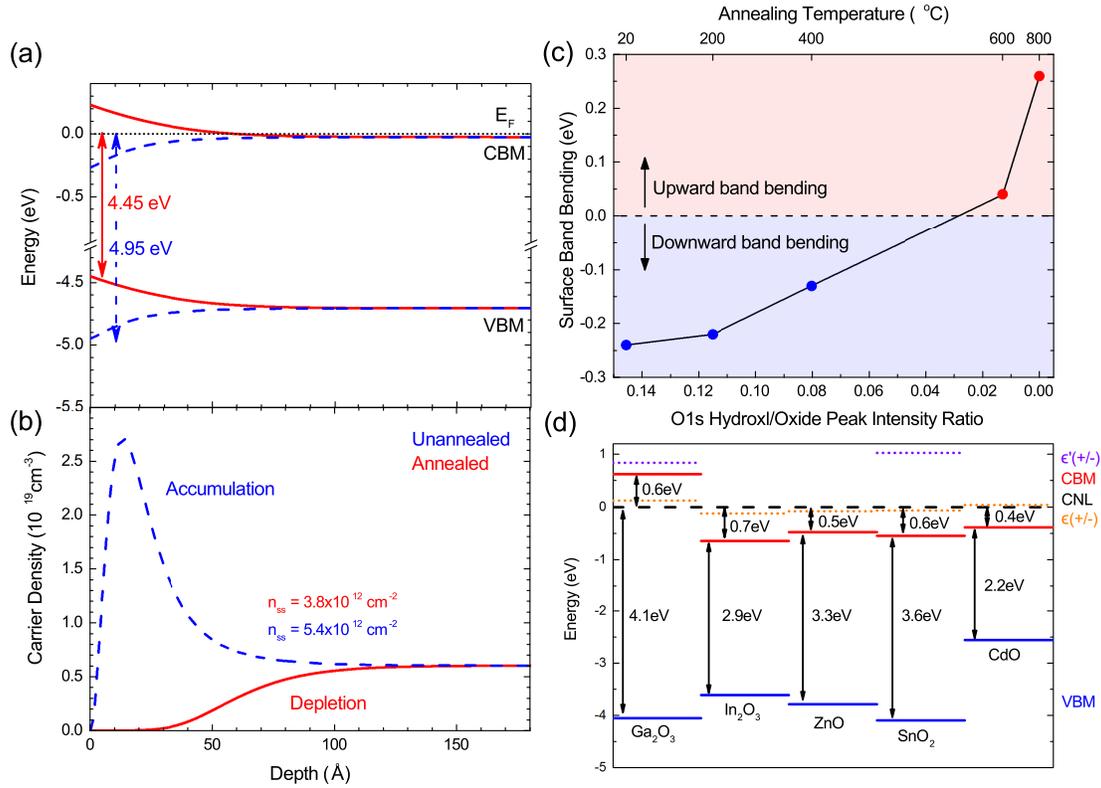
Figure 6.3: Poisson-MTFA calculations of (a) the band bending and (b) carrier density profiles in the electron accumulation (blue) and depletion (red) surface layers for β-Ga$_2$O$_3$ before and after cleaning. (c) The β-Ga$_2$O$_3$ surface band bending and the relative contribution of the hydroxyl component of the O 1s spectrum as a function of annealing temperature. The annealing temperature of 20 on the upper x-axis represents the unannealed sample. The lines between points are a guide to the eye. (d) Band line-up with respect to the charge neutrality level for β-Ga$_2$O$_3$ compared to that of other metal oxide semiconductors. The CNL for β-Ga$_2$O$_3$, In$_2$O$_3$,[59] ZnO[83] and CdO[59] are determined from valence band offsets from particle irradiation studies of other materials[363] and from calculations for SnO$_2$.[364] The CBM shown for CdO is for the Γ point rather than the indirect CBM at the the L point. The ε(+/-) transition levels for H$_i$ configurations which involve disrupting the cation-anion bond are shown as dotted orange lines for β-Ga$_2$O$_3$ (0.51 eV below the CBM - this work), In$_2$O$_3$ (0.53 eV above the CBM),[365] ZnO (0.40 eV above the CBM),[70] SnO$_2$ (0.48 eV above the CBM for the bond disrupting H$_i^+$ configuration - this work), and CdO (0.43 eV above the CBM).[366] Also shown, as purple dotted lines, are the ε'(+/-) transition levels for H$_i$ configurations that involve O lone-pairs capturing H$_i$ which are the lowest formation energy for β-Ga$_2$O$_3$ (ε'(+/-) at 4.90 eV above the VBM,[367] corresponding to 0.22 eV above the CBM for our room temperature band gap of 4.68 eV) and SnO$_2$ (ε'(+/-) at 1.53 eV above the CBM).[339]

in figure 6.4, where even at high energies there is no clear conduction band emission. The other charge profile shown in Fig. 6.3(a) is for the 0.26 eV upward band bending and surface electron depletion present after annealing at 800°C for 2 hours. The corresponding acceptor surface state density is 3.8×10^{12} cm^{-2}. The band bending as
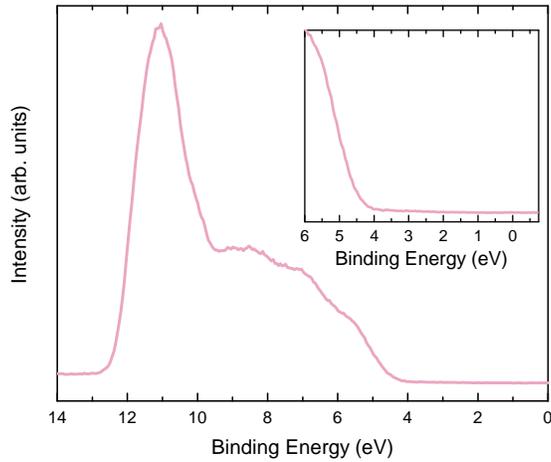
Figure 6.4: Valence band spectrum taking using hard (5921 eV) energy photons. Inset: shows a zoomed in picture of the valence band edge, displaying clearly no conduction band emission.

a function of annealing treatment is shown in Fig. 6.3(c) and is also correlated with the relative strength of the hydroxyl contribution to the O 1*s* spectrum, as discussed below. The uncertainty in the absolute band bending values is ±0.22 eV due to the aforementioned uncertainties in the band gap and in determining the VBM position. However, the change in band bending is independent of the band gap and, between the as-received surface and after 800 annealing, is 0.50±0.14 eV using the VBM positions with their ±0.10 eV uncertainty or 0.50±0.07 eV using the core-level positions with their ±0.05 eV uncertainty. The uncertainty in the band bending value would be much smaller than ±0.22 eV with a less conservative estimate of the band gap than the ±0.2 eV uncertainty stated here. Fig. 6.3(d) compares the natural band alignment and bulk interstitial H transition levels of Ga_2O_3 with other metal oxide semiconductors and will be discussed further below.

6.3.3 THE ROLE OF HYDROGEN IN Ga_2O_3

In order to confirm the changes in the surface electronic properties as a function of annealing and to investigate their origin, core level XPS spectra were also recorded. Fig. 6.5(a)-(d) shows XPS survey and core level spectra for β-Ga_2O_3 after successive in-
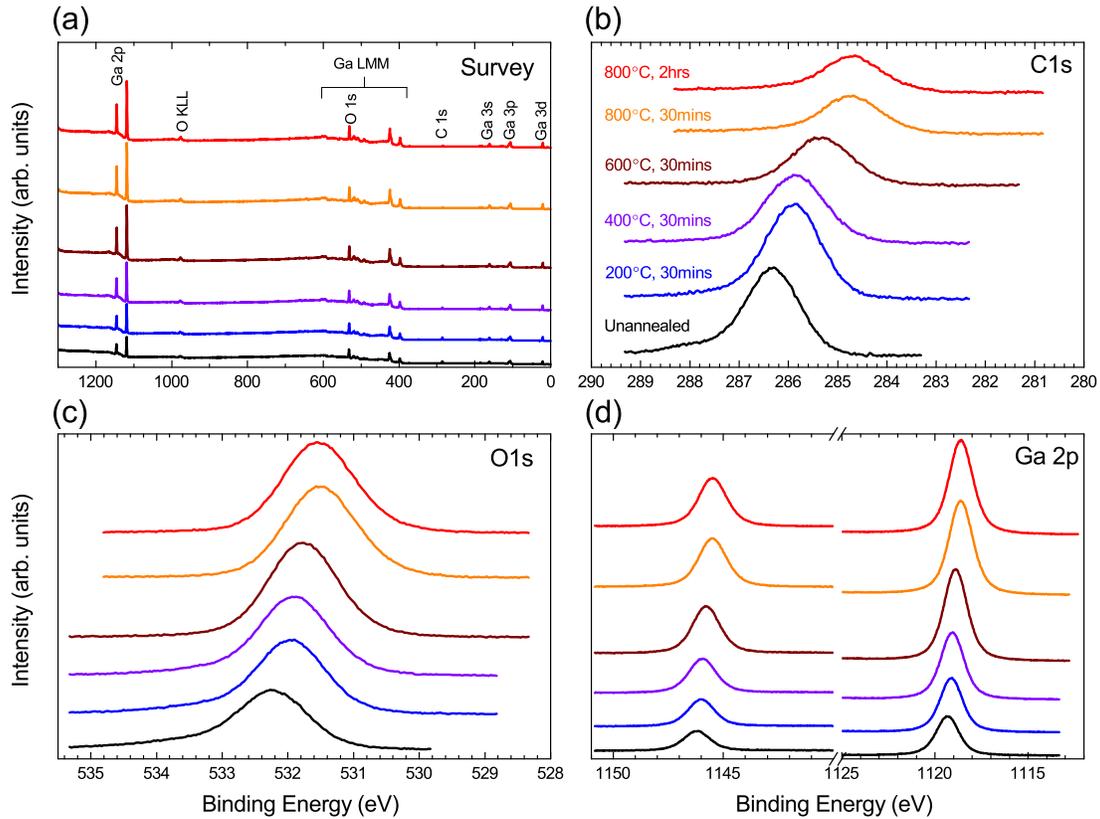
Figure 6.5: β-Ga$_2$O$_3$ XPS spectra following thermal annealing steps at 200, 400, 600 and 800 for 30 minutes and finally for 2 hrs at 800 showing (a) wide survey scan with visible peaks labeled. (b) C 1s core-level, (c) O 1s core-level and (d) Ga 2p core level doublets, all demonstrating a binding energy shift to lower energy upon surface cleaning up to 800°C. The binding energy scale is referenced to the Fermi level.

situ thermal annealing steps were performed to remove surface contamination. Starting from an unannealed sample (black spectra), 200°C temperature steps were performed for 30 mins up to 800°C, which was also done for 2 hours (red spectra). All survey spectra possess only gallium, oxygen and carbon lines. No peaks related to Sn could be discerned due to the low concentration of Sn in the samples ($\sim 6 \times 10^{18}$ cm^{-3}), well below the detection limit. As successive annealing steps are completed, Ga and O related peaks are seen to increase in intensity, whilst C related peaks are greatly reduced, indicating the removal of contamination from the surface. This is more clearly seen in the core level spectra in Fig. 6.5(b) C 1s, (c) O 1s and (d) Ga 2p.

The C 1s peak reduced in intensity after each annealing step up to 800°C where it could not be reduced further. Annealing at a higher temperature (1000°C) caused another carbon species to develop on the surface, possibly as a result of carbon reacting with Ga_2O_3. All core level peaks shifted to lower binding energy after each cleaning step, indicating changes in the surface electronic properties of the material occurring as a result of the surface cleaning.
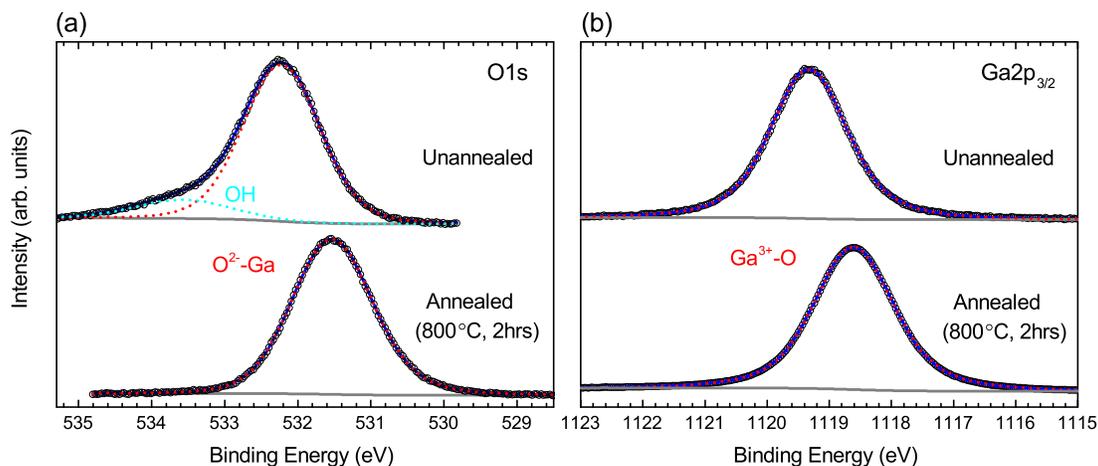
Figure 6.6: (a) O 1s and (b) Ga 2$p_{3/2}$ XPS core level peaks and curve fitting for both unannealed (as entered) samples and following a 2 hour 800°C annealing cycle.

This energy shift is investigated in more detail by peak fitting of the O 1s and Ga 2p core levels as shown for the unannealed and maximally annealed crystal in Fig. 6.6. The O 1s spectra prior to cleaning (top) and after a 2 hour 800°C cleaning cycle (bottom) can be seen in Fig. 6.6(a). Both spectra were fitted using a Shirley background and Voigt line shapes. The spectrum possessed a strong component with a high binding energy shoulder. The main component is associated with O bonded to Ga at 532.2 eV, with a shoulder component 1.3 eV higher in binding energy, which is commonly associated with a hydroxyl (-OH) group.[368–372] In this case, this is likely due to H adatoms bonding to O atoms on the Ga_2O_3 surface. Subsequent annealing reduces and then removes the shoulder component completely and a binding energy shift of $\sim$0.7 eV

to lower energy (531.6 eV) is observed. The FWHM of the O-Ga O $1s$ component remained the same (1.2 eV) before and after annealing, suggesting no new chemical species is formed. The binding energy shift of the O-Ga peak is associated with the change in the surface electronic properties resulting from the hydroxyl groups being removed from the surface. This correlation between the relative intensity of the hydroxyl component of the O $1s$ peak and the band bending is apparent in Fig. 6.3(c).

This suggests that the downward band bending and electron accumulation is due to the hydrogen on the surface. That is, the adsorbed hydrogen on the surface acts as an extrinsic source of positively charged donor surface states which provide the electrons in the accumulation layer. As the hydrogen is desorbed by annealing, the surface donor density decreases along with the downward band bending, and all the core levels shift accordingly. This effect has been seen previously in ZnO.[326,327] Indeed this is seen in Fig. 6.6(b) where the Ga $2p_{3/2}$ peak shifts after cleaning by the same amount also (from 1119.3 eV to 1118.6 eV with no change in the FWHM of 1.5 eV).

As the annealing steps progress, the initial decrease in the amount of downward band bending is followed by the development and then increase of upward band bending and surface electron depletion. The greatest upward band bending of 0.26 eV occurs when the hydroxyl component of the O $1s$ spectrum is absent after annealing at 800°C. This type of behaviour has also been reported for polar ZnO surfaces.[325-327] Despite the relative area between the hydroxyl and oxide components in figure 6.6a being known, it is difficult to determine the fraction of surface sites that act as donors from this as there may be many surface states that do not contribute to the space charge. The surface states will have an associated distribution of energies. Only the unoccupied donor surface states will be charged and contribute to the space charge by donating electrons. Any below the Fermi level will be occupied, will not contribute electrons, and be neutral. The proportion of OH-related surface states which con-

tribute to conductivity thus depends on the position of the Fermi level within the (unknown) distribution of surface states. This may explain why the relative intensity can be quite high, while the sheet density of carriers is quite low (which only include states that contribute to conductivity). Additionally, the OH species is expected to reside directly at the surface, and so will not be as attenuated in the photoemission process as subsequent layers which means this peak is made relatively more intense. These results suggest that electron depletion is the 'natural' surface space charge state of Ga_2O_3 surfaces. This corresponds to the presence of occupied negatively-charged acceptor-type intrinsic surface states.

Having electron depletion at the surface of moderately doped *n*-type Ga_2O_3 with roughly coincident bulk CBM and Fermi level, is consistent with the so-called charge neutrality level (CNL) lying below the CBM. The CNL, otherwise referred to as the Fermi level stabilization energy (of the amphoteric defect model[373,374]) or branch point energy,[375] is the energy at which the surface states change from being predominantly donor-like below the CNL, to being predominantly acceptor-like above the CNL.[59,323] Indeed, as shown in Fig. 6.3(d), the natural band alignments of the metal oxides indicate that CNL lies about 0.6 eV below the CBM for Ga_2O_3. Such a position high in the band gap and relatively close to the CBM for such a wide band gap material is consistent with other reported properties of Ga_2O_3. These include its *n*-type dopability,[376] the ability to form Schottky barriers at metal/Ga_2O_3 interfaces,[376] the sensitivity of surface conductivity to gas adsorption,[341] and the decrease in conductivity upon particle irradiation of already *n*-type material.[377]

However, a possible contradiction is presented by the bulk donor behaviour of interstitial hydrogen, H_i, in *n*-type Ga_2O_3. If the Fermi level is above the CNL, it would be expected that H_i would act as compensating acceptors. This is because for H_i configurations that involve disrupting the cation-anion bond, the H_i $\varepsilon(+/-)$ transition level,
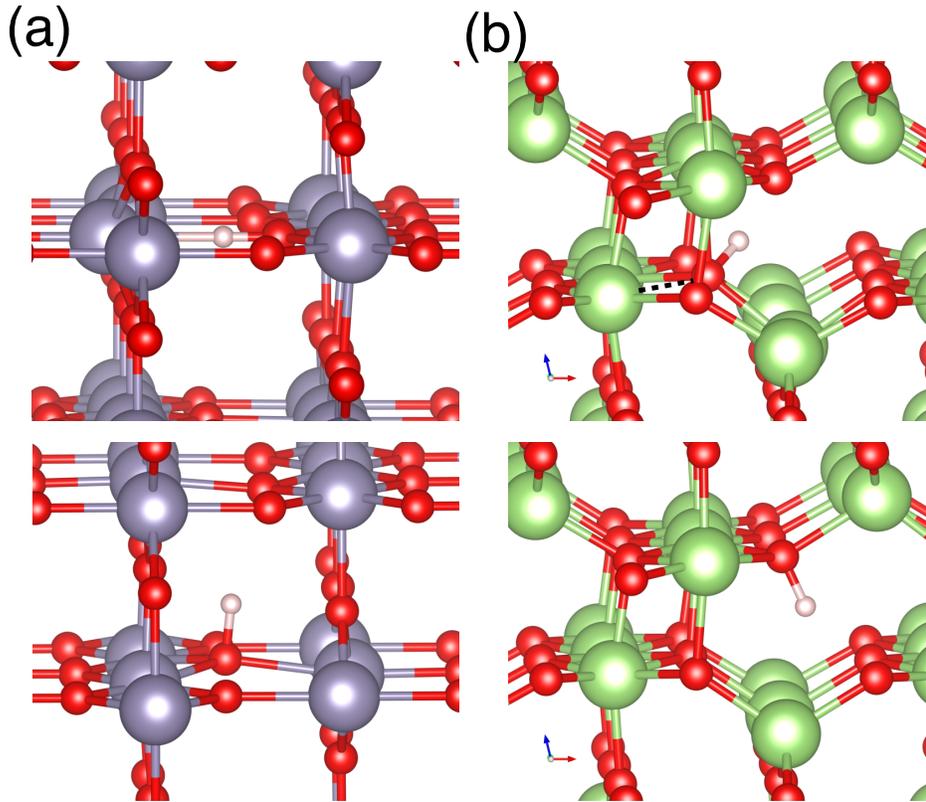
Figure 6.7: Comparison of H_i^+ configurations used in the alignment for (a) rutile SnO_2 and (b) monoclinic β-Ga_2O_3. The top panels represents the bond-center (a) or anti-bonding (b) configurations that disrupt the host bonding. The broken Ga-O bond in (b) is represented by a dashed line. The bottom panels represent the lowest energy configurations for H_i^+ in these structures, but they do not probe the cation-anion bonds and instead occupy anion-derived lone-pairs.

corresponds to a transition between anion- and cation-derived dangling bond states and so is very close to the CNL.[59,69,70] But H_i have been shown to act exclusively as shallow donors in Ga_2O_3 from theory[367] and studies of the electronic analogue muonium.[338] This apparent contradiction is resolved when the very different nature of the bonding of H_i in Ga_2O_3 compared to most other metal oxides is considered. The bonding environment of the O atoms in Ga_2O_3 results in O lone-pairs that may capture H_i and form favourable shallow donors without much influence on the lattice (bottom panels of Fig. 6.7). This is a phenomenon also exhibited by SnO_2. In the spirit of the usual relationship between the CNL and the behavior of H_i in other materi-

als, these lone-pair H_i^+ configurations for Ga_2O_3 and SnO_2 do not probe anion- and cation-derived dangling bond states and so their H_i transition level (denoted $\varepsilon'(+/-)$ in Fig. 6.3(d)) does not correspond to the CNL for these materials. But, if we instead consider higher-formation energy H_i^+ configurations that disrupt the bonding and lead to cation dangling bonds (top panels of Fig. 6.7), it is found that the H_i $\varepsilon(+/-)$ transition level[70] falls 0.51 eV below the CBM for Ga_2O_3, in good agreement with the experimental findings here and the CNL depicted 0.6 eV below the CBM in Fig. 6.3(d). (For SnO_2, it is found that a bond-center H_i^+ configuration that probes Sn dangling bond energies yields a H_i $\varepsilon(+/-)$ transition level 0.48 eV above the CBM, which is also qualitatively and quantitatively in good agreement with previous studies.[364])

6.4 Conclusion

In conclusion, the type and magnitude of band bending at the ($\bar{2}01$) surface of β-Ga_2O_3 was determined for the as-entered and annealed crystal from valence band and core level XPS measurements combined with hybrid DFT valence band calculations. The surface of the as-entered Ga_2O_3 crystal is found to be terminated by O-H groups, resulting in downward band bending of 0.24 eV and electron accumulation with a sheet density of $\sim 5 \times 10^{12}$ cm^{-3}. Cleaning the surface by annealing and removing the adsorbed hydrogen results in a transition from electron accumulation to depletion with upward band bending of 0.26 eV. The observation of electron accumulation at uncleaned and even moderately annealed Ga_2O_3 surfaces has been previously overlooked and the extent of depletion overestimated. This is due in different cases either to use of the wrong band gap energy and/or to an over-reliance upon linear extrapolation of the valence band leading edge for the determination of the VBM position with respect to the Fermi level. These findings are consistent with the charge neutrality level lying ~ 0.6 eV below the CBM, which is in line with other reported properties of Ga_2O_3 and

also the calculated $\varepsilon(+/-)$ transition level for bond disrupting H_i configurations.

7
Summary and Future Work

7.1 Summary

This thesis has primarily focussed on the electronic properties and structure of a class of *'semiconductors'* known as transparent conducting oxides (TCOs). Of the wide and varied selection of materials that make up the TCOs, only two have seen major commercial success. These are SnO_2 and In_2O_3 which are most commonly doped with F and Sn respectively. Understanding the role of the dopant in these materials is key to improving on the performance of current and indeed all TCOs. This task is complicated both experimentally, by very low quantities of the dopant being present, and theoretically, by the breaking of translational symmetry at a dopant centre. Much of the work in this thesis has been an attempt to advance our understanding regarding the effects of doping in TCOs, and provide new insights and ideas regarding both posi-

tive and negative effects of intentionally introducing dopant defects in these materials. A common rule of thumb that has endured for TCOs is that an element in the column on the right-hand-side of the host atom in the periodic table (be it anion or cation) is the natural choice of dopant. Whilst logical and clearly effective in many cases, there are numerous other choices of dopant element which have either gone unexplored, or in cases where their performance has been shown to be superior to the conventional doping method, the mechanism for this has been left unexplained. The work presented here helps explain some of the considerations necessary to improve upon the overly simplified assumptions of the conventional doping method.

Chapter 4 looks at the conventional doping route in SnO_2 doped with F (FTO). Through a combination of theoretical defect formation energy calculations, Hall effect measurements and modelling, infrared reflectivity modelling, and quantitative evidence from XPS and SIMS measurements, the fluorine interstitial was proposed as the likely compensating mechanism for free carriers in FTO. This extrinsic compensating mechanism limits the dopability of FTO, and reduces the mobility meaning the conductivity is self limiting.

In chapter 5, molybdenum is shown, to be a superior dopant in In_2O_3 than the conventional tin, as evidenced through a combination of Hall effect data and optical measurements. Optical modelling gives evidence of lower effective masses in IMO compared to ITO, which explains the relatively high carrier mobilities recorded for IMO, while the mobility of ITO for the same carrier density is much more limited. The transition metal Mo improves upon the carrier mobility in In_2O_3, and this is determined to be due to the minimal hybridization at the CBM of Mo dopant states. Sn dopant levels sit close in energy to the CBM states, and the Sn 5s dopant states share orbital character with the In 5s states at the CBM meaning they hybridize readily. This was determined using HAXPES and IPES measurements, in conjunction with DFT calcula-

tions. Mo has 4d dopant levels which due to structural and Coulombic distortion, are energetically remote and do not mix well with the CBM states in In_2O_3. This level of understanding would not be attainable using only the conventional doping method, as all dopant levels share orbital character with the host in this case.

Additionally, there is growing interest in a particular oxide semiconductor material, Ga_2O_3, which may find uses in high power electronic devices amongst other things. Because interest in Ga_2O_3 is in its relative infancy, many of the fundamental properties of this material have not been fully explored. Amongst these, the surface electronic properties had not been well understood.

In chapter 6 it is determined that Ga_2O_3 displays electron accumulation when the surface is terminated by $O-H$ groups resulting in downward band bending. After surface cleaning, the direction of band bending is reversed and Ga_2O_3 displays electron depletion at the surface. This observation is explained by invoking the theory of the charge neutrality level, which was determined to be $\sim$0.6 eV below the CBM. This is supported by the determination of the $\varepsilon(+/-)$ transition level for H interstitials, which unlike the materials previously studied, do not disrupt the bonding in Ga_2O_3, but are captured by O lone-pairs in the crystal forming favourable shallow donors.

7.2 Future Work

7.2.1 Transparent Conductor Design

The obvious progression of the work presented here is to look for other dopants in TCOs that improve upon the optoelectronic properties of the host relative to the conventional doping route. Luckily, many of these have been screened, albeit often unintentionally, with examples including Ce, Zr, Hf and W in In_2O_3 and Ta in SnO_2. However, the simplistic approach of selecting a transition metal with the correct oxidation

state as a dopant seems to give sporadic success, both in terms of run-to-run varia-
tion in material properties and vast variation of materials between research groups,
indicating there is still more to be understood.

Indeed looking back to figure 1.15, varying levels of success can be seen in achieving
both high mobility and high carrier densities using different dopants in In_2O_3. While
some transition metals give very high mobility films, some appear to be heavily lim-
ited in how easily incorporated into the material they are, limiting the free carrier
density, e.g. see Ce in In_2O_3 in figure 1.15. Why are some dopants more easily incor-
porated into materials than others? Logically, the size of the atom could play a role
in dopability. Larger (or smaller) dopants than the host atom will probably distort the
lattice more, which will hence have an increasing effect on the electronic structure as
more dopant is incorporated. It is not surprising then that the conventional doping
route is so successful at producing films with extremely high carrier densities, as the
atom directly on the right of the cations (or anion) will have a very similar atomic
size and so will not add much strain to the crystal structure. In fact the higher dopa-
bility is precisely the reason why ITO has become the industry standard, achieving
high enough carrier density to offset the poorer mobility in many cases. It seems that
careful selection and understanding of the precise nature of a dopant in a material is
imperative for improved material design, and the general 'rules of thumb' may give
some success, but are ultimately limited in their application.

With this in mind, figure 7.1 gives a schematic representation of how I believe the
work presented in this thesis can be extended and improved upon. This visualises
the research framework used during this work, and incorporates both theoretical and
experimental approaches harmoniously. Starting from the top of figure 7.1, a good un-
derstanding of structural information is imperative for all materials, but is especially
important when considering doping behaviour, and impact of the dopant on the crys-
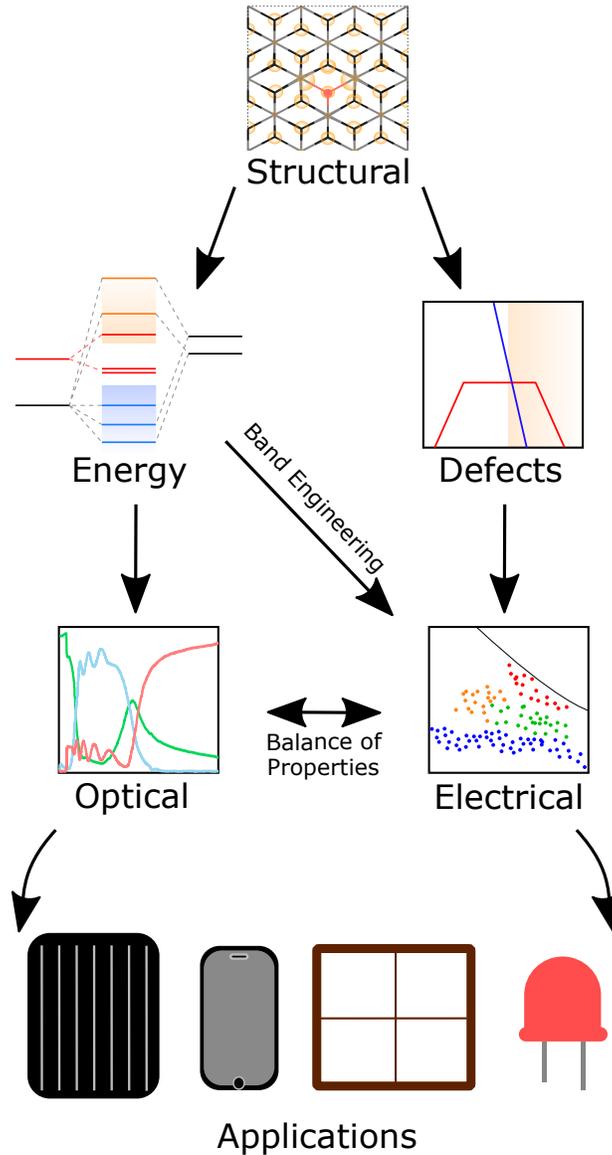
Figure 7.1: Flow diagram of all the considerations necessary for improved TCOs.

tal structure. This can often have a large impact on the electronic properties of the material, as has been shown throughout this thesis. On the next row, knowledge of the electronic structure is important. A great deal of papers report the electronic structure of doped TCO materials, but very few consider the effect on the band structure when the dopant level is ionized. It has been demonstrated that this is an important

consideration that should hopefully be accounted for in future works. Even simple considerations, such as how many dopant atoms per unit cell are required to achieve a realistic free carrier density, are ignored. Very much related to this, consideration of the limiting defects is extremely important also, dictating the behaviour of the materials, both electronically and optically. Only by considering all of these properties for each new material studied do we hope to improve upon, or explain the limitations of, the existing TCO materials. This should then directly feed into improving the the materials used, making them more functional or even creating new applications. These improvements can be evaluated experimentally through optical and electronic measurements, the results of which will back-inform the assumptions and predictions made in the previous parts of the schematic. Finally, the improved product can be incorporated in existing applications or new technologies.

7.2.2   FURTHER INVESTIGATIONS INTO Ga_2O_3

It may seem strange that the focus of the final results chapter of this thesis was quite different from the others. Whilst undeniably true, the reason I wanted to include these results in this thesis was because β-Ga_2O_3 is a very exciting material to work with currently. Its commercial utilization is in its infancy mainly because the most common material of choice for power electronics is currently Si. However, there is a large push to replace Si (which performs quite poorly due to its low breakdown voltage), with wide-band gap semiconductors. Out of these, β-Ga_2O_3 offers superior high power device performance (due to greater band gap and higher breakdown voltage) and the possibility of gains in cost effectiveness over its rival materials including GaN and SiC, suggesting that β-Ga_2O_3 is a good candidate for power switching. Indeed there are in fact diodes made from SiC commercially available, showing the demand for such high power devices is high.

One important factor that needs to be considered when forming devices using β-Ga_2O_3 is the electronic behaviour at the surface or interfaces. This was demonstrated in chapter 6 for the simplest case of unintentional contamination of the crystal surface. However, further investigation needs to be carrier out to get a better understanding of the how the surface of β-Ga_2O_3 behaves. Questions such as i) how does an atomically clean crystal surface behave when forming a device relative to a contaminated one ii) what happens to the surface properties when dosing the surface with other gases (or in a more controllable way such as using pure H_2 or H_2O for example) iii) does the crystal orientation make a difference to this iv) how to properly determine the level of band bending when a β-Ga_2O_3-semiconductor (or metal) interface is created in a device, are just some examples of works that could be carried out in the future for this exciting material.

A

Selection rules for electric dipole transitions

It was previously touched upon in chapter 1 that the optical transitions between the VBM and CBM in In_2O_3 are forbidden, and that the first allowed transition in this material originates from a valence band ~ 0.8 eV below the VBM to the CBM. This information is often still not known for this material, despite it now being quite well evidenced by experiment. However, we can perform a similar analysis to that suggested at by Walsh et al.[25] to show this is the case for In_2O_3.

In_2O_3 usually adopts the body centred cubic bixbyite structure, which has space group $Ia\bar{3}$, No. 206 (international point group symbol: $m\bar{3}$, Schoenflies space group symbol: T_h^7). The character table for the $m\bar{3}$ point group (Schoenflies point group T_h) is as follows[378]:

Table A.1: Character table for T_h point group, note: $w = exp\left(\frac{2i\pi}{3}\right)$

$T_h(m\bar{3})$	E	$4C_3$	$4(C_3)^2$	$3C_2$	i	$4(S_6)^5$	$4S_6$	$3\sigma_h$	functions
Mult.	1	4	4	3	1	4	4	3	-
A_g	1	1	1	1	1	1	1	1	$x^2 + y^2 + z^2$
^1E_g	1	w	w^2	1	1	w	w^2	1	$(2z^2 - x^2 - y^2,$
^2E_g	1	w^2	w	1	1	1	w^2	1	$x^2 - y^2)$
T_g	3	0	0	-1	3	0	0	-1	(xy, yz, xz)
A_u	1	1	1	1	-1	-1	-1	-1	-
^1E_u	1	w	w^2	1	-1	$-w$	$-w^2$	-1,	-
^2E_u	1	w^2	w	1	-1	$-w^2$	$-w$	-1	-
T_u	3	0	0	-1	-3	0	0	1	(x, y, z)

The rate of optical interband transitions at the Γ-point depends upon the momentum matrix element $\mathbf{p}_{if}$ which can be represented by

$$\mathbf{p}_{fi} = \int \Psi_f^*(\mathbf{r})\mathbf{p}\Psi_i(\mathbf{r})\,d\mathbf{r} \tag{A.1}$$

where Ψ_i and Ψ_f are the initial and final wave functions respectively. As the electron momentum operator $\mathbf{p}$ is a vector quantity, its three components (p_x, p_y and p_z) must belong to the irreducible representation T_u (note that in non cubic systems components of $\mathbf{p}$ may transform as one or two-dimensional irreducible representations rather than a single three-dimensional irreducible representation). The two wave functions Ψ_i and Ψ_f have their own parities, which when multiplied will give the parity of their product. The selection rules rely on the constraints imposed on the matrix elements of the electric-dipole operator. The important rule here is known as the matrix-element theorem and states that the matrix element between an operator $\mathbf{p}$ and two wave functions Ψ_i and Ψ_f can differ from zero only when the direct product of the representations of $\mathbf{p}$ and Ψ_i contains an irreducible representation of Ψ_f. The electric-dipole operator has odd parity, and so its matrix element is zero between two states of the same parity.

In representation theory, the transformation properties in equation A.1 are handled through what is called the direct product, allowing us to expand the integrand into irreducible representations of the point group[54]:

$$\Gamma_{equiv} = \Gamma_{final\ state} \otimes \Gamma_{operator} \otimes \Gamma_{initial\ state}. \tag{A.2}$$

Focussing on transitions at the zone centre, we can consider the irreducible representations of the VBM (T_g) and CBM (A_g). The electric dipole operator transforms with the irreducible representation T_u which allows us to define our selection rules through the direct product:

$$\Gamma_{equiv} = A_g \otimes T_u \otimes T_g \tag{A.3}$$

allowing us to determine the following characters for the equivalent representation of the integrand in equation A.1

Table A.2: Character table for the equivalent representation Γ_{equiv} of the T_h point group.

	E	$4C_3$	$4(C_3)^2$	$3C_2$	i	$4(S_6)^5$	$4S_6$	$3\sigma_h$
Γ_{equiv}	9	0	0	3	-9	0	0	-3

The orthogonality theorem presented in equation A.4 gives the basic orthonormality relations used to set up character tables.

$$\sum_k \chi_i(C_k)^* \chi_j(C_k) N_k = h\delta_{ij} \tag{A.4}$$

where h is the order of the group ($h = 24$ in this case), N_k is the number of elements of the class C_k (or the multiplicity), $\chi_{i,j}$ are the characters for the selected irreducible representations of a particular class and δ_{ij} is the Kronecker delta. This equation has a delta function which relates to the orthogonality of the irreducible representations

(i.e. the characters of an irreducible representation form a set of orthogonal vectors in group-element space). If a representation is reducible then the characters will generally not obey this orthogonality relation. In our case, where we may have constructed a reducible representation in table A.2, we need to adjust equation A.4 slightly to account for the possibility of this representation not being orthogonal to the irreducible representations.

$$\sum_k \chi(C_k)^* \chi_j(C_k) N_k = ha. \tag{A.5}$$

Here we replace the character associated with an irreducible representation from table A.1 with one from our equivalent representation $\chi(C_k)$ and replaced the delta function with a constant a that can take any integer value, which essentially accounts for the number of instances that a selected irreducible representation is found in the reducible representation. Using equation A.5, the T_h character table, and the characters for the equivalent representation above in table A.2 we see that the sum does not contain the totally symmetric irreducible representation A_g, i.e. this electric-dipole transition is forbidden by selection rules. If instead of the initial state selected (being T_g) we select the state with irreducible representation T_u and perform the same analysis, we find the equivalent representation contains the totally symmetric irreducible representation A_g, hence this transition is allowed. This is in agreement with the results of Walsh et al.[25] One useful table to refer to for selection rules is known as the multiplication table of irreducible representations[378].

This table shows the irreducible components of the direct product between irreducible representations.

Table A.3: Multiplication table of irreducible representations of the T_h point group

$T_h(m\bar{3})$	A_g	A_u	^1E_g	^1E_u	^2E_g	^2E_u	T_u	T_g
A_g	A_g	A_u	^1E_g	^1E_u	^2E_g	^2E_u	T_u	T_g
A_u	-	A_g	^1E_u	^1E_g	^2E_u	^2E_g	T_g	T_u
^1E_g	-	-	^2E_g	^2E_u	A_g	A_u	T_u	T_g
^1E_u	-	-	-	^2E_g	A_u	A_g	T_g	T_u
^2E_g	-	-	-	-	^1E_g	^1E_u	T_u	T_g
^2E_u	-	-	-	-	-	^1E_g	T_g	T_u
T_u	-	-	-	-	-	-	$A_g +^1 E_g +^2 E_g + 2T_g$	$A_u +^1 E_u +^2 E_u + 2T_u$
T_g	-	-	-	-	-	-	-	$A_g +^1 E_g +^2 E_g + 2T_g$

B

Plasma reflectivity in a free electron gas and the dipole oscillators

Much of the content of this section comes from undergraduate text books. In Particular I refer to the book on optical properties of solids by M. Fox [129]. This should provide the basic structure for creating a more complicated oscillator model applicable to semiconductor systems.

Free electron gas

We begin by giving some useful definitions. The response of a dielectric medium to an external electric field is characterized by three vectors. The electric field strength $\mathbf{E}$, the polarization $\mathbf{P}$ and the electric displacement $\mathbf{D}$. The polarization is defined as the net dipole moment per unit volume which comes from the application of a field

on the medium producing polarization by forces exerted on the positive and negative charges of the atoms. The field can force positive and negative charges apart and induce a dipole parallel to the field direction in atoms with no permanent dipole moment, else the field applies a torque to align the dipoles parallel to the field if they are naturally randomly oriented. Either way this generates a net dipole moment within the dielectric and hence a polarization

$$\mathbf{P} = \varepsilon_0 \chi \mathbf{E}. \tag{B.1}$$

Equation B.1 defines the polarization, which is parallel to $\mathbf{E}$, ε_0 is the electric permittivity of free space and χ the electric susceptibility of a medium. We assume in these equations an isotropic medium. The electric displacement is given as

$$\mathbf{D} = \varepsilon_0 \mathbf{E} + \mathbf{P} = \varepsilon_0 \varepsilon_r \mathbf{E} \tag{B.2}$$

where the relative dielectric constant $\varepsilon_r = 1 + \chi$. The microscopic response of a material is mostly determined from the $\mathbf{P}$, which then allows us to determine ε_r which determines the optical properties of a material.

The equation of motion for free electron in an AC field ($\mathbf{E}(t)$) is given by:

$$m_0 \frac{d^2 x}{dt^2} + m_0 \gamma_p \frac{dx}{dt} = -eE(t) = -eE_0 e^{-i\omega t} \tag{B.3}$$

Here the first term represents the acceleration of the free electron, while the second term is the frictional damping force of the medium (γ_p is the damping rate), and the final term is the driving force exerted by the light. Since the electric field of an AC electric field varies sinusoidally we can choose a solution to this equation of a similar sinusoidal form $x(t) = x_0 e^{-i\omega t}$ giving a solution to equation B.3:

$$x(t) = \frac{eE(t)}{m_0(\omega^2 + i\gamma_p\omega)}. \tag{B.4}$$

This is the displacement of an electron from its equilibrium position and so the polarization of a free electron gas is now given as $P = -Nex$ where N is the number of electrons per unit volume. Using equation B.2, we know for a free electron system:

$$\mathbf{D} = \varepsilon_0\varepsilon_r E = \varepsilon_0 E - \frac{Ne^2 E}{m_0(\omega^2 + i\gamma_p\omega)} \tag{B.5}$$

and therefore

$$\varepsilon_r(\omega) = 1 - \frac{Ne^2}{\varepsilon_0 m_0}\frac{1}{\omega^2 + i\gamma_p\omega} \tag{B.6}$$

which is usually written as

$$\varepsilon_r(\omega) = 1 - \frac{\omega_p^2}{\omega^2 + i\gamma_p\omega} \tag{B.7}$$

where the plasma frequency ω_p is given as

$$\omega_p = \left(\frac{Ne^2}{\varepsilon_0 m_0}\right)^{\frac{1}{2}}. \tag{B.8}$$

We can relate the relative dielectric constant of a material to the complex refractive index. the refractive index can be described by

$$\tilde{n} = n + ik \tag{B.9}$$

where n is the refractive index and k is the extinction coefficient. Because the dielectric function can be complex we can write this as

$$\tilde{\varepsilon}_r = \varepsilon_1 + i\varepsilon_2. \tag{B.10}$$

ε_1 and ε_2 are the real and imaginary parts of the dielectric function. These two equation are related through Maxwell's equations to give

$$\tilde{n} = \sqrt{\tilde{\varepsilon}_r}. \tag{B.11}$$

Combining equations B.9, B.10 and B.11 gives

$$\varepsilon_1 = n^2 - k^2$$

$$\varepsilon_2 = 2nk$$

$$n = \frac{1}{\sqrt{2}}\left(\varepsilon_1 + (\varepsilon_1^2 + \varepsilon_2^2)^{\frac{1}{2}}\right)^{\frac{1}{2}}$$

$$k = \frac{1}{\sqrt{2}}\left(-\varepsilon_1 + (\varepsilon_1^2 + \varepsilon_2^2)^{\frac{1}{2}}\right)^{\frac{1}{2}}. \tag{B.12}$$

Finally to relate these functions to an experimental observable we use the equation

$$R = \left|\frac{\tilde{n}-1}{\tilde{n}+1}\right|^2 = \frac{(n-1)^2 + k^2}{(n+1)^2 + k^2}. \tag{B.13}$$

For context, a lightly damped system (i.e. inserting $\gamma \sim 0$ in equation B.7) displays a reflectivity and dielectric response as shown in figure B.1. Because $\gamma = 0$ the dielectric constant $\tilde{\varepsilon}_r$ becomes negative only when $\omega < \omega_p$ ($\tilde{n}$ becomes imaginary), and takes a positive value when $\omega > \omega_p$. At $\omega = \omega_p$ the refractive index and the dielectric function are zero. Physically this means the reflectivity is 100% for frequencies below ω_p ($\tilde{n}$ has a large imaginary part), and approaches zero when $\omega >> \omega_p$. This idealized case of an undamped free electron gas can be adjusted for many real physical systems.
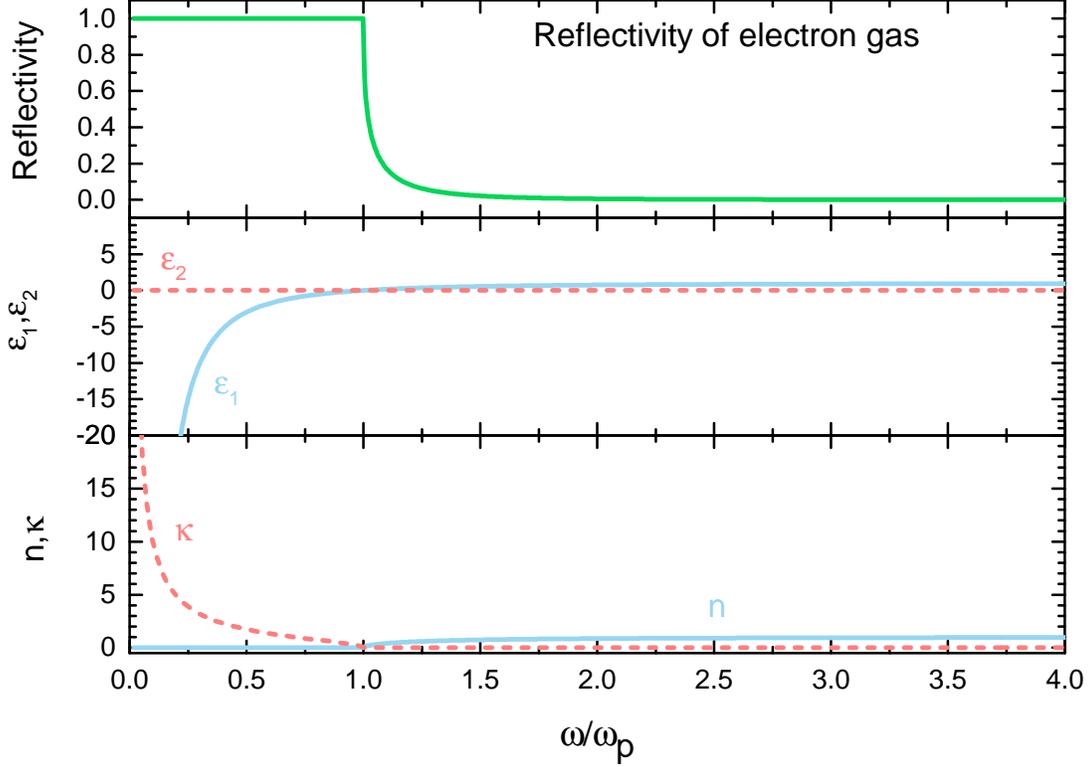
Figure B.1: Reflectivity, the real ε_1 and imaginary ε_2 parts of the relative dielectric constant, and the refractive index n and extinction coefficient k of an undamped free electron gas.

THE DIPOLE OSCILLATOR MODEL

In a very similar approach to that used above, we can create a model for the dielectric constant of an atom with bound electron with a single resonant frequency ω_0. Equation B.3 becomes

$$ m_0 \frac{d^2 x}{dt^2} + m_0 \gamma \frac{dx}{dt} + m_0 \omega_0^2 x = -eE(t) = -eE_0 e^{-i\omega t}. \tag{B.14} $$

Notice how this is exactly the same equation of motion as in equation B.3, only with an extra term representing the restoring force between a dipole oscillator (is this is now a bound system). Oscillating dipole can lose energy in collision processes which in solids is typically through a phonon interaction. Note that we assume the mass of

the electron is far less than the nucleus, i.e. the nucleus is essentially motionless. Now by following the exact same procedure outlined above, only with the extra resonant term in equation B.4 we can arrive at the relative dielectric constant

$$\varepsilon_r(\omega) = 1 + \chi - \frac{Ne^2}{\varepsilon_0 m_0} \frac{1}{\omega^2 + i\gamma\omega - \omega_0^2}. \tag{B.15}$$

Taking the limits to the two extremes, i.e. $\omega \to 0$ (the static case) and $\omega \to \infty$ (the high frequency case) we find

$$\varepsilon_r(0) = 1 + \chi - \frac{Ne^2}{\varepsilon_0 m_0 \omega_0^2} \tag{B.16}$$

$$\varepsilon_r(\infty) = 1 + \chi \tag{B.17}$$

and so we can write

$$\varepsilon_r(0) - \varepsilon_r(\infty) = \frac{Ne^2}{\varepsilon_0 m_0 \omega_0^2}. \tag{B.18}$$

giving

$$\varepsilon_r(\omega) = \varepsilon_r(\infty) + \frac{(\varepsilon_r(0) - \varepsilon_r(\infty))\omega_0^2}{\omega^2 + i\gamma\omega - \omega_0^2}. \tag{B.19}$$

Figure B.2 shows the dielectric response for a dipole oscillator. The resonant frequency can be seen at the peak position of ε_2, corresponding to a drop in reflectivity (increased absorption) around ω_0. ε_1 increases from the static regime $\varepsilon(0)$ approaching ω_0 and then drops to a negative value at ω_0, before being restored at the high frequency regime $\varepsilon(\infty)$. The width of ε_1 and ε_2 is determined by the damping γ, i.e. damping of the oscillators causes line broadening in reflectivity spectra.
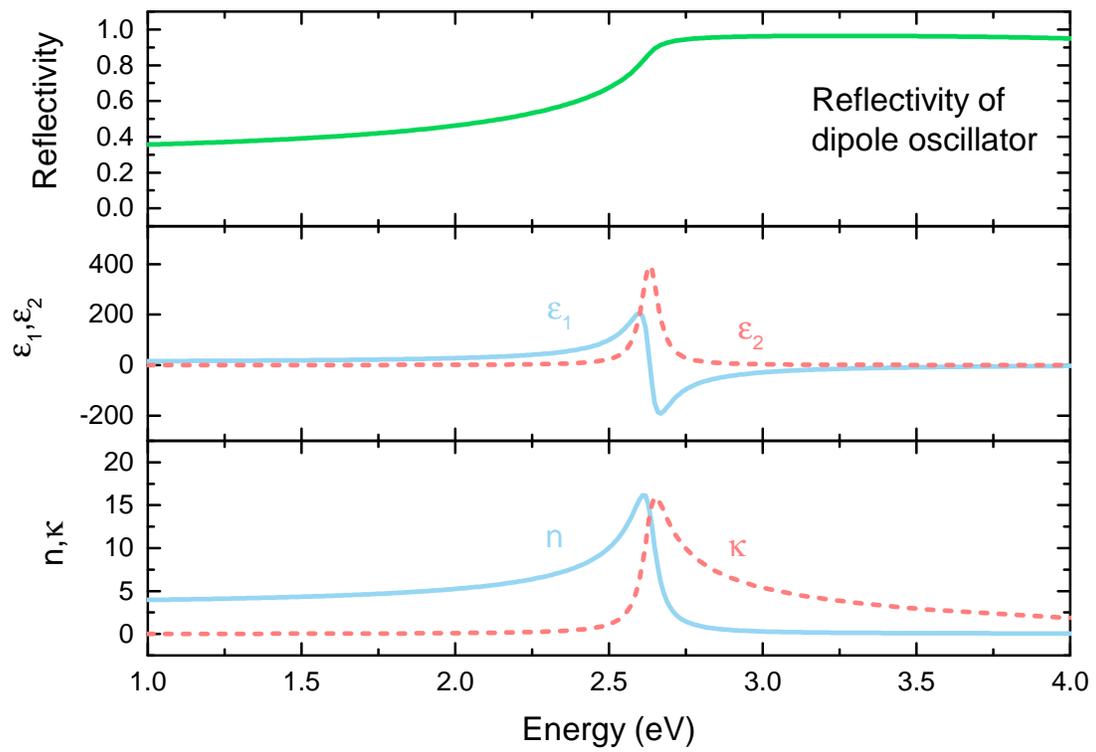
Figure B.2: Energy dependence of the Reflectivity, the real and imaginary parts of the dielectric constant and refractive index for a dipole oscillator around resonance.

C

Theory of Photoemission

Since the theory of PES is deeply seated in the physics of quantum mechanics and many-body interactions, it may be instructive to discuss these ideas here. The theory is applicable on a much grander scale than simply PES, but this gives a very good platform for direct application. Hopefully this chapter should provide a short but thorough introduction into the quantum mechanical description of PES, which may inform some of the discussion in chapter 2.1.1. Much of this discussion is based on that of Hüfner[95] and much of this chapter is mainly for the interested reader (as I am myself), and may not reflect directly onto the work carried out.

The transition probability of an electron being photo-excited from an initial state ϕ_i to final state ϕ_f is given by Fermi's golden

$$W \propto \frac{2\pi}{\hbar} |\langle \phi_f | H' | \phi_i \rangle|^2 \delta(E_f - E_i - h\nu) \tag{C.1}$$

where H' is a small perturbation produced caused by the light interaction which is evaluated by classical electromagnetism as the effect of the electromagnetic field on the atom. This can be related to the Hamiltonian of the electron before the field was applied $H_0 = \frac{\mathbf{p}^2}{2m_0} + V(\mathbf{r})$ (here $V(\mathbf{r})$ is the potential energy of the electron in the atom). The Hamiltonian of the electron in an electric field which is given by

$$
\begin{aligned}
H &= \frac{1}{2m_0}(\mathbf{p} + e\mathbf{A})^2 + V(\mathbf{r}) \\
&= H_0 + \frac{e}{2m_0}(\mathbf{A} \cdot \mathbf{p} + \mathbf{p} \cdot \mathbf{A}) + \frac{e^2}{2m_0}\mathbf{A} \cdot \mathbf{A}
\end{aligned}
\tag{C.2}
$$

where $\mathbf{A}$ is the vector potential and $\mathbf{p} = i\hbar\nabla$ is the momentum operator, and the change in momentum due to the applied field is given as $\mathbf{p} \to \mathbf{p} - e\mathbf{A}$. Therefore the perturbation due to the light interaction can be given as

$$
H' = \frac{e}{2m_0}(\mathbf{A} \cdot \mathbf{p} + \mathbf{p} \cdot \mathbf{A}) + \frac{e^2}{2m_0}\mathbf{A}^2.
\tag{C.3}
$$

If $\mathbf{A}^2$ is neglected (ignoring the the two photon interaction as it is weak), then because the first two terms in equation C.2 commute we can write $\mathbf{A} \cdot \mathbf{p} = \mathbf{p} \cdot \mathbf{A}$ and using the definition of $\mathbf{p} = -i\hbar\nabla$ for an arbitrary function $f(r)$ we see $(\mathbf{p} \cdot \mathbf{A}f(r) = \mathbf{A} \cdot (-i\hbar\nabla f) + (-i\hbar\nabla \cdot \mathbf{A})f)$. Now using the fact that $\nabla \cdot \mathbf{A} = 0$ in the Coulomb gauge, we can write

$$
H' = \frac{e}{m_0}(\mathbf{p} \cdot \mathbf{A}).
\tag{C.4}
$$

This returns us to the dipole approximation discussed briefly in chapter 2.1.1, where $A = A_0 e^{i(\mathbf{k} \cdot \mathbf{r} - \omega t)} = A_0(1 + i(\mathbf{k} \cdot \mathbf{r}) + \frac{1}{2}(i\mathbf{k} \cdot \mathbf{r})^2 + ...) \sim A_0$ for long wavelengths. So to evaluate the transition probability of a photoemission event we need to evauluate the matrix element in equation C.1 which is now described by

$$M_{12} = \frac{e}{m_0} |\langle \phi_f | \mathbf{p} \cdot \mathbf{A}_0 | \phi_i \rangle| \tag{C.5}$$

as well as knowledge of the initial and final wave functions of the electrons. [*] In the one electron view, the final state of the system is down a single electron so the matrix element consists of a product of a one-electron matrix element and an (N-1)-electron system. In the frozen orbital approximation, assuming that the initial and final state of the atom are the same Φ_{N-1} (the initial φ_i and final φ_f states of the photoemitted electron are treated explicitly only), that is no rearrangement takes place in the atom after photoemission, then the binding energy of the electron is equal to the orbital energy of the electron, sometimes known as Koopmans' theorem $E_B \sim -E_{orb.}$. This approximation is obviously too simplistic as relaxation must occur to minimise energy after photoemission, but is still useful to aid in basic understanding.

If the final state of the N-1 electron system has a number m of available excited states with wave function $\Phi_f^m(N-1)$ of energy $E_m(N-1)$ then the transition matrix element is given by the overlap integral of the initial states, and the sum of that of the final states

$$\begin{aligned} \langle \phi_f | \mathbf{r} | \phi_i \rangle &= \langle \varphi_f^{E_k} | \mathbf{r} | \varphi_{i,k} \rangle \sum \langle \Phi_{f,k}^m(N-1) | \Phi_{i,k}(N-1) \rangle \\ &= \langle \varphi_f^{E_k} | \mathbf{r} | \varphi_{i,k} \rangle \sum C(N-1). \end{aligned} \tag{C.6}$$

Here we explicitly include the orbital k from where the photoelectron originates. $|C(N-1)|^2$ is the probability that after the removal of an electron from orbital k of the N-electron system, the excited state of the system is left in the state m represented by

[*]note that using the commutation relations we can say $\langle \phi_f | \mathbf{A} \cdot \mathbf{p} | \phi_i \rangle \propto \langle \phi_f | \mathbf{A} \triangledown \cdot \mathbf{V} | \phi_i \rangle \propto \langle \phi_f | \mathbf{A} \cdot \mathbf{r} | \phi_i \rangle$. We can write the transition matrix element as the overlap integral of the initial and final states of the photoemitted electron, and the N-1 electron system $\langle \phi_f | \mathbf{r} | \phi_i \rangle = \langle \varphi_f | \mathbf{r} | \varphi_i \rangle \langle \Phi_f(N-1) | \mathbf{r} | \Phi_i(N-1) \rangle$.

$\Phi_f^m(N-1)$. For some systems $|C(N-1)|^2 \sim 1$ (i.e. $\Phi_f(N-1) \sim \Phi_i(N-1)$ meaning the state $k = m$) so there is only one available excited state and only one peak is observed, while for some other systems (such as strongly correlated ones) there are many non-zero values of $|C(N-1)|^2$ and so many satellites are observed as well as the main peak when $m = k$. From here we can write the intensity (or photocurrent) of a core-level from equation C.1 and C.6 as

$$I \propto \sum |\langle \varphi_f^{E_k}|\mathbf{r}|\varphi_{i,k}\rangle|^2 \sum |C(N-1)|^2 \delta(E_f - E_i - \hbar\omega) \qquad \text{(C.7)}$$

where E_f and E_i are the final and initial energy states of the system (E_f being comprised of the kinetic energy of the photoemitted electron and the final state (N-1) atomic system, while E_i is simply the N electron system). The photocurrent of core levels therefore relies heavily on the number of final states making up $\sum |C(N-1)|^2$, which in turn will determine the number of satellites also. This term is often called the spectral function. When looking at the valence states in photoemission we find heavy wave vector dependence of the states (as there are many hybridized lines, some displaying very parabolic behaviour with very de-localized electrons). The expression for the photocurrent in this case is given as

$$I \propto \sum |M_{if}|^2 \delta(E_f - E_i - \hbar\omega)\delta(E - E_f + \varphi)\delta(\mathbf{k_i} + \mathbf{G} - \mathbf{k_f}). \qquad \text{(C.8)}$$

For equation C.8 we make use of the periodicity of the Bloch function with respect the the reciprocal lattice vector $\mathbf{G}$. Here the first delta function ensures conservation of energy, as it did in equation C.7. The second delta function is added simply to ensure only electrons eith energy above the vacuum level (the work function φ is accounted for) are detected. Finally, the last delta function ensures conservation of momentum up to the reciprocal lattice vector. Due to the Fresnel equations we can

separate the momentum conservation term to conserve parallel momentum to the surface if desired.

$$I \propto \sum \left|M_{if}\right|^2 \delta(\mathbf{k_{i\parallel}} + \mathbf{G}_\parallel - \mathbf{k_{f\parallel}}) \delta(E_f - E_i - \hbar\omega)$$
$$\delta(E - E_f + \varphi)\delta(\mathbf{k_i} + \mathbf{G} - \mathbf{k_f}). \tag{C.9}$$

This geometric consideration leads to interesting physics such as being able to describe the angle resolved photoemission process, where the band structure can be mapped over the perpendicular electron momentum. This will not be considered here however. Also, further terms can be considered in equation C.9 to account for the probability of photoelectrons scattering, or not penetrating the sample surface etc. These are ignored here.

The last consideration we will make to gain a full description of the basic photoemission precess is that of the many-body interactions. Electrons in a solid interact with one-another, which means the created core-hole state from the photoemission process can exist in one of the m final states of the system. Returning to the definition of the spectral function

$$A(\mathbf{k}, E) = \sum_m \left|\langle\Phi_f^m(N-1)|\Phi_i(N-1)\rangle\right|^2 \tag{C.10}$$

we can rewrite this equation in terms of the annihilation operator a^- which in the language of second quantization acts to remove a single electron from the ground state system

$$A(\mathbf{k}, E) = \sum_m \left|\langle\Phi_f^m(N-1)|a^{-1}|\Phi_i(N)\rangle\right|^2. \tag{C.11}$$

Equation C.11 describes the probability of an electron being liberated from a system

in the ground state.

The spectral function is integral in describing the electronic behaviour of a many-body system. It is also convenient to write in terms of the single-particle greens function $G(\mathbf{k}, E)$, used to describe the probability that an electron at a position r_1 at time $t = 0$ will be found at position r_2 at a later time t. In terms of energy and momentum, it is equivalent to say the Green's function decribes the probability of an electon in state $\mathbf{k}_1$ is found in $\mathbf{k}_2$ if the a scattering process occurs with energy transfer E. The imaginary part of the Green's function can be shown to be equal to the spectral function[95]

$$A(\mathbf{k}, E) = \frac{1}{\pi}|ImG(\mathbf{k}, E)| \tag{C.12}$$

For a non-interacting system when $\Phi_f(N-1) = \Phi_i(N-1)$ and the spectral function is described by a delta function at $E = E^0(\mathbf{k})$, the one-electron energy of the system $(A^0(\mathbf{k}, E) = \frac{1}{\pi}\delta(E - E^0(\mathbf{k}))$.[†] This is another representation of Koopman's energy.

Now the electron-electron interaction can be accounted for by adding to the single-particle electron energy $E^0(\mathbf{k})$ with what is known as the self energy $\Sigma(\mathbf{k}, E)$ describing the renormalized structure

$$G(\mathbf{k}, E) = \frac{1}{E - E^0(\mathbf{k}) - \Sigma(\mathbf{k}, E)} \tag{C.13}$$

and the spectral function is now

$$A(\mathbf{k}, E) = \frac{1}{\pi}\frac{Im\Sigma(\mathbf{k}, E)}{(E - E^0(\mathbf{k}) - Re\Sigma(\mathbf{k}, E))^2 + (Im\Sigma(\mathbf{k}, E))^2}. \tag{C.14}$$

This reflects equation 2.30 not by coincidence. The lifetime of the electronic state is given by $\tau = \frac{1}{2Im\Sigma(\mathbf{k},E)}$ which gives rise to the finite width of the spectral feature

---

[†]The one-electron Green's function is given by $G^0(\mathbf{k}, E) = \frac{1}{E-E^0(\mathbf{k})-i\delta}$ where $E^0(\mathbf{k}) = \frac{\hbar^2\mathbf{k}^2}{2m_0}$.

through the uncertainty principle $\Gamma = \frac{\hbar}{\tau}$. If the self energy is small (for a weakly interacting system), the Green's function in equation C.13 can be decomposed into two parts, the coherent and incoherent parts.

$$A(\mathbf{k}, E) = \underbrace{\frac{1}{\pi} \frac{Z_k Im(E^1(\mathbf{k}))}{[E - Re(E^1(\mathbf{k}))]^2 + [Im(E^1(\mathbf{k}))]^2}}_{coherent \ part} + \underbrace{(1 - Z_k)A_{inc}}_{incoherent \ part} \qquad \text{(C.15)}$$

where Z_k is a normalisation constant and $E^1(\mathbf{k}) = E^0(\mathbf{k}) + \Sigma(\mathbf{k}, E^1(\mathbf{k}))$. The corresponding Green's function looks like

$$G(\mathbf{k}, E) = \underbrace{\frac{Z_k}{E - (Re(E^1(\mathbf{k})) + iIm(E^1(\mathbf{k})))}}_{coherent \ part} + \underbrace{(1 - Z_k)G_{inc}}_{incoherent \ part} \qquad \text{(C.16)}$$

The coherent part equation C.16 resembles that of the one electron greens function G^0 width a slightly renormalized mass. The electrons in the interacting system move as if free electrons with a renormalized energy, hence these are known as quasi-particles. The coherent part is approximated as the main emission line. The incoherent part of the equation makes up the satellite features (e.g. plasmons).

D

Optical transitions and the absorption coefficient

D.0.1 Absorption in Semiconductors

As with the photoemission process, some of the basic physical principles associated with optical spectroscopy should be discussed. This is in keeping with some previous section, and in fact many parallels between the photoemission process and photoabsorption can be made. As in the photoemission process, where an electron is photoexcited from a ground state to one outside of the material, in optical absorption an incoming photon excites a ground state electron into an unoccupied final state in the conduction band, usually just above the band gap. This is most commonly seen when looking at the transmission of light, which reduces greatly when the energy of incoming photons exceeds the band gap and absorption starts to occur. Many of the details

of this section can be found in text books such as Hamagucie[168], Basu[379] and Yu and Cardona[26], as well as a detailed account given in the PhD thesis of M. Birkett[380].

Much in the same treatment as in appendix C (because of its significance here we will reiterate some of what is discussed in appendix C), we can make use of the one-electron Hamiltonian in the presence of an a perturbation from electromagnetic radiation. The unperturbed Hamiltonian is

$$H_0 = \frac{\mathbf{p}^2}{2m} + V(\mathbf{r}) \tag{D.1}$$

where $\mathbf{p}$ is the electron momentum operator, and $V(\mathbf{r})$ is the one-electron potential. A perturbation due to the interaction of light can be included by adjusting the momentum of the electron*. This gives the perturbed Hamiltonian as

$$H = \frac{1}{2m} \left(\mathbf{p} + e\mathbf{A} \right)^2 + V(\mathbf{r}) \tag{D.2}$$

which we showed in appendix C to be equal to

$$H = H_0 + \frac{e}{m} \mathbf{A} \cdot \mathbf{p} = H_0 + H'. \tag{D.3}$$

Now we need to evaluate the transition probability, which involves evaluating the transition matrix element for a transition between an initial valence state to final conduction state given as

$$M_{fi} = \langle f|H'|i \rangle = \left(\frac{e}{m} \right) \langle f|\mathbf{A} \cdot \mathbf{p}|i \rangle \tag{D.4}$$

To evaluate the transition rate and hence determine the absorption coefficient for a given material we are therefore required to evaluate $|\langle f|\mathbf{A} \cdot \mathbf{p}|i \rangle|^2$. We can do this

---

*This is done by introducing the scalar and vector potentials $\varphi(\mathbf{r}, t)$ and $A(\mathbf{r}, t)$. Due to gauge invariance, we are bale to choose the Coulomb gauge which is mathematically most convenient here so that $\varphi = 0$ and $\nabla \cdot A = 0$.

under two assumptions. Firstly, we assume the forms of the initial (valence) and final (conduction) wave functions to be Bloch functions (which is already assumed when describing the interaction of light)

$$
\begin{aligned}
|i\rangle &= \frac{1}{\sqrt{V}} u_{i,k_i}(\mathbf{r}) e^{i(\mathbf{k_i} \cdot \mathbf{r})} \\
|f\rangle &= \frac{1}{\sqrt{V}} u_{f,k_f}(\mathbf{r}) e^{i(\mathbf{k_f} \cdot \mathbf{r})}.
\end{aligned}
\tag{D.5}
$$

Secondly we need to assume the form of the vector potential $\mathbf{A}$. This can be done by writing $\mathbf{A}$ as $A\hat{\mathbf{e}}$ where $\hat{\mathbf{e}}$ is a unit vector in the direction of $\mathbf{A}$.

$$
\mathbf{A} = \frac{A_0}{2} \left(e^{i(\mathbf{q} \cdot \mathbf{r} - \omega t)} + e^{-i(\mathbf{q} \cdot \mathbf{r} - \omega t)} \right) \hat{\mathbf{e}}
\tag{D.6}
$$

where A_0 is related to the amplitude of the incident electric field and $\mathbf{q}$ is the wavevector of the electromagnetic field. The two exponential terms essentially describe the same physical process in reverse, the first giving the absorption process, while the latter the stimulated emission process (an electron dropping from the conduction band to an unfilled state in the valence band, emitting a photon. The energy $\hbar\omega$ has a different sign in the delta function in equation D.7 as ωt does in equation D.6). Now substituting this into the equation for Fermi's golden rule which describes transition rates, allows us to determine the transistion probability per unit time for a transition from state $|i\rangle$ to $|f\rangle$ as

$$
\begin{aligned}
R &= \frac{2\pi}{\hbar} \left| \langle f | \frac{e}{m} \mathbf{A} \cdot \mathbf{p} | i \rangle \right|^2 \delta(E_f - E_i - \hbar\omega) \\
&= \frac{2\pi}{\hbar} \frac{e^2}{4m^2} |A_0^2| \left| \langle f | e^{i\mathbf{q} \cdot \mathbf{r}} \hat{\mathbf{e}} \cdot \mathbf{p} | i \rangle \right|^2 \delta(E_f - E_i - \hbar\omega)
\end{aligned}
\tag{D.7}
$$

where we have completely neglected the stimulated emission term in equation D.6 as

our results are for the ground state only. Normalising to the crystal volume (V) the matrix element can be written as

$$\langle f|e^{i\mathbf{q}\cdot\mathbf{r}}\hat{\mathbf{e}}\cdot\mathbf{p}|i\rangle = \frac{1}{V}\int_V u_{f,k_f}(\mathbf{r})^* e^{-i(\mathbf{k_f}\cdot\mathbf{r})} e^{i(\mathbf{q}\cdot\mathbf{r})}\hat{\mathbf{e}}\cdot\mathbf{p}u_{i,k_i}(\mathbf{r})e^{i(\mathbf{k_i}\cdot\mathbf{r})}\,d^3\mathbf{r}. \qquad \text{(D.8)}$$

Now since $\mathbf{p} = -i\hbar\nabla$ we see $\mathbf{p}u_{i,k_i}(\mathbf{r})e^{i(\mathbf{k_i}\cdot\mathbf{r})} = [\mathbf{p}u_{i,k_i}(\mathbf{r})]\,e^{i(\mathbf{k_i}\cdot\mathbf{r})} + u_{i,k_i}(\mathbf{r})\left[\hbar\mathbf{k_i}e^{i(\mathbf{k_i}\cdot\mathbf{r})}\right]$ and using the translational periodicity of the Bloch function (allowing us to rewrite $\mathbf{r}$ as $\mathbf{R}+\mathbf{r}$, where $\mathbf{R}$ is a primitive lattice vector and $\mathbf{r}$ is a unit cell vector), equation D.8 can be rewritten as an integration over the unit cell of volume Ω given by

$$\langle f|e^{i\mathbf{q}\cdot(\mathbf{r}+\mathbf{R})}\hat{\mathbf{e}}\cdot\mathbf{p}|i\rangle = \frac{1}{V}\sum_{\mathbf{R}} e^{i(\mathbf{k_i}-\mathbf{k_f}+\mathbf{q})\cdot\mathbf{R}}\int_\Omega e^{i(\mathbf{k_i}-\mathbf{k_f}+\mathbf{q})\cdot\mathbf{r}}u_{f,k_f}(\mathbf{r})^*\hat{\mathbf{e}}\cdot(\mathbf{p}+\hbar\mathbf{k_i})u_{i,k_i}(\mathbf{r})\,d^3\mathbf{r}.$$

$$\text{(D.9)}$$

The summation with respect to $\mathbf{R}$ is zero for all except when $(\mathbf{k_i}-\mathbf{k_f}+\mathbf{q})$ is equal to the reciprocal lattice vector $m\mathbf{G}$ (m is an integer), as $\mathbf{R}$ will cancel with the corresponding $-\mathbf{R}^\dagger$. If we assume umklapp processes will not occur as both electron wave vectors lie within the first Brillouin zone then we see that at $m\mathbf{G} = 0$ the integration in equation D.9 goes to unity. This means that if the photon wavevector $\mathbf{q}$ is much smaller than that of the electron wave vectors $\mathbf{k_i}$ and $\mathbf{k_f}$, then we can ignore the photon wavevector part and see $\mathbf{k_i}$ - $\mathbf{k_f}$ = 0 so that $\mathbf{k_i}$ = $\mathbf{k_f}$, i.e. we require a direct transition to conserve momentum. Electrons require the same $\mathbf{k}$ to perform an optical transition.

Using the above assumptions and due to the orthogonality of the Bloch wave functions $u_{f,k_f}(\mathbf{r})^* u_{i,k_i}(\mathbf{r})$, the integral with respect to $\hbar\mathbf{k}$ disappears. So equation D.9 can be massively simplified to give

---

†The summation $\sum_{\mathbf{R}} e^{i(\mathbf{k_i}-\mathbf{k_f}+\mathbf{q})\cdot\mathbf{R}}$ can be treated as a delta function $\delta(\mathbf{k_i} - \mathbf{k_f} + \mathbf{q})$ ensuring conservation of momentum.

$$\langle f|e^{i\mathbf{q}\cdot(\mathbf{r}+\mathbf{R})}\hat{\mathbf{e}}\cdot\mathbf{p}|i\rangle \approx \frac{1}{\Omega}\int_{\Omega}u_{f,k}(\mathbf{r})^{*}\hat{\mathbf{e}}\cdot\mathbf{p}u_{i,k}(\mathbf{r})d^{3}\mathbf{r} = \hat{\mathbf{e}}\cdot\mathbf{p}_{\mathrm{fi}} \tag{D.10}$$

In equation D.8 we have relabelled k_i and k_f in the subscripts for simply k in the expressions for the Bloch waves $u(\mathbf{r})$ as taking the Taylor series of $u_{f,k_f} = u_{f,k_i+q} = u_{f,k_i} + \mathbf{q}\cdot\nabla_k u_{f,k_i} + ...$ and we assume $\mathbf{q}$ is small. The crystal volume is also replaced by the unit cell volume $\Omega = \frac{V}{N}$ due to the sum over N unit cells.

Finally, inserting equation D.10 into D.7 gives the transition rate per unit volume

$$
\begin{aligned}
R &= \frac{\pi e^2}{2m^2\hbar}A_0^2\sum_{\mathbf{k}}|\hat{\mathbf{e}}\cdot\mathbf{p}_{\mathrm{fi}}|^2\delta(E_f(\mathbf{k})-E_i(\mathbf{k})-\hbar\omega) \\
&= \frac{\pi e^2}{2m^2\hbar}\frac{E_0^2}{\omega^2}\sum_{\mathbf{k}}|\hat{\mathbf{e}}\cdot\mathbf{p}_{\mathrm{fi}}|^2\delta(E_f(\mathbf{k})-E_i(\mathbf{k})-\hbar\omega)
\end{aligned}
\tag{D.11}
$$

where the Maxwell's relationships $\nabla\times\mathbf{E} = -\frac{\partial\mathbf{B}}{\partial t}$ and $\mathbf{B} = \nabla\times\mathbf{A}$ were used to give $\mathbf{E} = -\frac{\partial\mathbf{A}}{\partial t}$ and so $E_0 = \omega A_0$.

THE ABSORPTION COEFFICIENT

The transition rate is intimately connected to the absorption coefficient (the fraction of power of incident light absorbed in a unit length of a medium) in the Beer-Lambert law, which describes the attenuation of light intensity as a function of propagation length through a medium $I = I_0 e^{-\alpha x}$. Rearranging this formula gives

$$\alpha(\hbar\omega) \approx \frac{-\frac{\delta I}{I}}{\delta x} \tag{D.12}$$

where $-\frac{\delta I}{I}$ is the fraction of absorbed light intensity at a position δx. If the total incident intensity is given by the time averaged Poynting vector, which for an electromagnetic field described as a plane wave can be shown to be equal to $\langle\mathbf{S}\rangle = \frac{E_0^2}{2\mu_0\mu_r c}n$ [129,381]

(n is the dielectric constant) and the power loss is given as $\hbar\omega \cdot R$, then the absorption coefficient can be given by combining equations D.11 and D.12 to give

$$\alpha(\hbar\omega) = \frac{\hbar\omega \cdot R(\hbar\omega)}{\langle S(\hbar\omega) \rangle} = \frac{\pi\mu_r e^2}{\omega\varepsilon_0 m^2 cn} \sum_{\mathbf{k}} |\hat{\mathbf{e}} \cdot \mathbf{p_{fi}}|^2 \delta(E_f(\mathbf{k}) - E_i(\mathbf{k}) - \hbar\omega) \qquad \text{(D.13)}$$

If we assume that the matrix element $\hat{\mathbf{e}} \cdot \mathbf{p_{fi}}$ varies slowly, i.e. is not strongly $\mathbf{k}$ dependent then this term may be moved outside of the summation in equation D.13

$$\alpha(\hbar\omega) = \frac{\pi\mu_r e^2}{\omega\varepsilon_0 m^2 cn} |\hat{\mathbf{e}} \cdot \mathbf{p_{fi}}|^2 \sum_{\mathbf{k}} \delta(E_{fi} - \hbar\omega) \qquad \text{(D.14)}$$

where $E_{fi} = E_f(\mathbf{k}) - E_i(\mathbf{k})$. Because the summation in equation D.14 is over all available paired states $\mathbf{k}$, we can write this as an integration over all available states (over the joint density of states)

$$\alpha(\hbar\omega) = \frac{\pi\mu_r e^2}{\omega\varepsilon_0 m^2 cn} |\hat{\mathbf{e}} \cdot \mathbf{p_{fi}}|^2 \int g(E_{fi})\delta(E_{fi} - \hbar\omega) d^3 E_{fi}. \qquad \text{(D.15)}$$

D.0.2 Joint Density of States

Equation D.15 gives a convenient model for the absorption coefficient of a semiconductor. It does however, require knowledge of form of the joint density of states g_{fi}. Therefore, we introduce the concept of the joint density of states, and give the simplest example of the form of equation D.15 for a direct transition for a material with parabolic and spherically symmetric valence and conduction bands.

In the same way the density of states $g(E)$ counts the number of states in a given energy interval (most often per unit volume) for a single band, the joint density of states takes into account that both the initial and final electron states lie within continuous bands, with energy seperation $E_{fi} = E_f - E_i$. For electrons in a band the density of

states per within a given energy range $E + dE$ is given as

$$g(E)dE = 2g(k)dk \tag{D.16}$$

where the factor 2 accounts for the two spin states of the electron for each allowed k state. It follows that

$$g(E) = \frac{2g(k)}{\frac{dE}{dk}}. \tag{D.17}$$

$g(k)$ can be determined as the number of k-states in the incremental volume element between two constant energy surfaces separated by dE, see figure D.1 for a schematic guide.
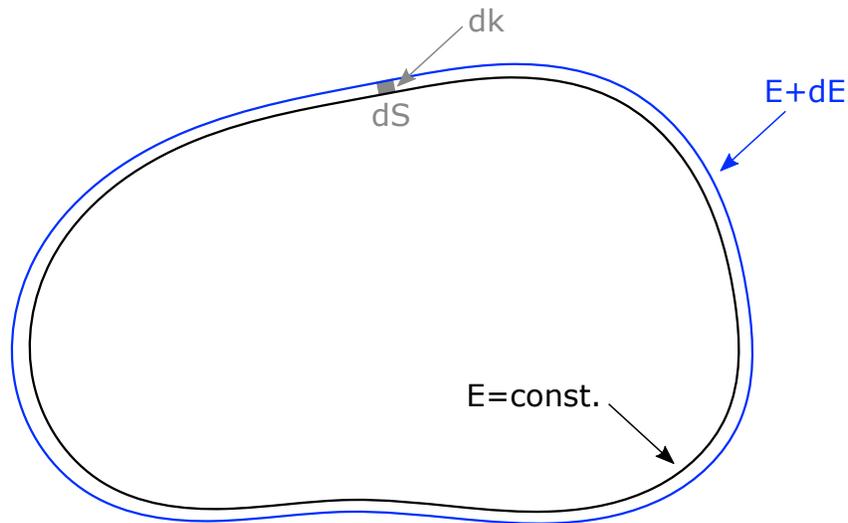
Figure D.1: Schematic derivation of the joint density of states between two constant energy surfaces E and $E + dE$. The volume element in k-space is defined by the product of the surface area dS and the distance dk perpendicular to the constant energy surface.

The number of states in this region per unit volume is given as

$$dN = \frac{dV_k}{(2\pi)^3}. \tag{D.18}$$

In three dimensions we write $\frac{dE}{d\mathbf{k}} = |\nabla_{\mathbf{k}} E|$, and the definition of the unit volume element pictured in figure D.1 is given as the infinitesimal surface area element (of a surface with constant energy) dS in k-space, multiplied by the perpendicular wavevector $\mathbf{k}$; $dV = dS \cdot d\mathbf{k}$. Combining all of the above and integrating over the surface gives the density of states (including the two spin sates)

$$g(E_{fi}) = \frac{2}{(2\pi)^3} \int_S \frac{dS}{|\nabla_{\mathbf{k}} E_{fi}|}. \tag{D.19}$$

As indicated by the no uniform shape of the energy surface in figure D.1, this expression is valid for any band (we liberally use E and E_{fi} interchangeably here). Singularities exist for equation D.19, when $|\nabla_k E_{fi}| = 0$, i.e. the bands are parallel. This leads to a maximum probability for optical transitions. These points can exist in multiple places in the Brillouin zone, and were originally classified by van Hove[382], so they are known as Van Hove singularities. Expanding equation D.19 as a Taylor series around $\mathbf{k} = \mathbf{k}_0$ gives

$$E_{fi}(\mathbf{k}) = E_{fi}(\mathbf{k}_0) + (\mathbf{k} - \mathbf{k}_0) \cdot \underbrace{\nabla_{\mathbf{k}} E_{fi}(\mathbf{k}_0)}_{=0} + \frac{1}{2}\left((\mathbf{k} - \mathbf{k}_0) \cdot \nabla_{\mathbf{k}}\right)^2 E_{fi}(\mathbf{k}_0) + \dots \tag{D.20}$$

It is common to ignore the higher order terms and note the first-order term is zero due to the condition $\nabla_{\mathbf{k}} E_{fi} = 0$.

Van Hove wrote an equation similar to

$$
\begin{aligned}
E_{fi}(\mathbf{k}) &= E_{fi}(\mathbf{k}_0) + \sum_{i=1}^{3} a_i k_i^2 + \dots \\
&= E_{fi}(\mathbf{k}_0) + \sum_{i=1}^{3} \frac{\hbar^2 k_i^2}{2\mu_i} + \dots
\end{aligned}
\tag{D.21}
$$

where $\frac{1}{\mu} = \frac{1}{m_e^*} + \frac{1}{m_h^*}$ is the reduced mass. Van Hove singularities are classified by the number of negative coefficients a_i which is essentially the sign of the reduced mass, and in three dimensional space there are four kinds of Van Hove singularity. The two extremes when all a_i are either -1 or 1 correspond to band extrema (minimum or maximum), while when one of the coefficients is opposite sign to the others, it represents one of two saddle points. These are conventionally labelled M_0 and M_3 for the two extrema, and M_1 and M_2 for the saddle points. The joint density of states for a a simple cubic lattice is shown in figure D.2 to illustrate this.
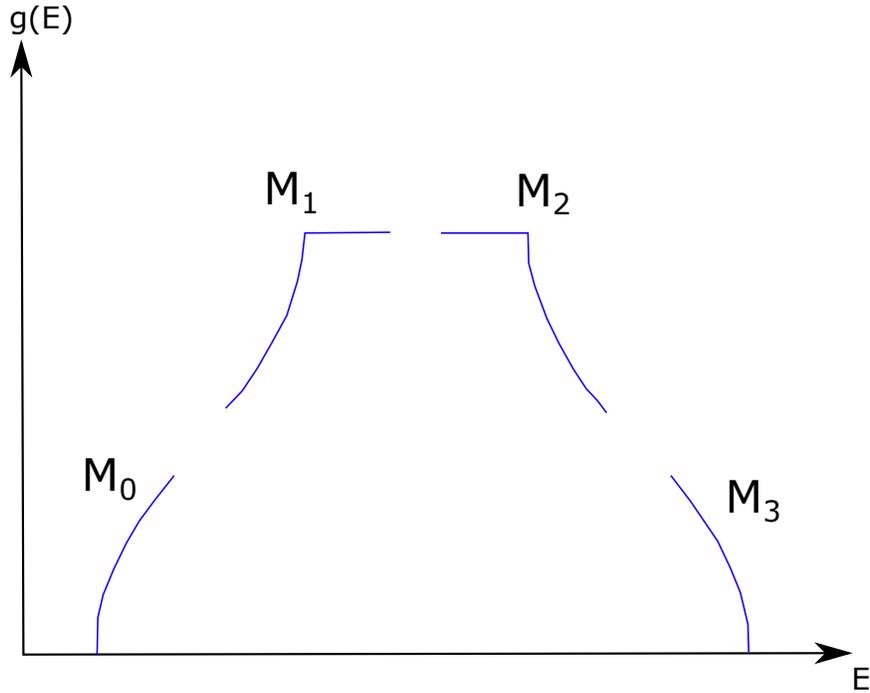
Figure D.2: Van Hove points for a cubic lattice with $E_{fi}(\mathbf{k}) = (E_{fi}(\mathbf{k_0})+3\gamma) - \gamma(cos(k_x a) + cos(k_y a) + cos(k_z a))$ taken from ref. 168.

To find an analytical expression for the joint density of states we will consider the M_0 (all $a_i > 0$) critical point with parabolic and spherically symmetric valence and conduction bands. The effective masses of such bands are given as m_v^* and m_c^*. The joint density of states transition energy from equation D.20 is therefore

$$E_{fi} = (E_g + E_f) - E_i == E_g + \frac{(\hbar k)^2}{2m_e^*} + \frac{(\hbar k)^2}{2m_h^*} = E_g + \frac{(\hbar k)^2}{2\mu}. \tag{D.22}$$

Inserting equation D.22 into D.19, given the area element S in spherical coordinates is given by $k^2 \sin\theta d\theta d\varphi$ and $\nabla_k E_{fi} = \frac{\hbar^2 k}{\mu}$ we can write the joint density of states as

$$g(E_{fi}) = \frac{2}{(2\pi)^3} \int_S \frac{dS}{\frac{\hbar^2 k}{\mu}} = \frac{k\mu}{4\pi^2\hbar^2} \int_0^\pi \sin\theta d\theta \int_0^{2\pi} d\varphi = \frac{\sqrt{2\mu^3}}{\pi^2\hbar^3} \sqrt{E_{fi} - E_g} \tag{D.23}$$

where we have rearranged equation D.22 to give $k^2 = \frac{2\mu(E - E_g)}{\hbar^2}$. Inserting this into equation D.15 we see

$$\alpha(\hbar\omega) = \frac{\mu_r e^2}{\pi\hbar^3\omega\varepsilon_0 m^2 cn} |\hat{\mathbf{e}} \cdot \mathbf{p_{fi}}|^2 \sqrt{2\mu^3(\hbar\omega - E_g)} \propto \sqrt{\hbar\omega - E_g}. \tag{D.24}$$

Equation D.24 tells us that for a direct transition between bands with parabolic nature $\alpha^2 \propto \hbar\omega - E_g$ which is why you will often see linear extrapolations used to predict the band gap, the so called Tauc analysis.

Of course many optical processes can occur that are not described by this simple model. Without going into too much details, another very important optical transition is that of an indirect band gap. In these transitions a photon is required to conserve momentum meaning the delta function from equation D.13 becomes $\delta(E_f(\mathbf{k}) - E_i(\mathbf{k}) \pm \hbar\mathbf{q} - \hbar\omega)$ where $\pm$ allows for absorption or emission of a photon with wavevector $\mathbf{q}$, so $\mathbf{k_f} = \mathbf{k_i} \mp \mathbf{q}$. The occupation and therefore absorption coefficient of phonons has to be considered using Bose-Einstein statistics. It can be shown in a similar (but more taxing) manner to what we have shown for a direct transition that the transition rate and absorption coefficient can be given as[26,168]

$$W_{if} = \frac{2\pi}{\hbar} \left| \sum_m \frac{\langle f|H|m\rangle \langle m|H|i\rangle}{E_i - E_m} \right|^2 \delta(E_i - E_f) \tag{D.25}$$

$$\alpha_i(\hbar\omega) \propto (\hbar\omega \mp \hbar\mathbf{q} - E_{ig})^2 \tag{D.26}$$

where the transition take place from an initial state i to a final state f through a virtual state m, and α is calculated in the vicinity if the indirect gap.

E

Electronic Transport

This section lays the ground work for what follows in section 2.3.2. Whilst the natural order would be to read this section first, like much of this thesis the content proved too in-depth a topic to fit concisely in a few pages in the main sections. Hence, I decided to separate it in the hopes of doing the subject justice, whilst acknowledging it is not necessary to fully understand the content in order to follow the procedures carried out in the main thesis. I aim on providing the tools to find analytical solutions to the Boltzmann equation, allow allow us to give one such solution as an example.

E.0.1 THE BOLTZMANN TRANSPORT EQUATION

Most often the important transport coefficients are calculated by solving the Boltzmann transport equation. To discuss this we need to introduce the distribution function $f(\mathbf{r}, \mathbf{k}, t)$ which gives the probability of occupation of a state by an electron at

position $\mathbf{r}$ and wave vector $\mathbf{k}$ at time t. If for the moment we assume no scattering events take place then, after a time interval dt an electron is changed into a new state with position $\mathbf{r} + \dot{\mathbf{r}}dt$ and wave vector $\mathbf{k} + \dot{\mathbf{k}}dt$. The change in the distribution function in the time interval dt can be written as

$$\left(\frac{df}{dt}\right)_{drift} = \frac{f(\mathbf{k} - \dot{\mathbf{k}}dt, \mathbf{r} - \dot{\mathbf{r}}dt, t - dt) - f(\mathbf{r}, \mathbf{k}, t)}{dt} \tag{E.1}$$

where the first term on the right hand side is essentially the distribution function for the initial electron state which transitions to the second term on the right hand side. Equation E.1 represents that for continuous flow of electrons since no scattering take place, hence this is called the drift term. By expansing $f(\mathbf{k} - \dot{\mathbf{k}}dt, \mathbf{r} - \dot{\mathbf{r}}dt, t - dt))$ in a Taylor series and keeping only the terms to first-order, it can be shown that equation E.1 becomes [155,168]

$$\left(\frac{df}{dt}\right)_{drift} = -\left[\dot{\mathbf{k}} \cdot \nabla_{\mathbf{k}}f + \mathbf{v} \cdot \nabla_{\mathbf{r}}f + \frac{\partial f}{\partial t}\right]. \tag{E.2}$$

Where we have used $\dot{\mathbf{r}} = \mathbf{v}$. Now using the fact that $\dot{\mathbf{k}} = \frac{\mathbf{F}}{\hbar}$ we can write

$$\left(\frac{df}{dt}\right)_{drift} = -\left[\frac{1}{\hbar}(\mathbf{F} \cdot \nabla_{\mathbf{k}}f) + \mathbf{v} \cdot \nabla_{\mathbf{r}}f + \frac{\partial f}{\partial t}\right]. \tag{E.3}$$

Now we can introduce scattering (collision) which changes an electrons state. We can define the rate of vhange in the distribution function due to collision as $\left(\frac{df}{dt}\right)_{collision}$ and since the distribution function has to satisfy the equilibrium (steady-state) condition

$$\left(\frac{df}{dt}\right)_{drift} + \left(\frac{df}{dt}\right)_{collision} = 0 \tag{E.4}$$

giving us

$$\frac{\partial f}{\partial t} + \frac{1}{\hbar}\mathbf{F} \cdot \nabla_{\mathbf{k}}f + \mathbf{v} \cdot \nabla_{\mathbf{r}}f = \left(\frac{df}{dt}\right)_{collision}. \tag{E.5}$$

Equation E.5 is probably the most common form of the Boltzmann equation.

E.0.2 Scattering Theory

The scattering rate (probability of transition per unit time) of an electron transitioning from state $\mathbf{k}$ to $\mathbf{k}'$ due to some perturbation (ΔV) is described by the familiar Fermi's golden rule

$$W_{\mathbf{k},\mathbf{k}'} = \frac{2\pi}{\hbar}|M(\mathbf{k},\mathbf{k}')|^2 \delta(E_{\mathbf{k}} - E_{\mathbf{k}'}) \tag{E.6}$$

where

$$M(\mathbf{k},\mathbf{k}') = \int_\Omega \phi^*_{\mathbf{k}'}(\mathbf{r})\Delta V \phi_{\mathbf{k}}(\mathbf{r})d\mathbf{r}. \tag{E.7}$$

Defect scattering events tend to affect the wave function of the electron being scattered. Hence, we can only consider the bloch wavefunction of the electron

$$\varphi_{\mathbf{k}}(\mathbf{r}) = U_{\mathbf{k}}(\mathbf{r})e^{i\mathbf{k}\cdot\mathbf{r}}. \tag{E.8}$$

The perturbing potential can be written as a Fourier series

$$\Delta V = \sum_{\mathbf{q}} A(\mathbf{q})e^{i\mathbf{q}\cdot\mathbf{r}}. \tag{E.9}$$

Therefore, we can write equation E.7 as

$$M(\mathbf{k}, \mathbf{k}') = \sum_{\mathbf{q}} \int_{\Omega} U_{\mathbf{k}'}^*(\mathbf{r}) e^{-i\mathbf{k}'\cdot\mathbf{r}} A(\mathbf{q}) e^{i\mathbf{q}\cdot\mathbf{r}} U_{\mathbf{k}}(\mathbf{r}) e^{i\mathbf{k}\cdot\mathbf{r}} d\mathbf{r}$$

$$= A(\mathbf{k}' - \mathbf{k}) \int_{\Omega} U_{\mathbf{k}'}^*(\mathbf{r}) U_{\mathbf{k}}(\mathbf{r}) d\mathbf{r} \tag{E.10}$$

$$= A(\mathbf{k}' - \mathbf{k}) I(\mathbf{k}, \mathbf{k}')$$

where Ω is the unit cell volume, the last two lines were allowed by setting $\mathbf{q} = \mathbf{k}' - \mathbf{k}$ (conservation of momentum), and $I(\mathbf{k}, \mathbf{k}')$ is known as the overlap integral. Scattering with lattice atoms is more complicated as changes in the lattice need to be accounted for. For the sake of brevity, this will not be discussed here, but can be seen in text books on the subject, such as Nag[155] and Look[160].

E.0.3 THE COLLISION TERM

The total transition probability of an electron being scattered from wave vector $\mathbf{k}$ to $\mathbf{k}'$ can be calculated as a sum over all states that an electron can scatter from $\mathbf{k}$ to

$$P(\mathbf{k}, \mathbf{k}') = \frac{V_c}{(2\pi)^3} \int \frac{2\pi}{\hbar} |M(\mathbf{k}, \mathbf{k}')|^2 \delta(E_{\mathbf{k}} - E_{\mathbf{k}'}) d\mathbf{k}'$$

$$= \frac{V_c}{(2\pi)^3} \int W_{\mathbf{k}, \mathbf{k}'} d\mathbf{k}' \tag{E.11}$$

where V_c is the crystal volume, and the factor $\frac{1}{(2\pi)^3}$ accounts for the density of states in $\mathbf{k}'$-space. Note that the spin is not (necessarily) altered in the scattering process so there is no factor of 2 required here. If we now want to know the number of scattered electrons per unit time from $\mathbf{k}$ to $\mathbf{k}'$, i.e. the scattering out rate we need to multiply equation E.11 by the number of states both occupied and available giving

$$\left(\frac{df}{dt}\right)_{out} = \underbrace{\left[\frac{2V_c}{(2\pi)^3}f(\mathbf{r},\mathbf{k},t)\,d\mathbf{k}\right]}_{scattered\ electrons}\underbrace{\left[\frac{V_c}{(2\pi)^3}\{1-f(\mathbf{r},\mathbf{k}',t)\}\,d\mathbf{k}'\right]}_{available\ states}P(\mathbf{k},\mathbf{k}') \qquad \text{(E.12)}$$

Notice the factor of 2 in the term describing electrons to be scattered. This is because either spin state can be scattered, but the available state will be filled by an electron with matching spin. Likewise, electrons can be scattered into the state $\mathbf{k}$, with a scattering in rate of

$$\left(\frac{df}{dt}\right)_{in} = \left[\frac{2V_c}{(2\pi)^3}f(\mathbf{r},\mathbf{k}',t)\,d\mathbf{k}\right]\left[\frac{V_c}{(2\pi)^3}\{1-f(\mathbf{r},\mathbf{k},t)\}\,d\mathbf{k}'\right]P(\mathbf{k}',\mathbf{k}) \qquad \text{(E.13)}$$

The net scattering rate into $d\mathbf{k}$ is given by the sum of equations E.12 and E.13

$$-\frac{V_c}{(2\pi)^3}\frac{2V_c}{(2\pi)^3}\int_{\mathbf{k}'}\left[f(\mathbf{k})(1-f(\mathbf{k}'))P(\mathbf{k},\mathbf{k}')-f(\mathbf{k}')(1-f(\mathbf{k}))P(\mathbf{k}',\mathbf{k})\right]\,d\mathbf{k}\,d\mathbf{k}' \qquad \text{(E.14)}$$

where we have written $f(\mathbf{r},\mathbf{k},t) = f(\mathbf{k})$ and $f(\mathbf{r},\mathbf{k}',t) = f(\mathbf{k}')$ as short-hand. The integration is carrier out over all $\mathbf{k}$ and $\mathbf{k}'$ to account for any transition to and from *bfk*.

Equation E.14 is related to the collision term as $-\frac{2V_c}{2\pi^3}\left(\frac{df}{dt}\right)_{collision}d\mathbf{k}$ which describes the change in the nuber of electrons per unit volume with wave vectors in the volume element $d\mathbf{k}$ around $\mathbf{k}$. Hence we arrive at

$$\left(\frac{df}{dt}\right)_{collision} = -\frac{V_c}{(2\pi)^3}\int_{\mathbf{k}'}\left[f(\mathbf{k})(1-f(\mathbf{k}'))P(\mathbf{k},\mathbf{k}')-f(\mathbf{k}')(1-f(\mathbf{k}))P(\mathbf{k}',\mathbf{k})\right]\,d\mathbf{k}'.$$
$$\text{(E.15)}$$

Equation E.15 is often called the collision integral

E.0.4 RELATING SCATTERING AND RELAXATION TIMES

We will now try an find a solution to the Boltzmann equation which we write here for convenience, as well as to reaffirm its importance (we substitute our expression in equation E.15 into E.5)

$$\frac{\partial f}{\partial t} + \frac{1}{\hbar}\mathbf{F} \cdot \nabla_{\mathbf{k}}f + \mathbf{v} \cdot \nabla_{\mathbf{r}}f =$$
$$-\frac{V_c}{(2\pi)^3}\int_{\mathbf{k}'}[f(\mathbf{k})(1-f(\mathbf{k}'))P(\mathbf{k},\mathbf{k}') - f(\mathbf{k}')(1-f(\mathbf{k}))P(\mathbf{k}',\mathbf{k})]d\mathbf{k}'. \tag{E.16}$$

Because equation E.15 is a non-linear integral, equation E.16 becomes an integro-differential equation and so impossible to analytically solve. Hence, simplifications are required to provide an approximate solution to this. This is done by linearising the Boltzmann equation.

To begin, in the absence of any external field, and at thermal equilibrium the distribution function of the electrons is a Fermi-Dirac distribution

$$f_0(\mathbf{k}) = \frac{1}{e^{\frac{E_{\mathbf{k}}-\mu_F}{K_BT}} + 1}. \tag{E.17}$$

In the absence of accelerating and diffusive forces the Boltzmann equation reduces to the steady-state form $\left(\frac{df}{dt}\right)_{collision} = 0$, meaning the Fermi function must make the collision integral equal to zero (as many electrons scattered in and out of $\mathbf{k}$ is balanced)

$$f_0(\mathbf{k})(1-f_0(\mathbf{k}'))P(\mathbf{k},\mathbf{k}') - f_0(\mathbf{k}')(1-f_0(\mathbf{k}))P(\mathbf{k}',\mathbf{k}) = 0 \tag{E.18}$$

Now, we switch on the electric field which forces a change in the distribution function. The electric field accelerates the electrons (and their distribution function) oppo-

site to the direction of the field. The distribution function can be given as a Legendre polynomial series, and if the external force is weak, we can approximate the function to the first two terms

$$f(\mathbf{k}) = \sum_n f_n(\mathbf{k}) P_n(cos\theta) \approx f_0(\mathbf{k}) + f_1(\mathbf{k}) cos\theta \tag{E.19}$$

where $P_n(cos\theta)$ is the Legendre polynomial of order n and θ is the angle between the electron wave vector $\mathbf{k}$ and the applied electric field $\mathbf{E}$. We see the distribution function is made up of the equilibrium distribution function and a small perturbation.

Substituting equation E.19 into E.5 and assuming terms involving $f_1(\mathbf{k})f_1(\mathbf{k}')$ are negligible, we see

$$\begin{aligned}
\left(\frac{df}{dt}\right)_{collision} = & -\frac{V_c}{(2\pi)^3} \int_{\mathbf{k}'} \Big[f_0(\mathbf{k})(1 - f_0(\mathbf{k}'))P(\mathbf{k}, \mathbf{k}') - f_0(\mathbf{k}')(1 - f_0(\mathbf{k}))P(\mathbf{k}', \mathbf{k}) \Big] \\
& + \Big[f_1(\mathbf{k}) cos\theta \left(P(\mathbf{k}, \mathbf{k}')\{1 - f_0(\mathbf{k}')\} + f_0(\mathbf{k}')P(\mathbf{k}', \mathbf{k}) \right) \\
& - f_1(\mathbf{k}') cos\theta' \left(P(\mathbf{k}', \mathbf{k})\{1 - f_0(\mathbf{k})\} + f_0(\mathbf{k})P(\mathbf{k}, \mathbf{k}') \right) \Big] d\mathbf{k}'.
\end{aligned} \tag{E.20}$$

We can split this equation into two parts relating back to equation E.16. If we assume that the semiconductor sample is homogeneous $\nabla_\mathbf{r} f = 0$, and that the accelerating force $\mathbf{F}$ experienced by the carriers is due to an electric field only $\mathbf{B} = 0$, under steady state conditions $\frac{\partial f}{\partial t} = 0$, the equation E.16 becomes

$$\frac{e}{\hbar} \mathbf{E} \cdot \nabla_\mathbf{k} f = \left(\frac{df}{dt}\right)_{collision}. \tag{E.21}$$

Because at $E = 0$ we know $f_1 = 0$, i.e. we retain our equilibrium distribution f_0, we can expect $f_1 \propto \mathbf{E}$. Hence, $f_1 \mathbf{E} \propto \mathbf{E}^2$ and so is negligible at low $\mathbf{E}$, we relate the

Δf_1 term with the first integration in equation E.20 which is equal to zero[160,383] in the limit of low field as this is just the equation when $E = 0$ where we know $\left(\frac{df}{dt}\right) = 0$ (steady-state). Therefore, under these assumptions we see

$$
\frac{e}{\hbar} \mathbf{E} \cdot \nabla_{\mathbf{k}} f_0 = -\frac{V_c}{(2\pi)^3} \int_{\mathbf{k}'} \Big[f_1(\mathbf{k}) cos\theta \left(P(\mathbf{k}, \mathbf{k}')\{1 - f_0(\mathbf{k}')\} + f_0(\mathbf{k}') P(\mathbf{k}', \mathbf{k}) \right)
$$
$$
- f_1(\mathbf{k}') cos\theta' \left(P(\mathbf{k}', \mathbf{k})\{1 - f_0(\mathbf{k})\} + f_0(\mathbf{k}) P(\mathbf{k}, \mathbf{k}') \right) \Big] d\mathbf{k}'. \tag{E.22}
$$

Now if we decide that $P(\mathbf{k}, \mathbf{k}') = P(\mathbf{k}', \mathbf{k})$ which is essentially to say $P(\mathbf{k}, \mathbf{k}') = P(|\mathbf{k}', \mathbf{k}|)$ i.e. the magnitude of $|\mathbf{k}', \mathbf{k}|$ is the important parameter to P, not the orientation of $\mathbf{k}$ and $\mathbf{k}'$, then using this and equation E.18 above, we can make the simplification

$$
\left(\frac{df}{dt}\right)_{collision} = -\frac{V_c}{(2\pi)^3} cos\theta \int_{\mathbf{k}} P(\mathbf{k}, \mathbf{k}') \left[f_1(\mathbf{k}) \frac{1 - f_0(\mathbf{k}')}{1 - f_0(\mathbf{k})} - \frac{cos\theta'}{cos\theta} f_1(\mathbf{k}') \frac{f_0(\mathbf{k})}{f_0(\mathbf{k}')} \right] d\mathbf{k}'. \tag{E.23}
$$

Now we can simplify further using the relation $cos\theta' = cos\theta cos\theta_{\mathbf{k}} - sin\theta sin\theta_{\mathbf{k}} cos\beta$ where $\theta_{\mathbf{k}}$ is the angle between $\mathbf{k}$ and $\mathbf{k}'$, β is the angle between the directions of $\mathbf{k} \rightarrow \mathbf{k}'$ and $\mathbf{k} \rightarrow \mathbf{E}$. If the integration over β between 0 and 2π forces those terms to zero, we are left with

$$
\left(\frac{df}{dt}\right)_{collision} = -\frac{V_c}{(2\pi)^3} cos\theta \int_{\mathbf{k}} P(\mathbf{k}, \mathbf{k}') \left[f_1(\mathbf{k}) \frac{1 - f_0(\mathbf{k}')}{1 - f_0(\mathbf{k})} - cos\theta_k f_1(\mathbf{k}') \frac{f_0(\mathbf{k})}{f_0(\mathbf{k}')} \right] d\mathbf{k}'. \tag{E.24}
$$

Finally, we can use what is known as the relaxation-time approximation

$$
\left(\frac{df}{dt}\right)_{collision} \approx \frac{f(\mathbf{k}) - f_0(\mathbf{k})}{\tau} = \frac{f_1(\mathbf{k}) cos\theta}{\tau} \tag{E.25}
$$

250

giving

$$\frac{1}{\tau} = \frac{V_c}{(2\pi)^3} \int_{\mathbf{k}} P(\mathbf{k}, \mathbf{k}') \left[\frac{1 - f_0(\mathbf{k}')}{1 - f_0(\mathbf{k})} - cos\theta_k \frac{f_1(\mathbf{k}')}{f_1(\mathbf{k})} \frac{f_0(\mathbf{k})}{f_0(\mathbf{k}')} \right] d\mathbf{k}'. \qquad \text{(E.26)}$$

In the case of elastic scattering mechanisms such as ionized impurity scattering, $E_{\mathbf{k}} \approx E_{\mathbf{k}'}$, so $f_0(\mathbf{k}) \approx f_0(\mathbf{k}')$ and $E_{\mathbf{k}} = E_{\mathbf{k}'}$, so $f_1(\mathbf{k}) \approx f_1(\mathbf{k}')$ allowing equatoin E.26 to be reduced to

$$\frac{1}{\tau} = \frac{V_c}{(2\pi)^3} \int_{\mathbf{k}} P(\mathbf{k}, \mathbf{k}') \left(1 - cos\theta_k\right) d\mathbf{k}'. \qquad \text{(E.27)}$$

Substituting equation E.11 into E.27 gives the scattering time for a an electron collision with a defect

$$\frac{1}{\tau} = \frac{V_c}{(2\pi)^3} \int_{\mathbf{k}} \frac{2\pi}{\hbar} |M(\mathbf{k}, \mathbf{k}')|^2 \left(1 - cos\theta_k\right) \delta(E_{\mathbf{k}} - E_{\mathbf{k}'}) d\mathbf{k}'. \qquad \text{(E.28)}$$

Note that for lattice scattering this equation is complicated by the different form of E.11. Equation E.28 can be solved for any defect scattering mechanism provided that $|M(\mathbf{k}, \mathbf{k}')|^2$ is known.

F

The Charge Neutrality Level

This section is concerned with semiconductor surface space charge, and much of this is based on the work of Mönch[81]. This is primarily of interest for section 6, where the surface properties of Ga_2O_3 are investigated in the context of the CNL[384]. It was already mentioned that wave functions related to surface states are evanescent, exponentially decay into the vacuum and into the solid, and so exhibit real energies but complex wave vectors. These surface states physically describe electronic levels that are localized near the surface. We can obtain information regarding these states by considering the complex band structure of a one dimensional lattice of nearly free electrons[81], see figure F.1. An electron of mass m_0 in a constant potential V_0 must satisfy the one-electron Schrödinger equation

$$H_0|\varphi_{\mathbf{k}}(\mathbf{r})\rangle = E_{\mathbf{k}}^0|\varphi_{\mathbf{k}}(\mathbf{r})\rangle \tag{F.1}$$
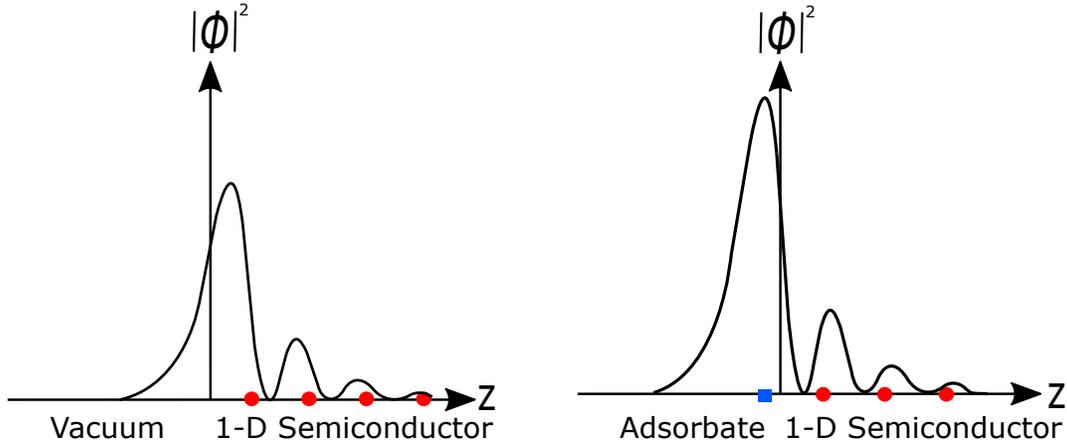
Figure F.1: Schematic of the wave function of surface states for a clean semiconductor, and one with an adborbate on the surface, both described by the one-dimensional lattice nearly free electron gas model. Figure reproduced from Mönch [81].

where $H_0 = \frac{\mathbf{p}^2}{2m_0} + V_0$ as in equation 3.1 and $E_\mathbf{k}^0 = \frac{\hbar^2 \mathbf{k}^2}{2m_0} + V_0$. A small periodic potential is introduced as a perturbation

$$H = H_0 + H_1 = H_0 + \sum_\mathbf{G} V_1 e^{-i\mathbf{G}\cdot\mathbf{r}} \tag{F.2}$$

where $\mathbf{G}$ is the reciprocal lattice vector (in the simplest case such as the linear one-dimensional lattice $\mathbf{G} = \frac{2\pi}{a}$ where a is the lattice parameter). Writing the Hamiltonian in matrix form gives [385]

$$\langle \varphi_{\mathbf{k}'} | H | \varphi_\mathbf{k} \rangle = \langle \varphi_{\mathbf{k}'} | H_0 + H_1 | \varphi_\mathbf{k} \rangle = E_{\mathbf{k}'}^0 \delta_{\mathbf{k}\mathbf{k}'} + \langle \varphi_{\mathbf{k}'} | H_1 | \varphi_\mathbf{k} \rangle \tag{F.3}$$

where the matrix elements are given by $\langle \varphi_{\mathbf{k}'} | H_1 | \varphi_\mathbf{k} \rangle = \sum_G V_1 \int e^{-i\mathbf{k}'\cdot\mathbf{r}} e^{-i\mathbf{G}\cdot\mathbf{r}} e^{i\mathbf{k}\cdot\mathbf{r}} d^3\mathbf{r}$, which gives V_1 for $\mathbf{k} = \mathbf{k}' + \mathbf{G}$ and zero everywhere else. The energy eigenvalues are then given by the secular equation

$$\begin{vmatrix} \frac{\hbar^2 \mathbf{k}^2}{2m_0} + V_0 - E(\mathbf{k}) & V_1 \\ V_1 & \frac{\hbar^2 (\mathbf{k}-\mathbf{G})^2}{2m_0} + V_0 - E(\mathbf{k}) \end{vmatrix} = 0. \tag{F.4}$$

It is convenient to write wave vector with respect to the Brillouin zone edge $\kappa = \frac{\pi}{a} - k$ in one dimension. Solving F.4 gives

$$E(\kappa) = V_0 + E_1 + \frac{\hbar^2 \kappa^2}{2m_0} \pm \sqrt{V_1^2 + 4E_1 \frac{\hbar^2 \kappa^2}{2m_0}} \tag{F.5}$$

where

$$E_1 = \frac{\hbar^2}{2m_0} \left(\frac{G}{2} \right)^2. \tag{F.6}$$
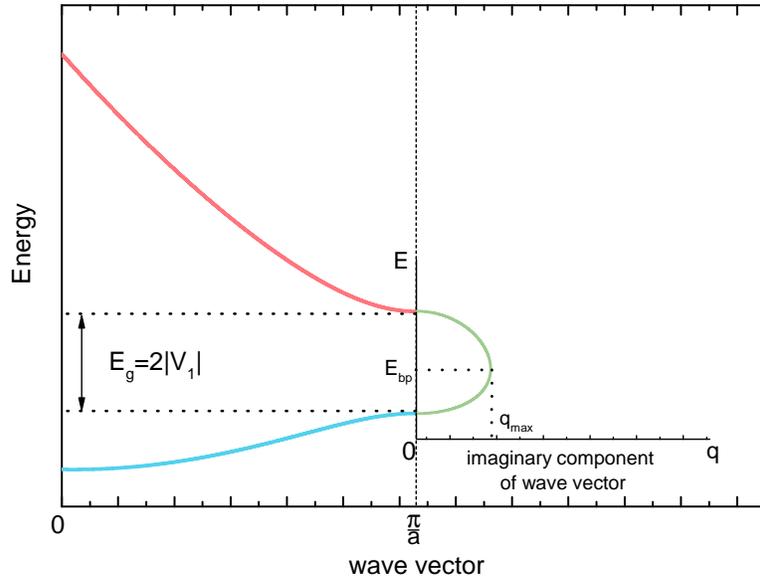
Figure F.2: Complex band structure in the reduced zone scheme for a one-dimensional nearly free electron solid. Within the band gap at the Brillouin zone edge, the imaginary component of the band structure is plotted. The maximum component of the wave vector q_{max} occurs just below mid gap.

Figure F.2 shows the familiar band dispersion for real wave vectors calculated using equation F.5, displaying an energy gap of $E_g = 2|V_1|$ at the Brillouin zone boundary.

This is only the band structure corresponding to real wave vectors however, and so now we consider the evanescent states, with complex wave vectors but real energies, i.e. those with negative κ^2 in equation F.5. These wave functions are not normally considered as in bulk materials, complex wave vectors mean the Bloch wave function cannot be normalized (is infinite when $z \rightarrow \infty$, see figure F.1). However, these wave functions physically describe the energy state at the surface, which exponentially decays into the vacuum and bulk of the solid due to the finite potential barrier at the surface. Within this model surface states may also exist within the band gap. Considering the imaginary component of the wave vector q, where $k = \frac{\pi}{a} + iq$, the dispersion becomes

$$E(q) = V_0 + E_1 - \frac{\hbar^2 q^2}{2 m_0} \pm \sqrt{V_1^2 - 4 E_1 \frac{\hbar^2 q^2}{2 m_0}}. \tag{F.7}$$

The dispersion calculated from equation F.7 is also plotted in figure F.2, where it clearly exists within the gap of the semiconductor. Because these states are meaningless in the bulk of the material, they are termed virtual gap states (ViGS) of the complex band structure. The complex band structure is related to the conventional band structure, only the periodicity (which is a boundary condition in conventional band structures) is broken, giving rise to solutions with complex k. These solutions are local in space and are occasionally referred to as edge states.

ViGS derive from the bulk band structure and so their character changes from predominantly donor-like near the VBM to acceptor like near the CBM. The energy at which the have equal donor and acceptor-like character is termed the branch-point energy E_{bp}. This happens when the decay length $\frac{1}{q}$ of the ViGS wave function is a minimum at q_{max}, usually somewhere close to mid-gap. In the one-dimensional lattice this occurs at $\frac{V_1^2}{3 E_1}$ below the mid-gap energy. Since the ViGS are predominantly donor-like below this energy, and acceptor like above (much like in the real band structure), this

branch-point energy can also be identified as the charge neutrality level (CNL) in a semiconductor[78,79], the energy level at which there is equal probability of a state (or defect) being a donor or an acceptor.
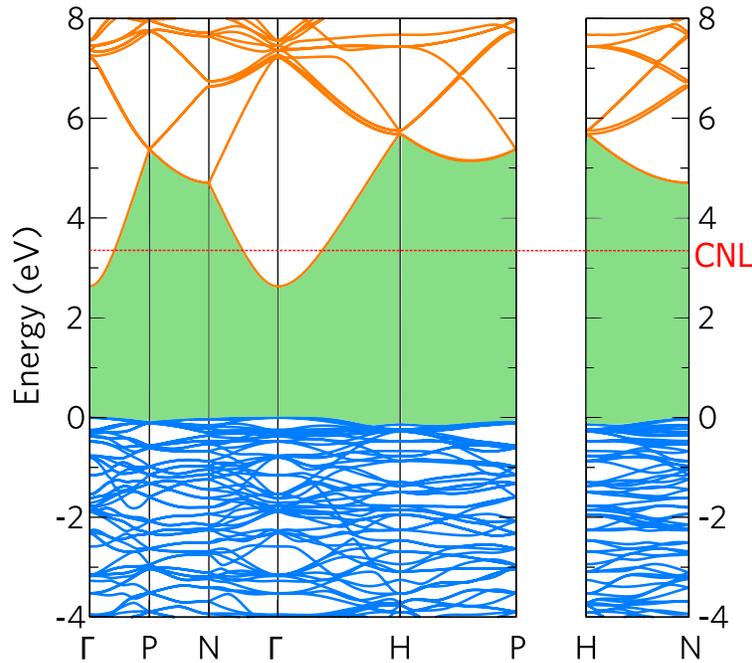
Figure F.3: DFT band structure and CNL of In_2O_3 according to King et al.[69]

Because ViGS are spatially very localized in real space, they are very dispersive in reciprocal space. Hence, their character is determined by contributions over a substantial portion of the entire Brillouin Zone, not just at the Γ-point band edges. Therefore, the branch point of the ViGS (CNL) will often be located around the mid-gap energy, averaged over the whole Brillouin zone. This includes all direct and indirect gaps, which explains why this point may often be located above the CBM in TCO materials, which most often display a single, narrow low lying CB, but all other bands are much flatter and CB and VB are much more greatly separated. This idea is displayed for In_2O_3 in figure F.3, with the value of the CNL taken from King et al.[69] Whilst this may not be directly the band-edge average, the CNL in figure F.3 is extremely close to

this value (deviations may arise due to a different band gap in the experimental work from King at al.[69] to that calculated here, Or possibly because we only have a sample of the calculated Brillouin zone.) Clearly, the low lying, and not very dispersive CB means the CNL is far from the middle of the band gap, and is infarct degenerate in the CB. This explains why the pinning position varies between semiconductors, whilst the CNL is thought to be universal for all materials when put on an absolute energy scale.

Many approaches have been taken to determine the CNL position in solids. Tersoff[78,386−388] determined the band structures for a number of materials using the LDA using the linearized augmented plane wave method (this is a way to approximate how the potential and so the wave function changes throughout the unit cell, for example within an atom to outside in the interstitial region[389]) in their calculations. Since they used a low level of theory in the LDA, a rigid energy shift was applied to correct the band gap. Tersoff determined the CNL to be the energy at which the valence and conduction band contributions to the real-space Green's functions were equal*. They averaged over 152 points within the Brillouin zone, which well represented the extended k-space behaviour associated with the ViGS, but also increased the computational expense. Cardona and Christensen[390], and Mönch[391] employed the mean-value points concept suggested by Baldereschi[392] to lower the computational cost. This method states that the value of any function which is periodic in wave vector, is representative of the average value of the function across the whole Brillouin zone at the mean-value point.

Mönch also showed that the energy gap at the mean-value point of the Brillouin zone is equal to the dielectric gap, and that the branch point energy of the ViGS, cal-

---

*Here the Green's function is given as $G(\mathbf{R}, E) = \int g(\mathbf{r}, \mathbf{r} + \mathbf{R}, E) d\mathbf{r} = \sum_{nk} \frac{e^{i(\mathbf{k} \cdot \mathbf{R})}}{E - E_{nk}}$ and $g(\mathbf{r}, \mathbf{r} + \mathbf{R}, E) = \sum_{nk} \frac{\varphi_{nk}^*(\mathbf{r}) \varphi_{nk}(\mathbf{r}')}{E - E_{nk}}$ where φ are Bloch wave functions of the band n, and $E_{n\mathbf{k}}$ is the energy of that band. $\mathbf{R}$ here is the lattice vector. As E increases $G_v(\mathbf{R}, E)$ decrease while $G_c(\mathbf{R}, E)$ increases. In this sense it may be thought of analogous to the probability of finding a carrier at a given energy, and where the two cross is where there is equal probability, i.e. the branch point energy.

culated through an empirical tight-binding method, has a linear relation with the dielectric band gap of the semiconductor[393]. The dielectric gap can be determined using a damped oscillator dielectric model in the low frequency limit, where

$$\varepsilon_\infty - 1 = \frac{\hbar\omega_p}{\hbar\omega_w} = \frac{\hbar\omega_p}{E_{dg}} \tag{F.8}$$

where E_{dg} is the dielectric gap energy.
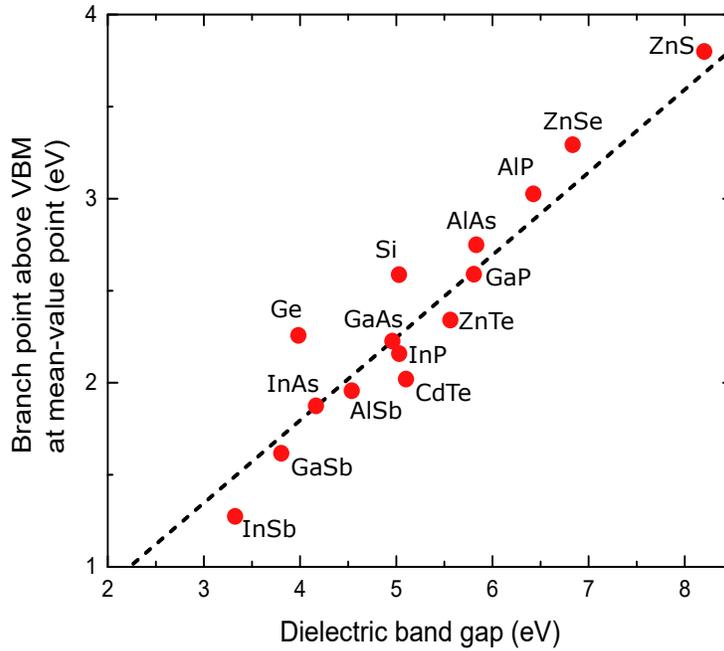
Figure F.4: Calculated branch point energy of the ViGS above the VBM versus the dielectric band gap for several materials. The dashed line is a linear fit with gradient 0.449. Figure adapted from Mönch[391].

This relationship is displayed in Figure F.4. Hence, for these simple binary materials investigated, the branch point energy is slightly below the mid gap, in agreement to the 1-dimensional linear lattice model discussed above. Tersoff[386] gave the semi-empirical formula $E'_{mid} = \frac{1}{2}(E'_V + E'_C)$ where $E'_V = E_V - \frac{1}{3}\Delta_{so}$ is the position of the 'effective' VBM in the absence of spin orbit splitting Δ_{so}, and E'_C is the indirect CBM. This give a very simple way to estimate the position of the CNL.

The physical origin of the ViGS for a clean semiconductor surface (i.e. the intrinsic surface states) is the dangling bonds which are created when a surface is generated, due to the termination of the bulk band structure. Because bonds are broken, each surface atom has fewer nearest neighbours to bond to, which means the surface atoms have to rearrange to minimize the energy lost by breaking said bonds. Hence, the surface is structurally different to the bulk. Extrinsic surface states may also be formed, which is when imperfections cause states to become localized at the surface. These can be formed from adatoms or defects for example. Adatoms at the surface will form bonds with surface atoms causing charge transfer to occurs, which is dictated by the electronegativity between surface atom and the adatom.

The surface ViGS, be they intrinsic or extrinsic, will be either donor or acceptor-like depending on whether they lie below or above the CNL. Additionally, they can display positively charged (unoccupied donor-like), negatively charged (occupied acceptor-like) or neutral (occupied donor-like/unoccupied acceptor-like) surface electronic behaviour, depending on the relative position of the Fermi level. When in the presence of a charged surface, carriers in this near-surface region rearrange to screen the electric field induced by these states, with the length scale over which this occurs being given by the Thomas-Fermi screening length. This length scale tends to be extremely short in metals due to their high carrier density, but is much longer in semiconductors leading to regions of charge redistribution known as space-charge regions. The total charge due to the surface states Q_{ss} must be compensated by an equal and opposite charge in the space-charge region Q_{sc}

$$Q_{ss} = -Q_{sc}.$$
(F.9)

As a result the Fermi-level shifts as a function of depth within the semiconductor. This is conventionally seen as a shifting of the valence and conduction bands rather than the Fermi level, which is known as band bending. The position of the surface Fermi level is determined by the condition for charge neutrality, given by equation F.9. In the case of an n-type semiconductor, where the surface Fermi level is located above (below) the CNL, the surface states are negatively (positively) charged, requiring an electron depletion (electron accumulation) layer near the surface to maintain charge neutrality. This is achieved by upward (downward) band bending to decrease (increase) the electron concentration at the surface with respect to the bulk. Figure 3.4 displays the types of band bending relevant to an n-type semiconductor. In certain situations there may be enough negatively charged surface states that the requirement for charge neutrality means the Fermi level is actually pushed below the middle of the band gap. This sets up a p-type surface, along with the n-type bulk, with the two region separated by a depletion layer, which is known as band inversion. Conversely, in some circumstances the Fermi level is directly at the CNL at the surface, and a flat band situation is set up, where the bands are not required to bend to maintain charge neutrality. Upward and downward band bending are also possible as displayed in figure 3.4.

References

[1] Streintz, F. *Annalen der Physik* **1902**, *314*, 854–885.

[2] Bädeker, K. *Annalen der Physik* **1907**, *327*, 749–766.

[3] Tiwari, A. N.; Khrypunov, G.; Kurdzesau, F.; Bätzner, D. L.; Romeo, A.; Zogg, H. *Progress in Photovoltaics: Research and Applications* **2004**, *12*, 33–38.

[4] Granqvist, C. G. *Solar Energy Materials and Solar Cells* **2007**, *91*, 1529 – 1598.

[5] Hamberg, I.; Granqvist, C. G. *Journal of Applied Physics* **1986**, *60*, R123–R160.

[6] Scott, R.; Yang, S.; Coombs, N.; Ozin, G.; Williams, D. *Advanced Functional Materials* **2003**, *13*, 225–231.

[7] Semancik, S.; Cavicchi, R. *Thin Solid Films* **1991**, *206*, 81 – 87.

[8] Wu, H.; Hu, L.; Carney, T.; Ruan, Z.; Kong, D.; Yu, Z.; Yao, Y.; Cha, J. J.; Zhu, J.; Fan, S.; Cui, Y. *Journal of the American Chemical Society* **2011**, *133*, 27–29.

[9] Granqvist, C. G. *Applied Physics A* **1993**, *57*, 19–24.

[10] Theis, T. N.; Solomon, P. M. *Science* **2010**, *327*, 1600–1601.

[11] Gordon, T. *Nature* **1997**, *389*, 907–908.

[12] Wager, J. F. *Science* **2003**, *300*, 1245–1246.

[13] Ramirez, A. P. *Science* **2007**, *315*, 1377–1378.

[14] Heber, J. *Nature* **2009**, *459*, 28–30.

[15] Takagi, H.; Hwang, H. Y. *Science* **2010**, *327*, 1601–1602.

[16] Mannhart, J.; Schlom, D. G. *Science* **2010**, *327*, 1607–1611.

[17] Ohta, H.; Hosono, H. *Materials Today* **2004**, *7*, 42 – 51.

[18] Nomura, K.; Ohta, H.; Ueda, K.; Kamiya, T.; Hirano, M.; Hosono, H. *Science* **2003**, *300*, 1269–1272.

[19] Tsukazaki, A.; Ohtomo, A.; Kita, T.; Ohno, Y.; Ohno, H.; Kawasaki, M. *Science* **2007**, *315*, 1388–1391.

[20] Tsukazaki, A.; Akasaka, K., S. amd Nakahara; Ohno, Y.; Ohno, H.; Maryenko, D.; Ohtomo, A.; Kawasaki, M. *Nature* **2010**, *9*, 889–893.

[21] Haynes, W. M. *CRC Handbook of Chemistry and Physics, 95th Edition*; CRC Press, 2014.

[22] Minami, T. *Semiconductor Science and Technology* **2005**, *20*, S35–S44.

[23] *Transparent Conductive Films Market By Application Type (Smartphones, Tablets, Notebooks, LCDs, Wearable Devices); On the basis of Material (ITO on Glass, ITO on PET, Silver Nanowires, Metal Mesh, Carbon Nanotubes) and by Regional Analysis - Global Forecast*; 2017; p 125.

[24] Pasquarelli, R. M.; Ginley, D. S.; O'Hayre, R. *Chem. Soc. Rev.* **2011**, *40*, 5406–5441.

[25] Walsh, A.; Da Silva, J. L. F.; Wei, S.-H.; Körber, C.; Klein, A.; Piper, L. F. J.; De-Masi, A.; Smith, K. E.; Panaccione, G.; Torelli, P.; Payne, D. J.; Bourlange, A.; Egdell, R. G. *Phys. Rev. Lett.* **2008**, *100*, 167402.

[26] Yu, P.; Cardona, M. *Fundamentals of Semiconductors*, 4th ed.; Springer, 2010.

[27] Burstein, E. *Phys. Rev.* **1954**, *93*, 632–633.

[28] Moss, T. S. *Proceedings of the Physical Society. Section B* **1954**, *67*, 775–782.

[29] Mott, N. F. *The Philosophical Magazine: A Journal of Theoretical Experimental and Applied Physics* **1961**, *6*, 287–309.

[30] Zhang, K. H. L.; Egdell, R. G.; Offi, F.; Iacobucci, S.; Petaccia, L.; Gorovikov, S.; King, P. D. C. *Phys. Rev. Lett.* **2013**, *110*, 056803.

[31] Fuchs, F.; Bechstedt, F. *Phys. Rev. B* **2008**, *77*, 155107.

[32] Lu, J. G.; Fujita, S.; Kawaharamura, T.; Nishinaka, H.; Kamada, Y.; Ohshima, T.; Ye, Z. Z.; Zeng, Y. J.; Zhang, Y. Z.; Zhu, L. P.; He, H. P.; Zhao, B. H. *Journal of Applied Physics* **2007**, *101*, 083705.

[33] Hamberg, I.; Granqvist, C. G.; Berggren, K. F.; Sernelius, B. E.; Engström, L. *Phys. Rev. B* **1984**, *30*, 3240–3249.

[34] Jefferson, P. H.; Hatfield, S. A.; Veal, T. D.; King, P. D. C.; McConville, C. F.; Zúñiga-Pèrez, J.; Muñoz-Sanjosè, V. *Applied Physics Letters* **2008**, *92*, 022101.

[35] Berggren, K. F.; Sernelius, B. E. *Phys. Rev. B* **1981**, *24*, 1971–1986.

[36] Inkson, J. C. *Journal of Physics C: Solid State Physics* **1976**, *9*, 1177–1183.

[37] Abram, R.; Rees, G.; Wilson, B. *Advances in Physics* **1978**, *27*, 799–892.

[38] Walsh, A.; Da Silva, J. L. F.; Wei, S.-H. *Phys. Rev. B* **2008**, *78*, 075211.

[39] Jain, S. C.; McGregor, J. M.; Roulston, D. J. *Journal of Applied Physics* **1990**, *68*, 3747–3749.

[40] Lindefelt, U. *Journal of Applied Physics* **1998**, *84*, 2628–2637.

[41] Auvergne, D.; Camassel, J.; Mathieu, H. *Phys. Rev. B* **1975**, *11*, 2251–2259.

[42] Feneberg, M.; Nixdorf, J.; Lidig, C.; Goldhahn, R.; Galazka, Z.; Bierwagen, O.; Speck, J. S. *Phys. Rev. B* **2016**, *93*, 045203.

[43] Pisarkiewicz, T.; Kolodziej, A. *physica status solidi (b)* **1990**, *158*, K5–K8.

[44] Pisarkiewicz, T.; Zakrzewska, K.; Leja, E. *Thin Solid Films* **1989**, *174*, 217 – 223.

[45] Young, D. L.; Coutts, T. J.; Kaydanov, V. I.; Gilmore, A. S.; Mulligan, W. P. *Journal of Vacuum Science & Technology A* **2000**, *18*, 2978–2985.

[46] Pickering, C. *J. Phys. C: Solid St. Phys.* **1980**, *13*, 2959–2968.

[47] Kittel, C. *Introduction to Solid State Physicss, 8th ed.*; John Wiley & Sons, 2005.

[48] Whalley, L. D.; Frost, J. M.; Morgan, B. J.; Walsh, A. *Phys. Rev. B* **2019**, *99*, 085207.

[49] Huy, H. A.; Aradi, B.; Frauenheim, T.; Deák, P. *Phys. Rev. B* **2011**, *83*, 155201.

[50] Wooten, F. *Optical Properties of Solids*; Academic Press, 1972.

[51] Egdell, R. G. In *Defects at Oxide Surfaces*; Jupille, J., Thornton, G., Eds.; Springer Series in Surface Sciences; Springer International Publishing: Cham, 2015; pp 351–400.

[52] Spitzer, W. G.; Fan, H. Y. *Phys. Rev.* **1957**, *106*, 882.

[53] Inaoka, T.; Newns, D. M.; Egdell, R. G. *Surf. Sci.* **1987**, *186*, 290–308.

[54] Dresselhaus, M. S.; Dresselhaus, G.; Jorio, A. *Group Theory*; Springer, 2009.

[55] Samson, S.; Fonstad, C. G. *J. Appl. Phys.* **1973**, *44*, 4618–4621.

[56] Ágoston, P.; Albe, K.; Nieminen, R. M.; Puska, M. J. *Phys. Rev. Letters* **2009**, *103*, 245501.

[57] Fonstad, C. G.; Rediker, R. H. *Journal of Applied Physics* **1971**, *42*, 2911–2918.

[58] Speaks, D. T.; Mayer, M. A.; Yu, K. M.; Mao, S. S.; Haller, E. E.; Walukiewicz, W. *Journal of Applied Physics* **2010**, *107*, 113706.

[59] King, P. D. C.; Veal, T. D.; Jefferson, P. H.; Zúñiga-Pérez, J.; Muñoz-Sanjosé, V.; McConville, C. F. *Phys. Rev. B* **2009**, *79*, 035203.

[60] Tuomisto, F.; Saarinen, K.; Look, D. C.; Farlow, G. C. *Phys. Rev. B* **2005**, *72*, 085206.

[61] Zubiaga, A.; Tuomisto, F.; Coleman, V. A.; Tan, H. H.; Jagadish, C.; Koike, K.; Sasa, S.; Inoue, M.; Yano, M. *Phys. Rev. B* **2008**, *78*, 035125.

[62] Singh, A. K.; Janotti, A.; Scheffler, M.; de Walle, C. G. V. *Phys. Rev. Letters* **2008**, *101*, 055502.

[63] de Walle, C. G. V. *Phys. Rev. Letters* **2000**, *85*, 1012.

[64] Janotti, A.; de Walle, C. G. V. *Appl. Phys. Letters* **2005**, *87*, 122102.

[65] Lany, S.; Zunger, A. *Phys. Rev. Lett.* **2007**, *98*, 045501.

[66] Lany, S.; Zunger, A. *Phys. Rev. B* **2005**, *72*, 035215.

[67] Janotti, A.; Van de Walle, C. G. *Phys. Rev. B* **2007**, *76*, 165202.

[68] Godinho, K. G.; Walsh, A.; Watson, G. W. *J. Phys. Chem. C* **2009**, *113*, 439–448.

[69] King, P. D. C.; Lichti, R. L.; Celebi, Y. G.; Gil, J. M.; Vilão, R. C.; Alberto, H. V.; Piroto Duarte, J.; Payne, D. J.; Egdell, R. G.; McKenzie, I.; McConville, C. F.; Cox, S. F. J.; Veal, T. D. *Phys. Rev. B* **2009**, *80*, 081201.

[70] de Walle, C. G. V. *Nature* **2003**, *423*, 626–628.

[71] King, P. D. C.; Veal, T. D. *Journal of Physics: Condensed Matter* **2011**, *23*, 334214.

[72] Thomas, D. G.; Lander, J. J. *The Journal of Chemical Physics* **1956**, *25*, 1136–1142.

[73] Cox, S. F. J.; Davis, E. A.; Cottrell, S. P.; King, P. J. C.; Lord, J. S.; Gil, J. M.; Alberto, H. V.; Vilão, R. C.; Piroto Duarte, J.; Ayres de Campos, N.; Weidinger, A.; Lichti, R. L.; Irvine, S. J. C. *Phys. Rev. Lett.* **2001**, *86*, 2601–2604.

[74] Koida, T.; Fujiwara, H.; Kondo, M. *Japanese Journal of Applied Physics* **2007**, *46*, L685–L687.

[75] Janotti, A.; de Walle, C. G. V. *Nature Materials* **2006**, *6*, 44–47.

[76] Lavrov, E. V.; Börrnert, F.; Weber, J. *Phys. Rev. B* **2005**, *71*, 035205.

[77] Lavrov, E. V.; Börrnert, F.; Weber, J. *Phys. Rev. B* **2005**, *72*, 085212.

[78] Tersoff, J. *Phys. Rev. Lett.* **1984**, *52*, 465–468.

[79] Tejedor, C.; Flores, F. *Journal of Physics C: Solid State Physics* **1977**, *11*, L19–L23.

[80] Heine, V. *Phys. Rev.* **1965**, *138*, A1689–A1696.

[81] Mönch, W. *Semiconductor Surfaces and Interfaces, Third Edition*; Springer-Verlag Berlin Heidelberg, 2001.

[82] King, P. D. C.; Veal, T. D.; Fuchs, F.; Wang, C. Y.; Payne, D. J.; Bourlange, A.; Zhang, H.; Bell, G. R.; Cimalla, V.; Ambacher, O.; Egdell, R. G.; Bechstedt, F.; McConville, C. F. *Phys. Rev. B* **2009**, *79*, 205211.

[83] Veal, T. D.; King, P. D. C.; Hatfield, S. A.; Bailey, L. R.; McConville, C. F.; Martel, B.; Moreno, J. C.; Frayssinet, E.; Semond, F.; Zúñiga Pérez, J. *Applied Physics Letters* **2008**, *93*, 202108.

[84] Falabretti, B.; Robertson, J. *Journal of Applied Physics* **2007**, *102*, 123703.

[85] Ginley, D.; Hosono, H.; Paine, D. C. *Handbook of Transparent Conductors*; Springer Science and Business Media, 2010.

[86] Hoyer, K. L.; Hubmann, A. H.; Klein, A. *physica status solidi (a)* **2017**, *214*, 1600486.

[87] Frischbiera, M. V.; Wardenga, H. F.; Weidner, M.; Bierwagen, O.; Jia, J.; Shigesato, Y.; Klein, A. *Thin Solid Films* **2016**, *614*, 62 – 68, 9th International Symposium on Transparent Oxide and Related Materials for Electronics and Optics (TOEO9).

[88] Bhachu, D. S.; Scanlon, D. O.; Sankar, G.; Veal, T. D.; Egdell, R. G.; Cibin, G.; Dent, A. J.; Knapp, C. E.; Carmalt, C. J.; Parkin, I. P. *Chemistry of Materials* **2015**, *27*, 2788–2796.

[89] Scanlon, D. O. *Phys. Rev. B* **2013**, *87*, 161201(R).

[90] Dingle, R.; Störmer, H. L.; Gossard, A. C.; Wiegmann, W. *Appl. Phys. Lett.* **1978**, *33*, 665.

[91] Siegbahn, K. *Science* **1982**, *217*, 111–121.

[92] Hofmann, S. *Auger- and X-Ray Photoelectron Spectroscopy in Materials Science*; Springer-Verlag, 2013.

[93] Briggs, D.; Seah, M. P. *Practical Surface Analysis: Auger and X-ray Photoelectron Spectroscopy*; Wiley-Blackwell; 2nd Edition, 1994.

[94] Briggs, D.; Grant, J. T. *Surface Analysis by Auger and X-ray Photoelectron Spectroscopy*; IM Publications and SurfaceSpectra Limited, 2003.

[95] Hüfner, S. *Photoelectron Spectroscopy Principles and Applications, 3ʳᵈ edt.*; Springer-Verlag Berlin Heidleberg, 2003.

[96] Steiner, P.; Höchst, H.; Hüfner, S. *Zeitschrift für Physik B Condensed Matter* **1978**, *30*, 129–143.

[97] Langreth, D. C. *Phys. Rev. B* **1970**, *1*, 471–477.

[98] Hedin, L. *Journal of Physics: Condensed Matter* **1999**, *11*, R489–R528.

[99] Hedin, L.; Lundqvist, B. I.; Lundqvist, S. *J. Res. NBS A* **1970**, *74A*, 417.

[100] Langreth, D. C. *Phys. Rev. Lett.* **1971**, *26*, 1229–1233.

[101] Chazalviel, J. N.; Campagna, M.; Wertheim, G. K.; Shanks, H. R. *Phys. Rev. B* **1977**, *16*, 697–705.

[102] Pardee, W. J.; Mahan, G. D.; Eastman, D. E.; Pollak, R. A.; Ley, L.; McFeely, F. R.; Kowalczyk, S. P.; Shirley, D. A. *Phys. Rev. B* **1975**, *11*, 3614–3616.

[103] Chang, J.-J.; Langreth, D. C. *Phys. Rev. B* **1973**, *8*, 4638–4654.

[104] Langreth, D. *Nobel Symp.* **1973**, *24*, 210.

[105] Egdell, R.; Walker, T.; Beamson, G. *Journal of Electron Spectroscopy and Related Phenomena* **2003**, *128*, 59 – 66.

[106] Egdell, R. G.; Rebane, J.; Walker, T. J.; Law, D. S. L. *Phys. Rev. B* **1999**, *59*, 1792–1799.

[107] Kotani, A.; Toyozawa, Y. *Journal of the Physical Society of Japan* **1974**, *37*, 912–919.

[108] Borgatti, F. et al. *Phys. Rev. B* **2018**, *97*, 155102.

[109] Weidner, M. Fermi Level Determination in Tin Oxide by Photoelectron Spectroscopy - Relation to Optoelectronic Properties; Band Bending at Surfaces and Interfaces; Modulation Doping. Ph.D. thesis, TU Darmstadt, 2016.

[110] Seah, M. P.; Dench, W. A. *Surface and Interface Analysis* **1979**, *1*, 2–11.

[111] Powell, C. J. *Journal of Electron Spectroscopy and Related Phenomena* **1988**, *47*, 197–214.

[112] Tanuma, S.; Powell, C. J.; Penn, D. R. *Surface and Interface Analysis* **2005**, *37*, 1–14.

[113] Tanuma, S.; Powell, C. J.; Penn, D. R. *Surface and Interface Analysis* **1994**, *21*, 165–176.

[114] Penn, D. R. *Phys. Rev. B* **1987**, *35*, 482–486.

[115] Tung, C.; Ashley, J.; Ritchie, R. *Surface Science* **1979**, *81*, 427 – 439.

[116] Quinn, J. J. *Phys. Rev.* **1962**, *126*, 1453–1457.

[117] Tanuma, S.; Powell, C. J.; Penn, D. R. *Surface and Interface Analysis* **1988**, *11*, 577–589.

[118] Ashley, J. C. *Journal of Electron Spectroscopy and Related Phenomena* **1982**, *28*, 177–194.

[119] Ashley, J. C. *Journal of Electron Spectroscopy and Related Phenomena,* **1988**, *46*, 199–214.

[120] Williams, M. W.; Young, D. W.; Ashley, J. C.; Arakawa, E. T. *Journal of Applied Physics* **1985**, *58*, 4360–4364.

[121] Yarzhemsky, V. G.; Nefedov, V. I.; Trzhaskovskaya, M. B.; Band, I. M.; Szargan, R. *Journal of Electron Spectroscopy and Related Phenomena* **2002**, *123*, 1 – 10.

[122] Moulder, J. F.; Stickle, W. F.; Sobol, P. E.; Bomben, K. D. *Handbook of X-ray Photoelectron Spectroscopy: A Reference Book of Standard Spectra for Identification and Interpretation of XPS Data*; Physical Electronics Division, Perkin-Elmer Corporation, 1992.

[123] Scofield, J. *Theoretical photoionization cross sections from 1 to 1500 keV.*; 1973.

[124] Yeh, J.; Lindau, I. *Atomic Data and Nuclear Data Tables* **1985**, *32*, 1 – 155.

[125] Trzhaskovskaya, M.; Nefedov, V.; Yarzhemsky, V. *Atomic Data and Nuclear Data Tables* **2002**, *82*, 257 – 311.

[126] Trzhaskovskaya, M.; Nefedov, V.; Yarzhemsky, V. *Atomic Data and Nuclear Data Tables* **2001**, *77*, 97 – 159.

[127] Trzhaskovskaya, M.; Nikulin, V.; Nefedov, V.; Yarzhemsky, V. *Atomic Data and Nuclear Data Tables* **2006**, *92*, 245 – 304.

[128] Bethe, H. A.; Salpeter, E. W. *Quantum Mechanics of One- and Two-Electron Atoms*; Springer-Verlag Berlin Heidelberg, 1957.

[129] Fox, M. *Optical Properties of Solids*; Oxford University Press, 2001.

[130] Pratt, R. H.; Ron, A.; Tseng, H. K. *Rev. Mod. Phys.* **1973**, *45*, 273–325.

[131] Goldberg, S.; Fadley, C.; Kono, S. *Journal of Electron Spectroscopy and Related Phenomena* **1981**, *21*, 285 – 363.

[132] Cooper, J. W. *Phys. Rev. A* **1993**, *47*, 1841–1851.

[133] Cooper, J. W. *Phys. Rev.* **1962**, *128*, 681–693.

[134] Manson, S. T.; Cooper, J. W. *Phys. Rev.* **1968**, *165*, 126–138.

[135] Payne, D. J. The electronic structure of post transition metal oxides. Ph.D. thesis, The University of Oxford, 2008.

[136] Cooper, J. W. *Phys. Rev. A* **1990**, *42*, 6942–6945.

[137] Körber, C.; Krishnakumar, V.; Klein, A.; Panaccione, G.; Torelli, P.; Walsh, A.; Silva, J. L. F. D.; Wei, S.-H.; Egdell, R. G.; Payne, D. J. *Physical Review B* **2010**, *81*, 165207.

[138] Vickerman, J. C.; Gilmore, I. S. *Surface Analysis The principal Techniques*; John Wiley & Sons, Ltd., 2009.

[139] Citrin, P. H.; Hamann, D. R. *Phys. Rev. B* **1977**, *15*, 2923–2928.

[140] Fujikawa, T.; Arai, H. *Journal of Electron Spectroscopy and Related Phenomena* **2009**, *174*, 85–92.

[141] Shirley, D. A. *Phys. Rev. B* **1972**, *5*, 4709–4714.

[142] Végh, J. *Journal of Electron Spectroscopy and Related Phenomena* **2006**, *151*, 159 – 164.

[143] Tougaard, S. *Surface and Interface Analysis* **1988**, *11*, 453–472.

[144] Evans, S. *Surface Interface Analysis* **1991**, *17*, 85–93.

[145] Fairley, N.; Carrick, A. *The Casa Cookbook: Part 1: Recipes for XPS data processing*; Acolyte Science, 2005.

[146] Doniach, S.; Sunjic, M. *Journal of Physics C: Solid State Physics* **1970**, *3*, 285–291.

[147] Dolgonos, A.; Mason, T. O.; Poeppelmeier, K. R. *Journal of Solid State Chemistry* **2016**, *240*, 43 – 48.

[148] Katsidis, C. C.; Siapkas, D. I. *Applied Optics* **2002**, *41*, 3978–3987.

[149] Born, M.; Wolf, E. *Principles of Optics*; Pergamon Press, 1975.

[150] Ashcroft, N. W.; Mermin, N. D. *Solid State Physics*; Brooks/Cole, 1976.

[151] Preissler, N.; Bierwagen, O.; Ramu, A. T.; Speck, J. S. *Physical Review B* **2013**, *88*, 085305.

[152] Fonstad, C. G.; Rediker, R. H. *Journal of Applied Physics* **1971**, *42*, 2911–2918.

[153] van der Pauw, L. J. *Philips Research Report* **1958**, *13*, 1–9.

[154] van der Pauw, L. J. *Philips Technical Report* **1958**, *20*, 220–224.

[155] Nag, B. R. *Electron Transport in Compound Semiconductors*; Springer-Verlag Berlin Heidelberg, 1980.

[156] Bate, R. T.; Baxter, R. D.; Reid, F. J.; Beer, A. C. *Journal of Physics and Chemistry of Solids* **1965**, *26*, 1205–1214.

[157] Appel, J. *Physical Review* **1962**, *125*, 1815.

[158] Conwell, E.; Weisskopf, V. F. *Physical Review* **1950**, *77*, 388–390.

[159] Chattopadhyay, D.; Queisser, H. J. *Reviews of Modern Physics* **1981**, *53*, 745–768.

[160] Look, D. C. *Electrical Characterization of GaAs Materials and Devices*; Wiley-Blackwell, 1988.

[161] Ridley, B. K. *Quantum Processes in Semiconductors, Fifth ed.*; Oxford University Press, 2013.

[162] Zawadzki, W.; Szymańska, W. *Physica Status Solidi B* **1971**, *45*, 415–432.

[163] Zawadzki, W. *Handbook of Semiconductors*; North-Holland, 1982; Chapter 12, pp 713–803.

[164] Brooks, H. *Advances in Electronics and Electron Physics* **1955**, *7*, 85–182.

[165] Dingle, R. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* **1955**, *46*, 831–840.

[166] Aida, H.; Nishiguchi, K.; Takeda, H.; Aota, N.; Sunakawa, K.; Yaguchi, Y. *Jpn. J. Appl. Phys.* **2008**, *47*, 8506–8509.

[167] Marchand, P.; Hassan, I. A.; Parkin, I. P.; Carmalt, C. J. *Dalton Transactions* **2013**, *42*, 9406–9422.

[168] Hamaguchi, C. *Basic Semiconductor Physics, second edition*; Springer-Verlag Berlin Heidelberg, 2010.

[169] Pidgeon, C. R.; Brown, R. N. *Phys. Rev.* **1966**, *146*, 575–583.

[170] Kane, E. O. *Journal of Physics and Chemistry of Solids* **1956**, *1*, 82–99.

[171] Kane, E. O. *Journal of Physics and Chemistry of Solids* **1957**, *1*, 249–261.

[172] Kane, E. In *Semiconductors and Semimetals*; Willardson, R., Beer, A. C., Eds.; Semiconductors and Semimetals; Elsevier, 1966; Vol. 1; pp 75 – 100.

[173] Martin, R. M. *Electronic Structure: Basic Theory and Practical Methods*; Cambridge University Press, 2004.

[174] Hohenberg, P.; Kohn, W. *Phys. Rev.* **1964**, *136*, B864–B871.

[175] Kohn, W.; Sham, L. J. *Phys. Rev.* **1965**, *140*, A1133–A1138.

[176] Löwdin, P.-O. *Phys. Rev.* **1955**, *97*, 1509–1520.

[177] Wigner, E. *Phys. Rev.* **1934**, *46*, 1002–1011.

[178] Fock, V. *Zeitschrift fur Physik* **1930**, *61*, 126–148.

[179] Slater, J. C. *Phys. Rev.* **1929**, *34*, 1293–1322.

[180] Heyd, J.; Peralta, J. E.; Scuseria, G. E.; Martin, R. L. *The Journal of Chemical Physics* **2005**, *123*, 174101.

[181] Furthmüller, J.; Bechstedt, F. *Phys. Rev. B* **2016**, *93*, 115204.

[182] Swallow, J. E. N.; Varley, J. B.; Jones, L. A. H.; Gibbon, J. T.; Piper, L. F. J.; Dhanak, V. R.; Veal, T. D. *APL Materials* **2019**, *7*, 022528.

[183] Scanlon, D. O.; Watson, G. W. *Journal of Materials Chemistry* **2012**, *22*, 25236–25245.

[184] Clark, S. J.; Robertson, J.; Lany, S.; Zunger, A. *Phys. Rev. B* **2010**, *81*, 115311.

[185] Perdew, J. P. *Phys. Rev. Lett.* **1985**, *55*, 1665–1668.

[186] Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.

[187] Perdew, J. P.; Ernzerhof, M.; Burke, K. *The Journal of Chemical Physics* **1996**, *105*, 9982–9985.

[188] Heyd, J.; Scuseria, G. E.; Ernzerhof, M. *The Journal of Chemical Physics* **2003**, *118*, 8207–8215.

[189] Adamo, C.; Barone, V. *The Journal of Chemical Physics* **1999**, *110*, 6158–6170.

[190] King, P. D. C.; Veal, T. D.; McConville, C. F. *Phys. Rev. B* **2008**, *77*, 125305.

[191] Veal, T. D.; Piper, L. F. J.; Schaff, W. J.; McConville, C. F. *Journal of Crystal Growth* **2006**, *288*, 268 – 272.

[192] Paasch, G.; Übensee, H. *physica status solidi (b)* **1982**, *113*, 165–178.

[193] Zöllner, J.-P.; Übensee, H.; Paasch, G.; Fiedler, T.; Gobsch, G. *physica status solidi (b)* **1986**, *134*, 837–845.

[194] King, P. D. C.; Veal, T. D.; McConville, C. F.; Zúñiga Pérez, J.; Muñoz Sanjosé, V.; Hopkinson, M.; Rienks, E. D. L.; Jensen, M. F.; Hofmann, P. *Phys. Rev. Lett.* **2010**, *104*, 256803.

[195] Elangovan, E.; Ramamurthi, K. *Applied Surface Science* **2005**, *249*, 183–196.

[196] Agashe, C.; Major, S. S. *Journal of Materials Science* **1996**, *31*, 2965–2969.

[197] Rey, G.; Ternon, C.; Modreanu, M.; Mescot, X.; Consonni, V.; Bellet, D. *Journal of Applied Physics* **2013**, *114*, 183713.

[198] Park, J. H.; Byun, D. J.; Lee, J. K. *Journal of Electroceramics* **2009**, *23*, 506–511.

[199] Kim, H.; Auyeung, R. C. Y.; Piqué, A. *Thin Solid Films* **2008**, *516*, 5052–5056.

[200] Geoffroy, C.; Campet, G.; Menil, F.; Portier, J.; Salardenne, J.; Couturier, G. *Active and Passive Electronic Components* **1991**, *14*, 111–118.

[201] Bhachu, D. S.; Waugh, M. R.; Zeissler, K.; Branford, W. R.; Parkin, I. P. *Chemistry - A European Journal* **2011**, *17*, 11613–11621.

[202] Reimann, K.; Steube, M. *Solid State Communications* **1998**, *105*, 649–652.

[203] Arlinghaus, F. J. *Journal of Physics and Chemistry of Solids* **1974**, *35*, 931–935.

[204] Rakhshani, A. E.; Makdisi, Y.; Ramazaniyan, H. A. *Journal of Applied Physics* **1998**, *83*, 1049–1057.

[205] Noor, N.; Chew, C. K. T.; Bhachu, D. S.; Waugh, M. R.; Carmalt, C. J.; Parkin, I. P. *J. Mater. Chem. C* **2015**, *3*, 9359–9368.

[206] Xu, J.; Huang, S.; Wang, Z. *Solid State Communications* **2009**, *149*, 527–531.

[207] Noor, N.; Parkin, I. P. *J. Mater. Chem. C* **2013**, *1*, 984–996.

[208] Shanthi, E.; Banerjee, A.; Dutta, V.; Chopra, K. L. *Journal of Applied Physics* **1982**, *53*, 1615–1621.

[209] Granqvist, C. G. *Solar Energy Materials and Solar Cells* **2007**, *91*, 1529–1598.

[210] Porch, A.; Morgan, D. V.; Perks, R. M.; Jones, M. O.; Edwards, P. P. *Journal of Applied Physics* **2004**, *95*, 4734–4737.

[211] Edwards, P. P.; Porch, A.; Jones, M. O.; Morgan, D. V.; Perks, R. M. *Dalton Transactions* **2004**, 2995–3002.

[212] Acosta, D.; Zironi, E.; Montoya, E.; Estrada, W. *Thin Solid Films* **1996**, *288*, 1–7.

[213] Fantini, M.; Torriani, I. *Thin Solid Films* **1986**, *138*, 255–265.

[214] Suffner, J.; Ágoston, P.; Kling, J.; Hahn, H. *Journal of Nanoparticle Research* **2010**, *12*, 2579–2588.

[215] Thangaraju, B. *Thin Solid Films* **2002**, *402*, 71–78.

[216] Bae, J. W.; Lee, S. W.; Yeom, G. Y. *Journal of The Electrochemical Society* **2007**, *154*, D34–D37.

[217] Canestraro, C. D.; Oliveira, M. M.; Valaski, R.; da Silva, M. V. S.; David, D. G. F.; Pepe, I.; Silva, a. F. D.; Roman, L. S.; Persson, C. *Applied Surface Science* **2008**, *255*, 1874–1879.

[218] Chinnappa, L.; Ravichandran, K.; Saravanakumar, K.; Muruganantham, G.; Sakthivel, B. *Journal of Materials Science: Materials in Electronics* **2011**, *22*, 1827–1834.

[219] Whittles, T. J.; Burton, L. A.; Skelton, J. M.; Walsh, A.; Veal, T. D.; Dhanak, V. R. *Chemistry of Materials* **2016**, *28*, 3718–3726.

[220] Pandey, R.; Cho, S. H.; Hwang, D. K.; Choi, W. K. *Current Applied Physics* **2014**, *14*, 850–855.

[221] Pan, Z.; Zhang, P.; Tian, X.; Cheng, G.; Xie, Y.; Zhang, H.; Zeng, X.; Xiao, C.; Hu, G.; Wei, Z. *Journal of Alloys and Compounds* **2013**, *576*, 31–37.

[222] Kresse, G.; Hafner, J. *Physical Review B* **1993**, *47*, 558–561.

[223] Kresse, G.; Hafner, J. *Physical Review B* **1994**, *49*, 14251–14269.

[224] Kresse, G.; Furthmüller, J. *Phys. Rev. B* **1996**, *54*, 11169.

[225] Kresse, G.; Furthmüller, J. *Comput. Mater. Sci.* **1996**, *6*, 15.

[226] Blöchl, P. E. *Physical Review B* **1994**, *50*, 17953–17979.

[227] Kresse, G.; Joubert, D. *Phys. Rev. B* **1999**, *59*, 1758.

[228] Paier, J.; Hirschl, R.; Marsman, M.; Kresse, G. *J. Chem. Phys.* **2005**, *122*, 234102.

[229] Kafizas, A.; Noor, N.; Carmichael, P.; Scanlon, D. O.; Carmalt, C. J.; Parkin, I. P. *Advanced Functional Materials* **2013**, *24*, 1758–1771.

[230] Williamson, B. A. D.; Buckeridge, J.; Brown, J.; Ansbro, S.; Palgrave, R. G.; Scanlon, D. O. *Chemistry of Materials* **2016**, DOI: 10.1021/acs.chemmater.6b03306.

[231] Marchand, P.; Sathasivam, S.; Williamson, B. A. D.; Pugh, D.; Bawaked, S. M.; Basahel, S. N.; Obaid, A. Y.; Scanlon, D. O.; Parkin, I. P.; Carmalt, C. J. *J. Mater. Chem. C* **2016**, *4*, 6761–6768.

[232] Sathasivam, S.; Williamson, B. A. D.; Kafizas, A.; Althabaiti, S. A.; Obaid, A. Y.; Basahel, S. N.; Scanlon, D. O.; Carmalt, C. J.; Parkin, I. P. *The Journal of Physical Chemistry C* **2017**, *121*, 202–210.

[233] Oba, F.; Togo, A.; Tanaka, I.; Paier, J.; Kresse, G. *Physical Review B* **2008**, *77*, 245202.

[234] Chen, H.; Dawson, J. A. *The Journal of Physical Chemistry C* **2015**, *119*, 15890–15895.

[235] Bhachu, D. S.; Sathasivam, S.; Sankar, G.; Scanlon, D. O.; Cibin, G.; Carmalt, C. J.; Parkin, I. P.; Watson, G. W.; Bawaked, S. M.; Obaid, A. Y.; Al-Thabaiti, S.; Basahel, S. N. *Advanced Functional Materials* **2014**, *24*, 5075–5085.

[236] Burbano, M.; Scanlon, D. O.; Watson, G. W. *Journal of the American Chemical Society* **2011**, *133*, 15065–15072.

[237] Ágoston, P.; Körber, C.; Klein, A.; Puska, M. J.; Nieminen, R. M.; Albe, K. *Journal of Applied Physics* **2010**, *108*, 053511.

[238] Ágoston, P.; Albe, K.; Nieminen, R. M.; Puska, M. J. *Physical Review Letters* **2009**, *103*, 245501.

[239] Scanlon, D. O.; Watson, G. W. *Journal of Materials Chemistry* **2012**, *22*, 25236–25245.

[240] Ganose, A. M.; Scanlon, D. O. *J. Mater. Chem. C* **2016**, *4*, 1467–1475.

[241] Vasheghani Farahani, S. K.; Veal, T. D.; Mudd, J. J.; Scanlon, D. O.; Watson, G. W.; Bierwagen, O.; White, M. E.; Speck, J. S.; McConville, C. F. *Physical Review B* **2014**, *90*, 155413.

[242] Lebens-Higgins, Z.; Scanlon, D. O.; Paik, H.; Sallis, S.; Nie, Y.; Uchida, M.; Quackenbush, N. F.; Wahila, M. J.; Sterbinsky, G. E.; Arena, D. A.; Woicik, J. C.; Schlom, D. G.; Piper, L. F. J. *Physical Review Letters* **2016**, *116*, 027602.

[243] Sallis, S.; Scanlon, D. O.; Chae, S. C.; Quackenbush, N. F.; Fischer, D. A.; Woicik, J. C.; Guo, J.-H.; Cheong, S. W.; Piper, L. F. J. *Applied Physics Letters* **2013**, *103*, 42105.

[244] Ellmer, K. *Journal of Physics D: Applied Physics* **2001**, *34*, 3097–3108.

[245] Look, D. C.; Leedy, K. D.; Vines, L.; Svensson, B. G.; Zubiaga, A.; Tuomisto, F.; Doutt, D. R.; Brillson, L. J. *Physical Review B* **2011**, *84*, 115202.

[246] Look, D. C.; Stutz, C. E.; Molnar, R. J.; Saarinen, K.; Liliental-Weber, Z. *Solid State Communications* **2001**, *117*, 571–575.

[247] Bardeen, J.; Shockley, W. *Physical Review* **1950**, *80*, 72–80.

[248] Low, F. E.; Pines, D. *Physical Review* **1955**, *98*, 414–418.

[249] Fröhlich, H.; Mott, N. F. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* **1939**, *171*, 496–504.

[250] Howarth, D. J.; Sondheimer, E. H. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* **1953**, *219*, 53–74.

[251] Bruneaux, J.; Cachet, H.; Froment, M.; Messad, A. *Thin Solid Films* **1991**, *197*, 129–142.

[252] Seto, J. Y. W. *Journal of Applied Physics* **1975**, *46*, 5247–5254.

[253] Button, K. J.; Fonstad, C. G.; Dreybrodt, W. *Physical Review B* **1971**, *4*, 4539–4542.

[254] Jarzebski, Z. M.; Morton, J. P. *Journal of The Electrochemical Society* **1976**, *123*, 333C–346C.

[255] Martínez, A. I.; Huerta, L.; de León, J. M. O.-R.; Acosta, D.; Malik, O.; Aguilar, M. *Journal of Physics D: Applied Physics* **2006**, *39*, 5091–5096.

[256] Haitjema, H.; Elich, J.; Hoogendoorn, C. *Solar Energy Materials* **1989**, *18*, 283–297.

[257] *Applied Physics Letters* **2010**, *96*, 062102.

[258] Rauch, C.; Tuomisto, F.; King, P. D. C.; Veal, T. D.; Lu, H.; Schaff, W. J. *Applied Physics Letters* **2012**, *101*, 011903.

[259] Makkonen, I.; Korhonen, E.; Prozheeva, V.; Tuomisto, F. *Journal of Physics: Condensed Matter* **2016**, *28*, 224002.

[260] Canestraro, C. D.; Roman, L. S.; Persson, C. *Thin Solid Films* **2009**, *517*, 6301–6304.

[261] Wahila, M. J. et al. *Chem. Mater.* **2016**, *28*, 4706–4713.

[262] Walsh, A.; Payne, D. J.; Egdell, R. G.; Watson, G. W. *Chemical Society Reviews* **2011**, *40*, 4455–4463.

[263] Ferraria, A. M.; Lopes da Silva, J. D.; Botelho do Rego, A. M. *Polymer* **2003**, *44*, 7241–7249.

[264] Sleigh, C.; Pijpers, A.; Jaspers, A.; Coussens, B.; Meier, R. J. *Journal of Electron Spectroscopy and Related Phenomena* **1996**, *77*, 41–57.

[265] Biesinger, M. C.; Payne, B. P.; Grosvenor, A. P.; Lau, L. W. M.; Gerson, A. R.; Smart, R. S. C. *Applied Surface Science* **2011**, *257*, 2717–2730.

[266] Payne, D. J.; Egdell, R. G.; Law, D. S. L.; Glans, P.-A.; Learmonth, T.; Smith, K. E.; Guo, J.; Walsh, A.; Watson, G. W. *J. Mater. Chem.* **2007**, *17*, 267–277.

[267] Cox, P. A.; Egdell, R. G.; Harding, C.; Patterson, W. R.; Tavener, P. J. *Surface Science* **1982**, *123*, 179–203.

[268] Veal, T. D.; McConville, C. F. *Physical Review B* **2001**, *64*, 085311.

[269] Vasheghani Farahani, S. K.; Veal, T. D.; King, P. D. C.; Zúñiga-Pérez, J.; Muñoz-Sanjosé, V.; McConville, C. F. *Journal of Applied Physics* **2011**, *109*, 073712.

[270] Taylor, J. A.; Lancaster, G. M.; Rabalais, J. W. *Journal of Electron Spectroscopy and Related Phenomena* **1978**, *13*, 435–444.

[271] Willemen, H.; Van De Vondel, D.; Van Der Kelen, G. *Inorganica Chimica Acta* **1979**, *34*, 175–180.

[272] Wu, S.; Yuan, S.; Shi, L.; Zhao, Y.; Fang, J. *Journal of Colloid and Interface Science* **2010**, *346*, 12–16.

[273] Nakao, S.; Yamada, N.; Hitosugi, T.; Hirose, Y.; Shimada, T.; Hasegawa, T. *Applied Physics Express* **2010**, *3*, 031102.

[274] Feneberg, M.; Lidig, C.; Lange, K.; White, M. E.; Tsai, M. Y.; Speck, J. S.; Bierwagen, O.; Goldhahn, R. *Physica Status Solidi (A) Applications and Materials Science* **2014**, *211*, 82–86.

[275] Kim, M.-G.; Kanatzidis, M. G.; Facchetti, A.; Marks, T. J. *Nature Materials* **2011**, *10*, 382–388.

[276] Ellmer, K. *Nature Photonics* **2012**, *6*, 808–816.

[277] Kumar, A.; Zhou, C. *ACS Nano* **2010**, *4*, 11–14.

[278] Zhang, L.; Zhou, Y.; Guo, L.; Zhao, W.; Barnes, A.; Zhang, H. T.; Eaton, C.; Zheng, Y.; Brahlek, M.; Haneef, H. F.; Podraza, N. J.; Chan, M. H.; Gopalan, V.; Rabe, K. M.; Engel-Herbert, R. *Nature Materials* **2016**, *15*, 204–210.

[279] Kim, Y. H.; Sachse, C.; Machala, M. L.; May, C.; Müller-Meskamp, L.; Leo, K. *Adv. Func. Mater.* **2011**, *21*, 1076–1081.

[280] Bonaccorso, F.; Hasan, Z. S. T.; Ferrari, A. C. *Nature Photon.* **2010**, *4*, 611–622.

[281] Fraser, D. B.; Cook, H. D. *J. Electrochem. Soc.* **1972**, *119*, 1368–1374.

[282] Barnes, T. M.; O., M.; Bergeson, J. D.; Larsen, B. A.; Blackburn, J. L.; Beard, M. C.; Bult, J.; van de Lagemaat, J. *Adv. Energ. Mat.* **2012**, *2*, 353–360.

[283] Szczyrbowski, J.; Dietrich, A.; Hartig, K. *Sol. Energ. Mat.* **1989**, *19*, 43–53.

[284] Kresse, G.; Furthmüller, J. *Computational Materials Science* **1996**, *6*, 15–50.

[285] Kresse, G.; Furthmüller, J. *Physical Review B* **1996**, *54*, 11169–11186.

[286] Krukau, A. V.; Vydrov, O. A.; Izmaylov, A. F.; Scuseria, G. E. *Journal of Chemical Physics* **2006**, *125*, 224106.

[287] Blöchl, P. E. *Physical Review B* **1994**, *50*, 17953–17979.

[288] Medeiros, P. V. C.; Stafström, S.; Björk, J. *Physical Review B* **2014**, *89*, 041407.

[289] Medeiros, P. V. C.; Tsirkin, S. S.; Stafström, S.; Björk, J. *Physical Review B* **2015**, *91*, 041116.

[290] Pickering, C. *Journal of Physics C: Solid State Physics* **1980**, *13*, 2959–2968.

[291] Parthiban, S.; Elangovan, E.; Ramamurthi, K.; Martins, R.; Fortunato, E. *Solar Energy Materials and Solar Cells* **2010**, *94*, 406–412.

[292] Meng, Y.; Yang, X.; Chen, H.; Shen, J.; Jiang, Y.; Zhang, Z.; Hua, Z. *Thin Solid Films* **2001**, *394*, 218–223.

[293] Parthiban, S.; Ramamurthi, K.; Elangovan, E.; Martins, R.; Fortunato, E. *Applied Physics Letters* **2009**, *94*, 212101.

[294] Yamada, N.; Yamada, M.; Toyama, H.; Ino, R.; Cao, X.; Yamaguchi, Y.; Ninomiya, Y. *Thin Solid Films* **2017**, *626*, 46–54.

[295] Meng, Y.; Yang, X.-l.; Chen, H.-x.; Shen, J.; Jiang, Y.-m.; Zhang, Z.-j.; Hua, Z.-y. *Journal of Vacuum Science & Technology A* **2002**, *20*, 288–290.

[296] Tahar, R. B. H.; Ban, T.; Ohya, Y.; Takahashi, Y. *Journal of Applied Physics* **1998**, *83*, 2631–2645.

[297] Groth, R. *physica status solidi* **1966**, *14*, 69–75.

[298] Mizuhashi, M. *Thin Solid Films* **1980**, *70*, 91–100.

[299] Clanget, R. *Applied Physics* **1973**, *2*, 247–256.

[300] Song, P. K.; Shigesato, Y.; Kamei, M.; Yasui, I. *Japanese Journal of Applied Physics* **1999**, *38*, 2921–2927.

[301] Shigesato, Y.; Takaki, S.; Haranoh, T. *Journal of Applied Physics* **1992**, *71*, 3356–3364.

[302] Bierwagen, O.; Speck, J. S. *Phys. Stat. Sol. (a)* **2014**, *211*, 48–53.

[303] Köstlin, H.; Jost, R.; Lems, W. *Physica Status Solidi (a)* **1975**, *29*, 87–93.

[304] Dobberstein, P.; Hampe, A.; Sauerbrey, G. *Physics Letters A* **1968**, *27*, 256–257.

[305] Zawadzki, W. *Advances in Physics* **1974**, *23*, 435–522.

[306] Swallow, J. E. N.; Williamson, B. A. D.; Whittles, T. J.; Birkett, M.; Featherstone, T. J.; Peng, N.; Abbott, A.; Farnworth, M.; Cheetham, K. J.; Warren, P.; Scanlon, D. O.; Dhanak, V. R.; Veal, T. D. *Advanced Functional Materials* **2018**, *28*, 1701900.

[307] Scanlon, D. O.; Watson, G. W.; Payne, D. J.; Atkinson, G. R.; Egdell, R. G.; Law, D. S. L. *The Journal of Physical Chemistry C* **2010**, *114*, 4636–4645.

[308] Sallis, S.; Scanlon, D. O.; Chae, S. C.; Quackenbush, N. F.; Fischer, D. A.; Woicik, J. C.; Guo, J.-H.; Cheong, S. W.; Piper, L. F. J. *Applied Physics Letters* **2013**, *103*, 42105.

[309] Goyal, A.; Gorai, P.; Peng, H.; Lany, S.; Stevanovic, V. *Computational Materials Science* **2017**, *130*, 1–9.

[310] Xu, J.; Liu, J.-B.; Liu, B.-X.; Li, S.-N.; Wei, S.-H.; Huang, B. *Advanced Electronic Materials* **2018**, *4*, 1700553.

[311] González, G. B. *Materials* **2012**, *5*, 818–850.

[312] Nadaud, N.; Lequeux, N.; Nanot, M.; Jové, J.; Roisnel, T. *Journal of Solid State Chemistry* **1998**, *135*, 140–148.

[313] Ueda, N.; Hosono, H.; Waseda, R.; Kawazoe, H. *Appl. Phys. Lett.* **1997**, *71*, 933.

[314] Rebien, M.; Henrion, W.; Hong, M.; Mannaerts, J. P.; Fleischer, M. *Appl. Phys. Lett.* **2002**, *81*, 250.

[315] Kokubun, Y.; Miura, K.; Endo, F.; Nakagomi, S. *Appl. Phys. Lett.* **2007**, *90*, 031912.

[316] Ogita, M.; Saika, N.; Nakanishi, Y.; Hatanaka, Y. *Appl. Surf. Sci.* **1999**, *142*, 188–191.

[317] Higashiwaki, M.; Sasaki, K.; Murakami, H.; Kumagai, Y.; Koukitu, A.; Kuramata, A.; Masui, T.; Yamakoshi, S. *Semicond. Sci. Technol.* **2016**, *31*, 034001.

[318] Vasheghani Farahani, S. K.; Veal, T. D.; Mudd, J. J.; Scanlon, D. O.; Watson, G. W.; Bierwagen, O.; White, M. E.; Speck, J. S.; McConville, C. F. *Phys. Rev. B* **2014**, *90*, 155413.

[319] Piper, L. F. J.; Colakerol, L.; King, P. D. C.; Schleife, A.; Zúñiga-Pérez, J.; Glans, P.-A.; Learmonth, T.; Federov, A.; Veal, T. D.; Fuchs, F.; Muñoz-Sanjosé, V.; Bechstedt, F.; McConville, C. F.; ; Smith, K. E. *Phys. Rev. B* **2008**, *78*, 165127.

[320] Allen, M. W.; Swartz, C. H.; Myers, T. H.; Veal, T. D.; McConville, C. F.; Durbin, S. M. *Phys. Rev. B* **2010**, *81*, 075211.

[321] King, P. D. C.; Veal, T. D.; Payne, D. J.; Bourlange, A.; Egdell, R. G.; McConville, C. F. *Phys. Rev. Lett.* **2008**, *101*, 116808.

[322] Mahboob, I.; Veal, T. D.; McConville, C. F.; Lu, H.; Schaff, W. J. *Phys. Rev. Lett.* **2004**, *92*, 036804.

[323] King, P. D. C.; Veal, T. D.; Jefferson, P. H.; Hatfield, S. A.; Piper, L. F. J.; McConville, C. F.; Fuchs, F.; Furthmüller, J.; Bechstedt, F.; Lu, H.; Schaff, W. J. *Phys. Rev. B* **2008**, *77*, 045316.

[324] Heiland, G.; Kunstmann, P. *Surf. Sci.* **1969**, *13*, 72–84.

[325] Heinhold, R.; Allen, M. W. *J. Mater. Res.* **2012**, *27*, 2214–2219.

[326] Heinhold, R.; Williams, G. T.; Cooil, S. P.; Evans, D. A.; Allen, M. W. *Phys. Rev. B* **2013**, *88*, 235315.

[327] Heinhold, R.; Cooil, S. P.; Evans, D. A.; Allen, M. W. *J. Phys. Chem. C* **2014**, *118*, 24575–24582.

[328] Batzill, M.; Diebold, U. *Prog. Surf. Sci.* **2005**, *79*, 47–154.

[329] Nagata, T.; Bierwagen, O.; White, M. E.; Tsai, M.-Y.; Speck, J. S. *J. Appl. Phys.* **2010**, *107*, 033707.

[330] Pearton, S. J.; Yang, J.; Cary, P. H.; F. Ren, J. K.; Tadjer, M. J.; Mastro, M. A. *Appl. Phys. rev.* **2018**, *5*, 011301.

[331] Navarro-Quezada, A.; Alamé, S.; Esser, N.; Furthmüller, J.; Bechstedt, F.; Galazka, Z.; Skuridina, D.; Vogt, P. *Phys. Rev. B* **2015**, *92*, 195306.

[332] Navarro-Quezada, A.; Galazka, Z.; Alamé, S.; Skuridina, D.; Vogt, P.; Esser, N. *Appl. Surf. Sci.* **2015**, *349*, 368–373.

[333] Lovejoy, T. C.; Chen, R.; Zheng, C.; Villora, E. G.; Shimamura, K.; Yoshikawa, H.; Tamashita, Y.; Ueda, S.; Kobayashi, K.; Dunham, S. T.; Ohuchi, F. S.; Olmstead, M. A. *Appl. Phys. Lett.* **2012**, *100*, 181602.

[334] Mohamed, M.; Unger, I.; Janowitz, C.; Manzke, R.; Galazka, Z.; Uecker, R.; Fornari, R. *J. Phys. Conf. Ser.* **2011**, *286*, 012027.

[335] Lovejoy, T. C.; Yitamben, E. N.; Shamir, N.; Morales, J.; Villora, E. G.; Shima-mura, K.; Zheng, S.; Ohuchi, F. S.; Olmstead, M. A. *Appl. Phys. Lett.* **2009**, *94*, 081906.

[336] Mohamed, M.; Janowitz, C.; Unger, I.; Manzke, R.; Galazka, Z.; Uecker, R.; Fornari, R.; Weber, J. R.; Varley, J. B.; Van de Walle, C. G. *Appl. Phys. Let.* **2010**, *97*, 211903.

[337] Piper, L. F. J.; Preston, A. R. H.; Fedorov, A.; Cho, S. W.; DeMasi, A.; Smith, K. E. *Phys. Rev. B* **2010**, *81*, 233305.

[338] King, P. D. C.; McKenzie, I.; Veal, T. D. *Appl. Phys. Lett.* **2010**, *96*, 062110.

[339] Varley, J. B.; Peelaers, H.; Janotti, A.; Van de Walle, C. G. *J. Phys.: Condens. Matter* **2011**, *23*, 334212.

[340] Li, H.; Robertson, J. *J. Appl. Phys.* **2014**, *115*, 203708.

[341] Fleischer, M.; Giber, J.; Meixner, H. *Appl. Phys. A* **1992**, *54*, 560–566.

[342] Geistlinger, H. *Sensors and Actuators B* **1994**, *18*, 125–131.

[343] Whittles, T. J.; Burton, L. A.; Skelton, J. M.; Walsh, A.; Veal, T. D.; Dhanak, V. R. *Chem. Mater.* **2016**, *28*, 3718–3726.

[344] Ahn, S.; Ren, F.; Patrick, E.; Law, M. E.; Pearton, S. J. *ECS J. Solid state Sci. Technol.* **2017**, *6*, Q3026–Q3029.

[345] Ahn, S.; Ren, F.; Patrick, E.; Law, M. E.; Pearton, S. J.; Kuramata, A. *Appl. Phys. Lett.* **2016**, *109*, 242108.

[346] Heyd, J.; Scuseria, G. E.; Ernzerhof, M. *J. Chem. Phys* **2003**, *118*, 8207.

[347] Heyd, J.; Scuseria, G. E.; Ernzerhof, M. *J. Chem. Phys.* **2006**, *124*, 219906.

[348] Blöchl, P. E. *Phys. Rev. B* **1994**, *50*, 17953.

[349] Peelaers, H.; Varley, J. B.; Speck, J. S.; Van de Walle, C. G. *Appl. Phys. Lett.* **2018**, *112*, 242101.

[350] Rafique, S.; Han, L.; Mou, S.; Zhao, H. *Opt. Mater. Express* **2017**, *7*, 3561.

[351] Freysoldt, C.; Neugebauer, J.; Van de Walle, C. G. *Phys. Rev. Lett.* **2009**, *102*, 016402.

[352] Ingebrigtsen, M. E.; Varley, J. B.; Kuznetsov, A. Y.; Svensson, B. G.; Alfieri, G.; Mihaila, A.; Badstübner, U.; Vines, L. *Appl. Phys. Lett.* **2018**, *112*, 042104.

[353] Kraut, E. A.; Grant, R. W.; Waldrop, J. R.; Kowalczyk, S. P. *Phys. Rev. Lett.* **1980**, *44*, 1620.

[354] Fu, H.; Chen, H.; Huang, X.; Baranowski, I.; Montes, J.; Yang, T.-H.; Zhao, Y. *IEEE Trans. Electron Devices* **2018**, *65*, 3507–3513.

[355] Sun, S.-M.; Liu, W.-J.; Wang, Y.-P.; Huan, Y.-W.; Ma, Q.; Zhu, B.; Wu, S.-D.; Yu, W.-J.; Horng, R.-H.; Xia, C.-T.; Sun, Q.-Q.; Ding, S.-J.; Zhang, D. W. *Appl. Phys. Lett.* **2018**, *113*, 031603.

[356] Sturm, C.; Schmidt-Grund, R.; Kranert, C.; Furthmuller, J.; Bechstedt, F.; Grundmann, M. *Phys. Rev. B* **2016**, *94*, 035148.

[357] Edwards, P. P.; Sienko, M. J. *J. Am. Chem. Soc* **1981**, *103*, 2967–2971.

[358] Peelaers, H.; Van de Walle, C. G. *Appl. Phys. Lett.* **2017**, *111*, 182104.

[359] Mock, A.; Korlacki, R.; Briley, C.; Darakchieva, V.; Monemar, B.; Kumagai, Y.; Goto, K.; Higashiwaki, M.; Schubert, M. *Phys. Rev. B* **2017**, *96*, 245205.

[360] Knight, S.; Mock, A.; Korlacki, R.; Darakchieva, V.; Monemar, B.; Kumagai, Y.; Goto, K.; Higashiwaki, M.; Schubert, M. *Appl. Phys. Lett.* **2018**, *112*, 012103.

[361] Paasch, G.; Übensee, H. *Phys. Status Solidi B* **1982**, *113*, 165–178.

[362] Zöllner, J.-P.; Übensee, H.; Paasch, G.; Fiedler, T.; Gobsch, G. *Phys. Status Solidi B* **1986**, *134*, 837–845.

[363] Brudnyi, V. N.; Grinyaev, S. N.; Kolin, N. G. *Semiconductors* **2003**, *37*, 537–545.

[364] Falabretti, B.; Robertson, J. *J. Appl. Phys.* **2007**, *102*, 123703.

[365] Limpijumnong, S.; Reunchan, P.; Janotti, A.; Van de Walle, C. G. *Phys. Rev. B* **2009**, *80*, 193202.

[366] Amini, M. N.; Saniz, R.; Lamoen, D.; Partoens, B. *J. Appl. Phys.* **2011**, *110*, 063521.

[367] Varley, J. B.; Weber, J. R.; Janotti, A.; Van de Walle, C. G. *Appl. Phys. Lett.* **2010**, *97*, 142106.

[368] Al-Kuhaili, M. F.; Durrani, S. M. A.; Khawaja, E. E. *Appl. Phys. Lett.* **2003**, *83*, 4533.

[369] Lay, T.; Liao, Y.; Hung, W.; Hong, M.; Kwo, J.; Mannaerts, J. *J. Cryst. Growth* **2005**, *278*, 624–628.

[370] Ramana, C. V.; Rubio, E. J.; Barraza, C. D.; Gallardo, A. M.; McPeak, S.; Kotru, S.; Grant, J. T. *J. Appl. Phys.* **2014**, *115*, 043508.

[371] Rubio, E. J.; Mates, T. E.; Manandhar, S.; Nandasiri, M.; Shutthanandan, V.; Ramana, C. V. *J. Phys. Chem. C* **2016**, *120*, 26720–26735.

[372] Sheng, J.; Park, E. J.; Shong, B.; Park, J.-S. *ACS Appl. Mater. Interfaces* **2017**, *9*, 23934–23940.

[373] Walukiewicz, W. *Phys. Rev. B* **1988**, *37*, 4760.

[374] Walukiewicz, W. *Appl. Phys. Lett.* **1989**, *54*, 2094.

[375] Tersoff, J. *Phys. Rev. Lett.* **1984**, *52*, 465.

[376] von Wenckstern, H. *Adv. Electron. Mater.* **2017**, *3*, 1600350.

[377] Vines, L.; Bhoodoo, C.; von Wenckstern, H.; Grundmann, M. *J. Phys.: Condens. Matter* **2018**, *30*, 025502.

[378] Aroyo, M. I.; Kirov, A.; Capillas, C.; Perez-Mato, J. M.; Wondratschek, H. *Acta Cryst.* **2006**, *A62*, 115–128.

[379] Basu, P. K. *Theory of Optical Processes in Semiconductors: Bulk and Microstructures*; Oxford University Press, 2003.

[380] Birkett, M. Optical properties of earth-abundant semiconductors for renewable energy. Ph.D. thesis, The University of Liverpool, 2016.

[381] Brooker, G. *Modern Classical Optics*; Oxford University Press, 2003.

[382] Van Hove, L. *Phys. Rev.* **1953**, *89*, 1189–1193.

[383] Rode, D. L. In *Low-field electron transport*; Willardson, R. K., Beer, A. C., Eds.; Semiconductors and Semimetals, Vol. 10; Academic, New York, 1975; pp 1–89.

[384] King, P. D. C. Charge neutrality level in significantly cation-anion mismatched semiconductors. Ph.D. thesis, The University of Warwick, 2009.

[385] Gasiorowicz, S. *Quantum physics, 3rd edition*; Wiley, New York, 2003.

[386] Tersoff, J. *Phys. Rev. B* **1984**, *30*, 4874–4877.

[387] Tersoff, J. *Surface Science* **1986**, *168*, 275 – 284.

[388] Tersoff, J. *Journal of Vacuum Science & Technology B: Microelectronics Processing and Phenomena* **1986**, *4*, 1066–1067.

[389] Slater, J. C. *Phys. Rev.* **1953**, *92*, 603–608.

[390] Cardona, M.; Christensen, N. E. *Phys. Rev. B* **1987**, *35*, 6182–6194.

[391] Mönch, W. *Journal of Applied Physics* **1996**, *80*, 5076–5082.

[392] Baldereschi, A. *Phys. Rev. B* **1973**, *7*, 5212–5215.

[393] Penn, D. R. *Phys. Rev.* **1962**, *128*, 2093–2097.