First-Principles Calculations of Auger Recombination in Optoelectronic Materials

by

Andrew McAllister

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Applied Physics)
in the University of Michigan
2019

Doctoral Committee:

Associate Professor Emmanouil Kioupakis, Chair
Professor Cagliyan Kurdak
Professor Zetian Mi
Associate Professor Donald Siegel

Andrew McAllister
mcala@umich.edu
ORCID: 0000-0001-9842-4159

For mom, who never said I had picked out too many books at the library.
For dad, who taught me about computers from before I can remember.

Acknowledgments

> Let me not be so vain to think that I am the sole author of my victories and a victim of my defeats.
>
> *An Invocation For Beginnings*
> *Ze Frank*

I have many, *many* people to thank for helping me be able to get my PhD. This list is certainly not exhaustive, but remember that I did not finish this without support that started when I was very small and continues to this day.

Thank you to my parents, who bought me as many books as I wanted, took me to the library often, let me explore the things I was interested in, drove me places, gave me food, put a roof over my head and reminded me that I *did not* work for the electric company. If only we had LED bulbs 20 years ago. I could write an entire dissertation on the many ways you have given me love and support over the years. Thank you for everything.

Thank you to my brother Stephen, for being my constant companion growing up and now. From smashing Lego cars to playing Rock Band to reading job applications. Thank you for always being there for me.

My family as a whole has been so supportive and encouraging throughout my life. Even when they don't understand exactly what I'm doing. Thank you to Steve, Steph, Henry, Ed, Christine, John, Barbara[2], Kenny, Sandy, Brooke, Sean, Sandra, Alex, Sam, Scott, Ben and Abby. You all are a part of what makes coming home special. I'd also like to thank Gene, Archie, Andrew, Stella, Geraldine, and Tom who won't be able to see my graduate, but I know would have been proud.

Thank you to Liz, who got me the *The Big Book of Tell Me Why* (all three volumes!) and for always being shorter than me. Let me know if you need another set of eyes for your own thesis.

Thank you to all the teachers that that I have had – both those that made the deep impacts I can enumerate here and those whose impacts I have forgotten or didn't even notice.

I'm thankful for the many mentors I've had over the years. Including Peter Kramer, Mark Holmes and Brad Lister for first letting me get my feet wet with research in the Math department at RPI. I appreciate the guidance you gave me even if I did wind up going to physics. For Peter Persans who helped me throughout my time at RPI. You gave me good advice, let me do research in your lab, have written (many) letters of recommendation and generally have been an amazing source of knowledge and help. And of course, my advisor Emmanouil Kioupakis who has been amazingly helpful throughout this process. You've taught me much about physics, doing research, and being a scientist, but you also let me explore things that I was interested in on my own. Without that freedom graduate school would not have been as enjoyable. I also really appreciate that you didn't pick out a big snake to fight during my defense.

For my colleagues in the Kioupakis research group. You've made time in the office and at conferences not just great for expanding my knowledge but also fun. From Dylan and Guangsha who were there to teach me the basics when I joined to to Kyle and Olivia who I've only know for a few months. Thank you to everyone: Dylan, Guangsha, Jihang, Alex, Logan, Christina, Kelsey, Nocona, Sahil, Kevin, Mike, Zihao, Suvadip, Sieun, Woncheol, Kyle and Olivia. It's amazing to be able to go to work and know that you'll get to see a bunch of friends throughout the day.

Thank you to the Applied Physics program for being my family while so far away from home. For Cagliyan, Cindy, and Lauren who have helped me out throughout my time here. For all my fellow graduate students, especially my 2012 cohort, thanks for the board games, the D&D, and the late night homework sessions.

Thank you to my science communication friends here at Michigan. You've helped me gain so many skills and figure out better ways to explain the complicated world of research. Thank you in particular to the RELATE program (and Elyse, Katie, Brandon and Brian), Michigan Science Writers and the team that put on ComSciCon Michigan. All of these programs (and more) have helped me figure out what I'll be doing now that my PhD is over.

Thank you to my many friends for sticking with me over the past 6 years. In particular Hanh and Diana who have been friends for *so* long and constantly challenge me to think differently. Thank you for being so thoughtful and supportive. Thank you to Padula who just a few months ago reminded me of how long I have been working on getting my PhD and how exciting it was to finally be there. Thank you to Adam, Dave, and Sara for Birthday Smash weekends and many other things. And of course thanks to Ramon, my roommate and friend throughout my PhD.

I can't forget to thank my partner and best friend Chrissy. I've been lucky to have you by my side now for three years and I'm looking forward to many more. Your support and love has made this whole thing (and especially the last few months) much less stressful. Not only have I found a best friend and partner with you, but I've been welcomed by your family who have also been supportive in such a short time.

While this section is long it is also incomplete. I cannot include everyone who has helped me over the years. Even if I tried I'd be bound to forget someone. So again, thank you to everyone who has been a part of supporting me throughout my PhD or my life, either in big or small ways.

Professionally, I also need to thank many funding sources for helping me throughout my PhD. The National Science Foundation (Graduate Research Fellowship (DGE-1256260), CAREER (DMR-1254314) and DMREF (1534221)) funded most of my work. The research on radiation detection was funded by the National Nuclear Security Administration Office of Nonproliferation and Verification (NA-22).

In addition, I must thank the National Energy Research and Computing Center (NERSC), University of Michigan Advanced Research Computing (ARC-TS), Lawrence Livermore National

Laboratory (LLNL), and the Extreme Science and Engineering Discovery Environment (XSEDE) for providing me with access to high-performance computing resources. Without these supercomputers my research would not be possible. I also want to thank the support staff for these various centers for helping me when issues arose.

Finally, I thank you to the developers of Quantum Espresso, BerkeleyGW and Wannier90. These codes, made freely available, formed the foundation of this work and wouldn't be possible without it.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

DFT density functional theory

LDA local density approximation

GGA generalized gradient approximation

SQS special quasi-random structure

IQE internal quantum efficiency

NSCF non-self consistent field

BZ Brillouin zone

GW Somewhat confusingly, this is **not** an acronym or initialism. G represents use of **G**reen's Functions and **W** is the symbol often used for the screened Coulomb Interaction.

LDA+U local density approximation with a +U Correction

DFT+U density functional theory (any functional) with a +U correction

IR infrared

CPU central processing unit

CFL compact fluorescent

LED light-emitting diode

lmW^{-1} lumen per watt

GaN gallium nitride

AlN aluminum nitride

InN indium nitride

CRI color rendering index

UV ultraviolet

IR infrared

MPI message-passing interface

GPU graphics processing unit

ABSTRACT

Lighting currently uses around 15% of global energy expenditure. Reducing this energy usage would be an important part of reducing the effects of global climate change. LEDs (light-emitting diodes) could be a more efficient light source than traditional incandescent or compact-fluorescent bulbs. However, they currently suffer from 'efficiency droop' — as the power through the device is increased, the efficiency goes down.

While the exact cause of droop is not known completely, one effect is the Auger recombination process. Auger recombination involves an electron and hole recombination which transfers energy to another carrier (rather than emitting a photon as in the desired radiative process). Auger recombination comes in many forms and can be assisted by phonons and alloys which allow Auger to occur when it would otherwise be prevented by momentum conservation. Reducing the Auger recombination process in LED materials would allow for manufacturers to use fewer LEDs in their bulbs, reducing costs and encouraging widespread LED adoption.

The challenge to reducing Auger is quantifying its impact. Experimental determination of Auger can be confounded by many factors and reported rates for Auger often vary over orders of magnitude. Theoretical determination of Auger recombination is also difficult because of the complexity of

the equations involved. In this situation, computation can be an important tool to understand the underlying physics. In particular, this thesis focuses on first-principles calculations, which solve Schödinger's equation directly rather than relying on existing experimental data. This allows these calculations to be predictive of experiment and can guide future research on more efficient materials.

While many open-source and commercial options exist for codes that will solve Schödinger's equation using Density Functional Theory, there is not an available code to solve for recombination rates. As part of this thesis, code used to find the Auger recombination rates was expanded and improved. This was then used to study various optoelectronic materials.

The group-III nitrides are widely used for making LEDs. GaN is popular in making blue LEDs, but other materials in this family have uses. InN has a band gap in the infrared, and could be used for telecommunication purposes. We studied the Auger process in InN and found that it is dominated by the direct Auger process, as expected for its small band gap. We also found that at high carrier densities, Auger was primarily reduced by carrier screening rather than phase-space filling. On the other hand, AlGaN alloys could be used to create UV LEDs, with applications in sterilization and sensing. We studied Auger in AlGaN alloys of three compositions and modeled expanding these Auger values throughout the entire alloy spectrum. We expected to find the maximum alloy-assisted Auger at the 50/50 alloy, but found the

opposite trend. This unintuitive result warrants further study.

Scintillators are another type of device that emits light when struck by radiation. Auger affects these devices by suppressing light output and making it difficult to identify what the elemental source of the original radiation was. We studied Auger in NaI and found that the phonon-assisted process dominates as expected by the large gap of the material.

This methodology and code has shed light on nonradiative carrier combination in optoelectronic materials and devices and can continue to be used in future studies

CHAPTER 1

Introduction

Specialized materials are all around us. From the steel that keeps up our to the silicon parts that make up our smartphones. A new field called materials science was formed from the need to understand and fine-tune the properties of materials. These properties vary widely and range from structural (how a material will bend or break when put under different forces), optoelectronic (how a material interacts with light), and chemical (what a material will do or how it will change when in the presence of a specific element) as only a few examples. The materials also vary widely, from the crystalline silicon in central processing units (CPUs) to the long, disorganized polymers in plastics.

Our understanding of semiconducting materials helped change computers from massive, specialized devices to small, general tools that everyone can use in less than a century. What once was a warehouse sized machine is now an easy-to-use device in nearly everyone's pocket. As computers have become more widespread in our day-to-day lives, they also have become an important tool for doing scientific research. Computational methods allow us to study a phenomenon even if the theoretical equations are too hard to solve by hand or when the phenomenon is too difficult to measure precisely. One way of using computers is to sift through or fit experimental data. While this is useful, we need to already have the data to do this. Instead, the methods used in this thesis are called *first-principles* methods and rely on solving the fundamental equations of the system that we are interested in. These methods allow us to *predict* experimental results and save time and money by only investigating the most promising areas.

The main method used in this thesis has been applied to an extraordinarily diverse array of applications and is called density functional theory (DFT). It has been used to understand the magnetic properties of complicated systems, how to speed up chemical reactions, and to design new drugs (to name just a few areas). DFT has become an incredible tool across many different fields. While DFT has limitations, there are many possibilities for what to study, but lighting has been the focus of most of this thesis.

1.1 Lighting

Illumination is something that is frequently taken for granted.[1,2] It has been over 100 years since electric light was first used for illumination and 50-80[1] years of "normal" homes having a light bulb in every room. It is assumed that we will have the light we need after the sun goes down or in a room without windows. But for most of human history, this was not the case. Fire, animal fat, and kerosene are some of the early ways we lit up the world. But these methods were unwieldy, unsafe and extremely expensive[2]. Moving to electric lights decreased the cost of light by a factor of 500,000[3,4] and has lead to major advances in human productivity and quality of life. But it has also been a major part of our increased energy usage. Around 15% of global energy expenditure is on lighting[3]. From our bedrooms to our 100,000 person football stadiums to our parking lots – we use an enormous amount of energy to light things up. To combat global climate change reducing our overall energy use is key. So it is worthwhile to consider what technologies we use to do all this lighting and how they may be improved.

Modern light bulbs come in three varieties fig. 1.1. The incandescent light bulb has been used for over 100 years but it is still incredibly inefficient. Modern incandescents transform only 5-7% of electricity into visible light, the rest being wasted as heat or light that our eyes cannot see. Fluorescent and eventually compact fluorescent (CFL) bulbs started to become popular in the early 1990s. While more efficient that incandescents, they have a maximum efficiency in the 20% range – still a huge portion of energy being unused. In addition, CFLs contain mercury, which is a hazard for both people and the environment.

Light-emitting diodes (LEDs) do not have the same problems. Their maximum luminous efficiency is much higher (~40%)[7]and they do not contain toxic materials. They also last much longer, meaning that the materials and energy put into making them in the first place has less of an impact on the environment (not to mention making the minor inconvenience of changing a light bulb much rarer). But even though LEDs are becoming more commonplace (fig. 1.2) they are still incredibly expensive in comparison to other options. Part of that cost is because LEDs have a long way to go before reaching that maximum theoretical efficiency.

[1]Depending on whether you lived in rural or urban areas.

[2]Think about how much wood you would have to chop to light up *one* room of your house for a day. How dim and smoky the fire would be compared to a light bulb. How not being careful could lead to your house burning down.

[3]Calculated by using data from the International Energy Agency statistics from 2016/2017.[5,6]

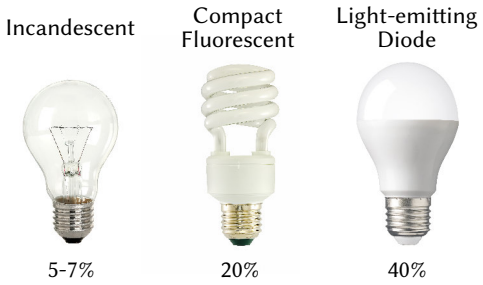

Figure 1.1 Household lightbulbs that you may have seen over the years. Efficiency numbers are based on *luminous efficiency* which is based on how the human eye percieves color. Our eyes are most responsive to green light, and the maximum efficiency (100%) is defined as ~685 lumen per watt (lmW^{-1}), which is equivalent to one watt of 555nm (green) light. Typically incandescents emit about 15–20 lmW^{-1}, CFLs emit around 60–90 lmW^{-1}, while modern LEDs emit around 150 lmW^{-1}, with theoretical efficiencies of 250-300 lmW^{-1}.[7]

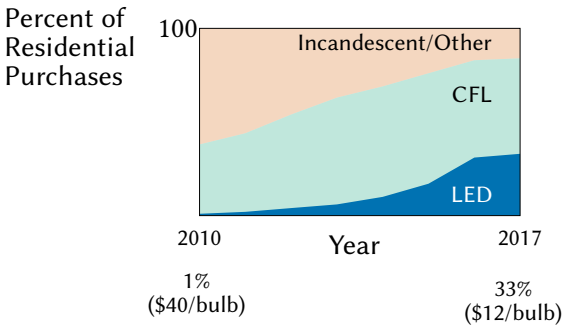

Figure 1.2 LED bulbs are being quickly adopted in for personal use, moving from 1% in 2010 to 33% in 2017. As the price for LEDs continues to decline, we can expect this trend to continue. (Figure made using data from International Energy Agency [8] and LED Inside [9].)

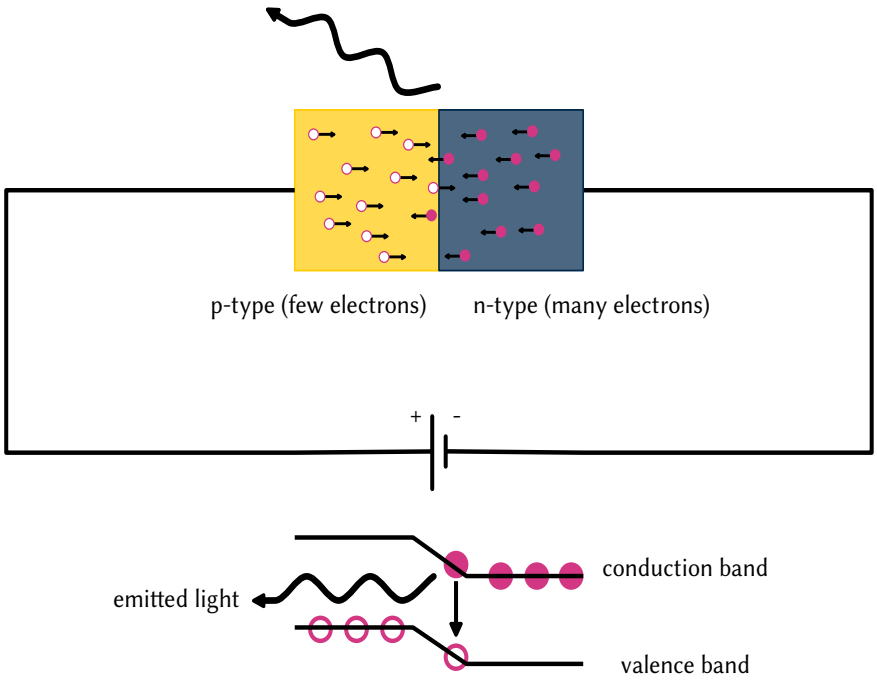

Figure 1.3 Diagram of how an LED works. Applying a potential difference (connecting to a battery for example) will push electrons from one side of the device to the other causing them to lose energy in the process. This energy is emitted in the form of light.

1.2 What are Light-emitting Diodes (LEDs)?

Light-emitting diodes are semiconducting devices that convert electrons into light. This is accomplished by stacking three materials together. In the middle is the "active region" light is emitted. Sandwiching the active region are two materials, one with excess electrons and one with very few electrons. This creates a device called a *p-n junction*. When an electrical current is applied in the right way, electrons move from the crowded region to the empty region, losing energy in the active region (fig. 1.3). How much energy the electrons lose is determined by a property of the material used in the active region called the *band gap* (E_g. The band gap determines what wave length of light comes out of the LED. When we talk about tuning the wave length of light in an LED, this is the property that we are trying to change. Of course, the other layers of materials matter and not just any combination will create an LED. Most importantly, these materials must be similar to each other so that they can be grown[4] together in one continuous block. If the materials are too different, the whole device won't work.

[4]If you have ever made rock candy you have "grown" sugar crystals. Growing these kinds of materials is much more complicated, but the idea of slowly atoms slowly bonding together in a structured way is the same.

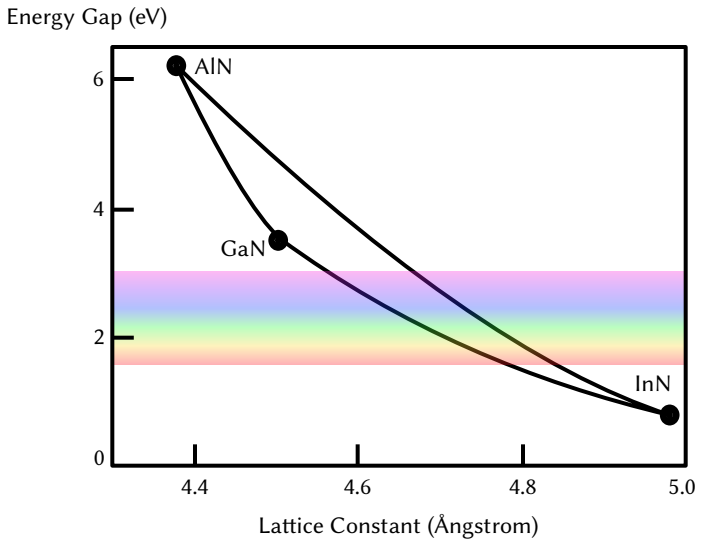

Figure 1.4 The band gaps of the group-III nitrides spans a large ranges of wave lengths. By alloying these materials together we can make LEDs of any visible color and even some that we cannot see.[13]

The first LED was made of a material called silicon carbide in 1928.[10] These first devices were not well understood. As time went on, a better understanding of semiconductors allowed for controlling what kind of light was emitted. Early devices starting in the 1960's were warmer colors[11] and were were very dim and only useful as small indicator lights. Breakthroughs in a material called gallium nitride (GaN) in the late 1980's and early 1990's allowed the creation of efficient blue LEDs. This was such an important discovery that the 2014 Nobel Prize in physics was awarded to Isamu Akasaki, Hiroshi Amano, and Shuji Nakamura[12] for figuring out how blue light could be generated.

A major part of why creating blue LEDs was so important was because we can make white light from blue LEDs(see the next section). But the materials that the Nobel prize winners studied could also be used to make LEDs that emit a large range of different wave lengths. These materials are part of the *group-III nitrides*, which consists of aluminum nitride (AlN), GaN, and indium nitride (InN). Because they are all of this same "family" of materials, it's pretty easy to create *alloys* of these materials. Alloys are combinations – rather than *pure* GaN, we can mix it with InN to get properties that are in-between the two. The band gaps (that determine the wave length of light emitted) have a big range (fig. 1.4) meaning that by combining these materials, we can make LEDs that are red, blue, green or even of wavelengths that we can't see – like ultraviolet (UV) or infrared (IR) light. All of this makes the group-III nitrides the most important LED material today.

1.2.1 Creating White Light Using LEDs

None of these colors are "white" light though – how do we make the LEDs which light up my living room? Creating white light using LEDs relies on how our eyes perceive light. Normal human eyes have structures called *cones* which are sensitive to red, blue, and green wave lengths of light. Various combinations of these three cone types are interpreted by our brain as different colors. Red and blue light together will appear purple. While red and green together light will appear yellow. When red, blue and green light all appear together we perceive white light (see fig. 1.5).

Knowing this you may already have one idea for creating a white LED – create a device with red, blue and green LEDs and turn them all on at once. This is one possibility and is certain how computer monitors, televisions and other screens work[5]. These kinds of devices are limited by the efficiency of the least efficient LED. So we must be able to make highly efficient devices for all three colors, which turns out to be a challenge[6].

Another way of creating white light is to use one color and convert it into the other needed colors using *phosphor* coatings. These coatings allow some of the original light through, but also change some of the light into another color. The most popular devices (and the ones you likely buy today) use blue LED and a phosphor coating which changes blue light in to yellow light. The blue and yellow appear white to our eyes.

There are disadvantages with this method. First, this light can appear unappealing to the human eye. It may appear "too blue" and objects illuminated by this light may not appear the color they are when they are outside in natural sunlight[7]. Second, the phosphor conversion is not 100% efficient, and lights relying on this technology suffer from various losses related to light being reflected back at the LED and the blue-to yellow conversion process itself only being 80% efficient.[15]

1.3 The Efficiency Droop Problem

With a handle on how LEDs work and how we can use them to make white light, we can talk about the focus of this thesis, the efficiency droop problem. When you buy an LED bulb today, you aren't buying *one* LED. Inside of the frosted over glass are many LEDs, which all have to work together to light up your room. This involves many engineering challenges – focusing the light

[5] Although those LEDs are organic LEDs and a different type than the ones studied in this thesis.

[6] The "Green-Gap" problem is one of the two major challenges facing LEDs today. We can make efficient red and blue LEDs, but not green. The other major challenge is the focus of this thesis and discussed in sec. 1.3.

[7] This is known as the color rendering index (CRI). A perfect CRI is 100 and represents sunlight and incandescent bulbs. Lower CRI values are common for outdoor lighting. Values above 80 are reasonable for indoor lighting, and LEDs range from 60-95 depending on the quality of the device.[14]

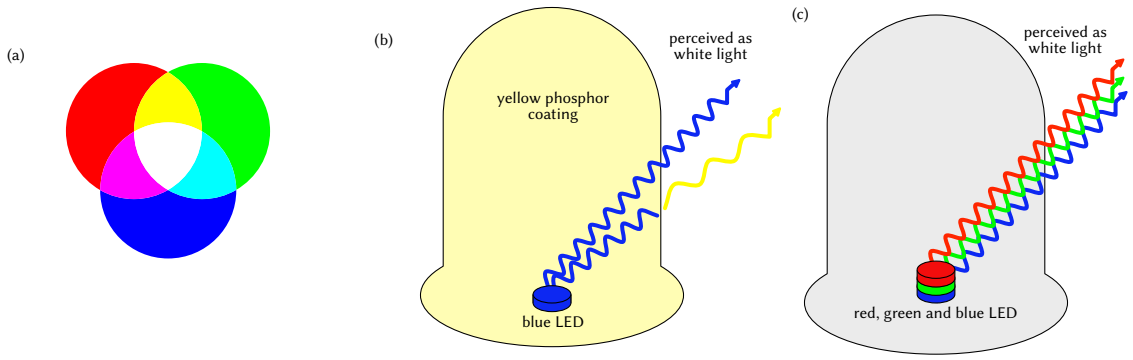

Figure 1.5 How we create white light using LEDs. (a) shows how additive color mixing works. (b) shows how most commercial LEDs work by using yellow phosphor coatings to convert some of the blue light to yellow. (c) shows how a tri-color LED would create white light without the need for phosphor coatings.

so that it is uniform and appealing in all directions. Plus, LEDs do generate heat, and the heat from multiple LEDs needs to be dissipated effectively otherwise the device won't work as well. These challenges increase the cost of the lightbulb. With only one LED there would be fewer materials needed and less engineering design to deal with focusing light and dealing with heat.

But why are there so many LEDs inside of the lightbulb? Why *not* just one? It turns out that when we try to put more electric current through an LED (which makes it brighter) we *also* reduce it's efficiency. So when making an LED bulb we have a choice to make. Put a few LEDs into the bulb but make them very bright. This would be cheap and it wouldn't be very efficient. But we already have cheap inefficient light bulbs! Instead, we can put many LEDs into the bulb that are all very dim. But by focusing these lights in the right way, they can together produce enough light for an entire room.

This tradeoff between brightness and electrical current is called the *efficiency droop* problem. fig. 1.6 shows the problem more visually. The reasons for droop aren't completely understood and could be due to multiple factors. There is still some debate over defects in the LED or carrier leakage causing the issue.[16] But a major contributor is called the *Auger process*, which is the focus of this thesis.

The Auger process occurs when many electrons are in a material – so many that they frequently "bump into" each other. This affects how the carriers move through the material and, most importantly, reduces the efficiency of devices that we expect to emit light. You can think about *Auger* like trying to get out of a crowded place. The more people there are, the harder it is to get out without jostling some other folks on your way. Electrons have the same problem.

Early experimental work by Shen et al. [17] brought attention to Auger as a cause for droop.

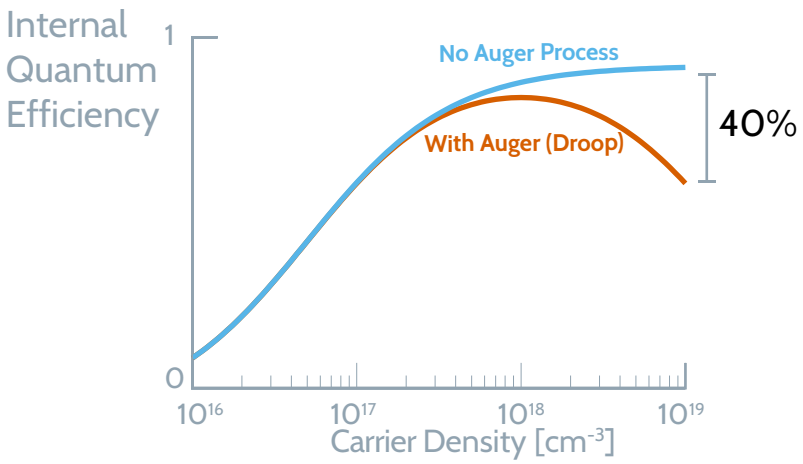

Figure 1.6 The efficiency droop problem, plotted here for InGaN using numbers experimental measurements from Shen et al. [17].

However, Auger is difficult to measure directly through experiment, and frequently there is a wide range in experimental values for the Auger process. Theoretical work showed much smaller Auger coefficients than those shown by experiment. This discrepancy was resolved by Kioupakis et al. [18]'s computational work that showed that Auger could take place through phonons or alloy scattering enhancing the Auger process.

1.4 Other Devices

While much of the above discussion has been about LEDs to light up a room, there are LEDs which can emit light that humans cannot see. UV LEDs are a subject of major research[19] and have uses in sterilization and sensing. On the opposite end of the electromagnetic spectrum are IR LEDs. Devices emitting at these wave lengths (in particular at 1550 nm) would be useful for telecommunication.

Other devices can emit light besides LEDs. Lasers might be the first to come to mind, and there is plenty of overlap in applications between lasers and LEDs. Another type of device that is not as well known is a scintillator. These devices emit light after being exposed to high-energy radiation. This be useful for non-proliferation, medicine, and basic science. These devices suffer from Auger recombination as well, when they are excited they create very high carrier densities which frequently undergo Auger recombination causing them to "keep" some of the high-energy radiation. This kept energy is not detected and affects how the material works as a detector.

References

[1] Stacey Vanek Smith and Cardiff Garcia. *Messy Desks, Light Bulbs & Dune*. 2018. URL: https://www.npr.org/sections/money/2018/08/13/638293909/messy-desks-light-bulbs-dune.

[2] David Kestenbaum and Jacob Goldstein. *The History Of Light, In 6 Minutes And 47 Seconds*. 2014. URL: https://www.npr.org/2014/05/02/309040279/in-4-000-years-one-thing-hasnt-changed-it-takes-time-to-buy-light..

[3] Tim Harford. "Epilogue: The Lightbulb." In: *50 Inventions that Shaped the Modern Economy*. Penguin, 2017.

[4] William D Nordhaus. "Do real-output and real-wage measures capture reality? The history of lighting suggests not." In: *The economics of new goods*. University of Chicago Press, 1996, pp. 27–70. URL: http://j-bradford-delong.net/Teaching_Folder/Econ_210c_spring_2002/Readings/Nordhaus_light.pdf.

[5] International Energy Agency. *Energy Efficiency: Lighting*. 2018. URL: https://www.iea.org/topics/energyefficiency/buildings/lighting/.

[6] International Energy Agency. *World Energy Balance 2018*. 2018. URL: https://www.iea.org/statistics/?country=WORLD%5C&year=2016%5C&category=Energy%5C%20consumption%5C&indicator=TFCbySource%5C&mode=chart%5C&categoryBrowse=false%5C&dataTable=BALANCES%5C&showDataTable=true.

[7] Thomas W. Murphy. "Maximum spectral luminous efficacy of white light." In: *Journal of Applied Physics* 111.10 (2012), p. 104909. DOI: 10.1063/1.4721897.

[8] International Energy Agency. *Lighting: Tracking Clean Energy Progress*. 2018. URL: https://www.iea.org/tcep/buildings/lighting/.

[9] LED Inside. *LEDinside: Prices of LED Light Bulbs for 40W Incandescent Light Bulbs Replacement Hit a Low of USD 18*. 2011. URL: https://www.ledinside.com/pricequotes/2011/8/led_bulb_price_201107.

[10] O.V. Lossev. "Luminous carborundum detector and detection effect and oscillations with crystals." In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 6.39 (1928), pp. 1024–1044. DOI: 10.1080/14786441108564683.

[11] Nick Holonyak and S. F. Bevacqua. "COHERENT (VISIBLE) LIGHT EMISSION FROM Ga(As1-xPx) JUNCTIONS." In: *Applied Physics Letters* 1.4 (1962), pp. 82–83. DOI: 10.1063/1.1753706.

[12] NobelPrize.org. *The Nobel Prize in Physics 2014*. 2014. URL: https://www.nobelprize.org/prizes/physics/2014/summary/.

[13] I. Vurgaftman and J. R. Meyer. "Band parameters for nitrogen-containing semiconductors." In: *Journal of Applied Physics* 94.6 (2003), pp. 3675–3696. DOI: 10.1063/1.1600519.

[14] Siddha Pimputkar et al. "Prospects for LED lighting." In: *Nature photonics* 3.4 (2009), p. 180. DOI: 10.1038/nphoton.2009.32.

[15] Steven C. Allen and Andrew J. Steckl. "A nearly ideal phosphor-converted white light-emitting diode." In: *Applied Physics Letters* 92.14 (2008), p. 143309. DOI: 10.1063/1.2901378.

[16] Claude Weisbuch et al. "The efficiency challenge of nitride light-emitting diodes for lighting." In: *Physica Status Solidi (a)* 212.5 (2015), pp. 899–913. DOI: 10.1002/pssa.201431868.

[17] YC Shen et al. "Auger recombination in InGaN measured by photoluminescence." In: *Applied Physics Letters* 91.14 (2007), p. 141101. DOI: 10.1063/1.2785135.

[18] Emmanouil Kioupakis et al. "Indirect Auger recombination as a cause of efficiency droop in nitride light-emitting diodes." In: *Applied Physics Letters* 98.16 (2011), p. 161107. URL: https://aip.scitation.org/doi/full/10.1063/1.3570656.

[19] JY Tsao et al. "Ultrawide-Bandgap Semiconductors: Research Opportunities and Challenges." In: *Adv. Electron. Mater.* 4.1 (2017), p. 1600501. DOI: 10.1002/aelm.201600501.

CHAPTER 2

Theoretical Background for Calculating Properties of Solids

Before being able to calculate anything using a computer, we first have to mathematically explain what we are trying to study. This chapter introduces some of the important concepts and theory behind the first-principles calculations that we use throughout this thesis. It is not meant to be comprehensive, but to serve as an introduction. Much more detailed information can be found in the references. We begin by making the many-body Schrödinger equation more manageable through approximations relevant to solids. From there, we show the fundamental idea behind DFT – that the ground state electron density rather than the wave functions of the electrons – can allow us to calculate important (ground state) properties. I discuss other approximations that are necessary for DFT and the theory's limitations. I discuss the theory of Wannier functions, which is an important part of being able to calculate the Auger recombination rates. Finally, I discuss how the GW method can be used to correct the band gap found from DFT.

2.1 Quantum Mechanics

Everything we interact with in a day is made up of atoms. While the word atom comes from the Greek word for *uncuttable*, scientists in the late 19th and early 20th century gradually revealed that atoms were actually made up of smaller pieces. In the center is the nucleus made up of protons and neutrons. Surrounding the nucleus are smaller particles called electrons. Understanding how electrons and nuclei interact with each other and with outside forces has allowed us to develop many new technologies over the past century.

The starting point for understanding these tiny particles is quantum mechanics. Quantum mechanics describes particles using *wave functions*. Wave functions allow us to calculate prop-

erties about particles – like where it's most likely to be located. We can find the wave functions by using Schrödinger 's equation:

$$\hat{H}\Psi(\{\mathbf{r}_N\},\{\mathbf{R}_M\}) = E\Psi(\{\mathbf{r}_N\},\{\mathbf{R}_M\})$$

A deceptively simple way of writing the equation we are hoping solve[1]. Here we consider a general system made up of N electrons and M nuclei. The positions of electron n and nucleus m are described by the vectors $\mathbf{r_n}$ and $\mathbf{R}_m$, with the curly brackets representing that there are n or m dfferent vectors for *each* one of the two types of particles: $\{\mathbf{r}_N\} = \mathbf{r}_1, \mathbf{r}_2 \ldots \mathbf{r}_N$.

The complexity becomes more apparent if we write out the Hamiltonian, or total energy operator $\hat{H}$ out explicitly[2].

$$\hat{H} = \underbrace{-\sum_{i=1}^{N} \frac{\nabla_i^2}{2}}_{\text{Kinetic Electron}} + \overbrace{\frac{1}{2}\sum_{i\neq j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}}^{\text{Electron-Electron Potential}} \underbrace{-\sum_{I=1}^{M} \frac{\nabla_I^2}{2M_I}}_{\text{Kinetic Nuclei}} + \overbrace{\frac{1}{2}\sum_{I\neq J} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|}}^{\text{Nuclei-Nuclei Potential}} \underbrace{-\sum_{\substack{i=1 \\ I=1}}^{N,M} \frac{Z_I}{|\mathbf{r}_i - \mathbf{r}_I|}}_{\text{Electron-Nuclei Potential}} \tag{2.1}$$

The Hamiltonian describes the different energies in the collection of electrons and nuclei. There is kinetic energy from both particles and three types of electrostatic potential energy: between electrons, between nuclei and between electrons and nuclei.

Solving this (complex) equation would give us the general, many-body, wave function (called Ψ)) of our system. There are many possible Ψ that satisfy this equation and one that would be particularly useful is the one with the lowest energy, or the ground-state wave function Ψ_0. Having this wave function would allow us to understand how the system behaves in equilibrium – what phases it might take, how it behaves when stretched or compressed, whether it will be easy to create in a lab or not. Unfortunately, this equation is still far too general to be of much use. This equation encompasses more than just solids and would hold for gases and liquids as well. Not only that, but finding a specific "closed" solution is possible only for extremely specific cases.

To make things simpler, let's think about a crystal like in fig. 2.1. Because crystals repeat over and over again, we can get away with just describing one part (called the unit-cell). This alone vastly reduces the number of electrons we need to worry about. In addition, we often know where the nuclei are located. This means that for the most part, the nuclei are not important because they are stationary and their positions are well known. It's how the electrons move (or do not move) through the material that determines many of its interesting properties.

[1]To non-expert readers, this is probably the most math heavy chapter. Feel free to skim over equations. Just looking at how the terms are are labeled can be helpful in understanding what I'm talking about.

[2]Here written in atomic units.

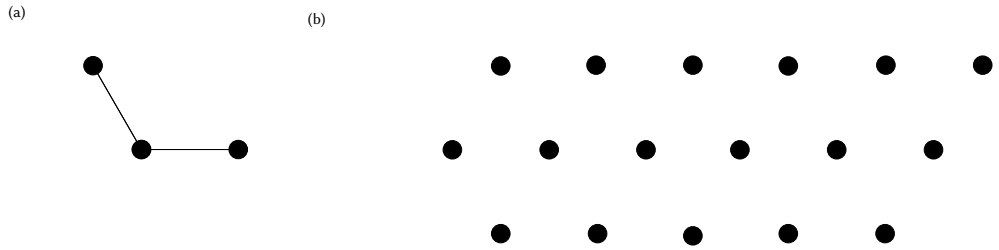

Figure 2.1 An example of a 2D crystal. In (a), a simple unit cell is shown. By repeating that one element over and over again, we get the hexagonal crystal structure shown in (b).

We can represent that the nuclei are stationary in the equation by setting the nuclei masses to ∞, causing the kinetic energy term in our Hamiltonian to be zero. The fixed positions cause the $\{\mathbf{R}_I\}$ to be constants that are determined by the type of material. This means that the nuclei-nuclei energy is a constant, and we can incorporate it into our expression for the total energy:

$$E = E_{\text{tot}} - \frac{1}{2}\sum_{I\neq J}\frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|}$$

The nuclei-electron potential energy is also altered. Since the R_I are constant this term only depends only on the positions of the electrons $\{\mathbf{r}\}_i$, it makes sense to explicitly hide the nuclei coordinates by calling the Coulomb potential from nuclei a function of r_i:

$$V(\{\mathbf{r}\}_i) = -\sum_{\substack{i=1\\I=1}}^{N,M}\frac{Z_I}{|\mathbf{r}_i - \mathbf{R}_I|}$$

These approximations are known as the *aiabatic* or *Born-Oppenheimer* approximations, and allow us to rewrite the Hamiltonian in eq. 2.1:

$$\hat{H}\Psi(\{\mathbf{r}\}) = \left[\sum_{i=1}^{N}\frac{\nabla_i^2}{2} + \frac{1}{2}\sum_{i\neq j}\frac{1}{|\mathbf{r}_i - r_j|} + V(\mathbf{r}_i)\right]\Psi(\{\mathbf{r}\}) = E\Psi(\{\mathbf{r}\}) \tag{2.2}$$

This is the fundamental equation for *electronic structure* calculations.

Is this equation solvable? The most popular LED material is gallium nitride and there are 4 nuclei and 76 electrons that we need to keep track of in this material's unit cell. So eliminating the 4 nuclei helped, but not much. We went from 80 position vectors to 76 position vectors (each made up of three components in x,y, and z). To solve this on a computer, we need to break up

the unit cell into discrete points. The unit cell is *about* 3 ×3 ×5 ångströms[3]. To start, we might divide this space evenly with each axis of the box divided into ten pieces giving us 1000 total points. For *each* of those thousand points, we would need to calculate $\Psi(\mathbf{r}_1, \mathbf{r_2}, ...\mathbf{r_{76}})$. Meaning that we would need $1000^7 6 = 1 \times 10^{228}$ numbers to describe this system. It is hard to comprehend how large this number is, but just *storing* this many numbers on a computer would require an impossible amount of space[4]. Solving it would be even more difficult. And keep in mind that this is a pretty rough approximation. We might need to divide the unit cell even more finely to get accurate properties.

So we have to continue trying to make this equation simpler but still useful. The first thing we will do is make a very large approximation – what if the electrons didn't interact at all? Then the Coulomb term would disappear and the many-body wave function would separate into many independent wave functions $\Psi(\{\mathbf{r}_N\} = \psi(\mathbf{r}_1)\psi(\mathbf{r}_2)...\psi(\mathbf{r}_n)$. Of course this is a *huge* approximation to make, and much of the actual physics and interesting behavior is lost. But by going to an extreme, we can add pieces back into our model individually until we find something that is still relatively simple in comparison to eq. 2.1 but also incorporates the physics we care about.

We'll start by realizing that we can average out the way that the electrons interact with each other. We won't consider each pair of electrons explicitly – instead only considering the average potential that each electron feels from the other electrons. This is called the *mean-field* approach, and still keeps the wave functions independent of each other. But we now add this average potential for each electron as:

$$V_H(\mathbf{r}) = \int d\mathbf{r}' \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \tag{2.3}$$

This potential is known as the Hartree potential and takes into account the effects of all of the electrons from the electron density, $n(\mathbf{r})$[5].

What is missing from this potential? There are two quantum effects that we care about but are not incorporated. The first is *electron exchange*. The Pauli exclusion principle states that electrons cannot be in exactly the same state[6]. But with the equation above, electrons can have exactly the same energy. We can add a term V_x, called the exchange energy, which allows us to model how electrons do not occupy the same state while keeping the wave functions independent. The

[3]An ångström is 1×10^{-10} meters. Human hair is around 1×10^{-5} meters in width. This five order of magnitude difference is similar to the difference between a small leap and traveling once around the world. Atoms are tiny!

[4]One number takes around 4 bytes to store on a computer. We'd need 8×10^{228} bytes to store that many-body wave function. According to,[1] this could store all the information humans have ever stored on digital devices many times over.

[5]I'm glossing over needing to now solve the equations *self-consistently*, a topic covered shortly in sec. 2.2.2.

[6]Or more formally, that since electrons are fermions the many-body wave function must change signs if exchange the variables of any two electrons.

second effect is *electron correlation.* When we removed the Coulomb potential from our equations, we removed how electrons repel each other. This can be an important effect, and just like with the exchange effect, we can add a term to account for this called the correlation energy to our equation V_c.

This leaves us with the following equation:

$$\hat{H}\psi_i(\mathbf{r}) = \left[\frac{\nabla_i^2}{2} + V_H(\mathbf{r}) + V_x(\mathbf{r}) + V_c(\mathbf{r}) + V(\mathbf{r})\right]\psi_i = E\psi_i(\mathbf{r}) \tag{2.4}$$

Which isn't all that different from what we started with. All of the potentials here stem from the same Coulomb interaction in eq. 2.2, but we've just written them out as different explicit parts which allow us to approximate them better. Indeed, you may be upset at "just add another term" to the equations as a possible solution! It turns out that for all but the most simple combinations of atoms, V_x and V_c do not have exact equations we can write down. A major part of the next few sections will be to figure out good approximations to these terms.

2.2 Density Functional Theory

Where do we go from eq. 2.4? Your first instinct might be to figure out those extra terms we added in the last section. But before we do that there is one more big insight we can make. Up until now we've had to consider each electron's wave function individually. The more electrons the more equations we have to solve. It would be much easier to solve this problem if we only had a few equations to solve regardless of how many electrons there are. This is where density functional theory (DFT) comes into play. The central realization of DFT is that rather than having to solve for each individual electron wave function, (which requires 3N variables) we can instead find the electron charge density, a function of only 3 variables.

2.2.1 Hohenberg-Kohn

This simplification was mathematically shown by Hohenberg and Kohn in 1964.[2] Their central realization was that the electron density entirely determines the external potential of the system (aside from a constant). Since other parts of the Hamiltonian do not depend on the electron density, this means the Hamiltonian is entirely determined by the electron density and that the ground state wave function and energy are also determined by the electron density.

This is written more formally by using functionals. A *functional* is similar to a function. But where a function takes a number and gives a number output, a functional takes a whole function and gives a number as an output. A simple example from calculus is the process of definite

integration, which takes a function and returns a number.

$$I[f(x)] = \int_a^b f(x) = F(b) - F(a) \tag{2.5}$$

In the case of quantum mechanics, we can write the total energy of the system as:

$$E = \langle \psi[n] | \hat{H} | \psi[n] \rangle = \langle \psi[n] | \hat{T} + \hat{V}_{ext} + \hat{V}_H + \hat{V}_{xc} + | \psi[n] \rangle \tag{2.6}$$

$$= T[n] + E_H[n] + E_{xc}[n] + E_{ext}[n] \tag{2.7}$$

Where I've used the Hamiltonian from eq. 2.4. This can be broken into two parts, a "universal" part that comes from the non-system dependent general terms (kinetic, Hartree, and the exchange-correlation[7]) $\mathscr{F}[n] = \hat{T} + \hat{V}_H + \hat{V}_{xc}$ and a system dependent external potential term:

$$E_{ext}[n] = \int d\mathbf{r}\, n(\mathbf{r}) V_{ext}(\mathbf{r})$$

Combining we have the total functional:

$$E[n] = \mathscr{F}[n] + E_{ext}[n]$$

Minimizing this functional will give us the ground state density and energy corresponding to the external potential.

2.2.2 Kohn-Sham Equations

With the Hohenberg-Kohn theorems we know that we *can* use the electron density to obtain properties of the many-body wave function...but we don't have a good recipe to actually do so. The Kohn-Sham method will give us a way to do this.[3] This method establishes that we can use the ground state density of an *equivalent* single particle system (like in eq. 2.4) to find the ground state density of the many-body wave function. This equivalent system of electrons are all affected by a different external potential called the Kohn-Sham potential, V_{KS} which incorporates all of the Coulomb potentials we've been discussing, including the many-body effects:

$$V_{\text{KS}} = V_{ext}(\mathbf{r}) + V_H(\mathbf{r}) + V_{xc}[n](\mathbf{r})$$

[7] I've merged the exchange and correlation effects into one potential that hides these many-body effects. The reason for merging them is that treating them separately winds up removing cancellations between the two terms.

2.2.3 The Exchange-Correlation Functional

Choosing the exchange-correlation potential proves to be the major approximation we make in DFT. I will briefly review the two main methods of coming up with V_{xc} , but there are many more complex methods.[4,5] In my work, I have made use of both of these simple approximations for V_{xc} .

2.2.3.1 The Local Density Approximation

The first and simplest way to approximate V_{xc} is called the local density approximation (LDA). This approximation considers the electrons to be a uniform electron gas of density $n(\mathbf{r})$. For metals, where electrons are nearly free to move throughout the solid, this is close to how electrons actually behave. Even for non-metals it is a very useful approximation because exchange and correlation are usually short-range anyway, so the "big picture" of the electrons moving around is very similar to a uniform electron gas.

Because this is such as simple system, the exchange energy can be calculated exactly. While there is not an *analytic* expression for the correlation energy, simulations allow for writing down a quantitative form.[5] With these two pieces, we can calculate V_{xc} .

For systems with electrons in s and p orbitals, LDA is very successful. Bond lengths, angles and vibrational frequencies can be calculated to within a few percent. Elastic properties are also accurate. LDA tends to underestimate bond lengths because of overestimating the binding energy. LDA does not do well with predicting properties of molecules, modeling core states or dealing with metals with d and f orbitals. This makes sense because these cases are often more localized, which is the opposite of our assumption that the electrons behave like a free electron gas. One method of improving LDA is to add a correction to the energies of the problematic localized orbitals. This method is called local density approximation with a +U Correction (LDA+U) or, more accurately density functional theory (any functional) with a +U correction (DFT+U) since it can be used on functionals other than in the LDA approximation. In this method, only certain orbitals are given a correction energy (the +U) similar to Hubbard models of solids.[6]

Another more general method is the subject of the next section.

2.2.3.2 The Generalized Gradient Approximation

The LDA can be expanded by considering higher order derivatives of the electron density, similar to how you would construct a Taylor series to approximate a function in mathematics. Here we can write the V_{xc} as:

$$E_{\text{xc}} = \int d\mathbf{r}\, n(\mathbf{r}) \epsilon_{\text{xc}}(n(\mathbf{r}))$$

If we take only the first new term, based on the gradient of the electron density we are considering the generalized gradient approximation or GGA. There are many ways of doing this expansion and it winds up being tricky to get right.[5] One of the most popular methods is the PBE method which tries to satisfy as many theoretical constraints as possible (rather than fitting to experimental set of data). generalized gradient approximation (GGA) is better at dealing with molecules than the LDA. It has the opposite trend in bond lengths and binding energy (overestimate, underestimate).

2.2.4 Planewave Basis

There are many basis sets we could choose to represent the wave functions we calculate using DFT. A very general set of basis functions are plane waves. Plane waves are useful in particular because of the periodic boundary conditions in the systems we are trying to describe. We also can systematically increase the accuracy of our calculations by increasing the number of plane waves we use to describe the system. This systematic way to increase accuracy is helpful when trying to obtain very accurate calculations. However planewaves can also be very costly in terms of describing rapidly electronic states near the core of atoms. To get the fine detail required for these core states, a prohibitive amount of plane waves can be required. This lead to the development of pseudopotentials.

2.2.5 Pseudopotential Method

One of the challenges to using a plane wave basis is describing the potentials that the electrons experience from the nuclei. These potentials exhibit discontinuities at the origin or change rapidly – both of which require using *many* plane waves to accurately describe (+fig. 2.2).

To get around this, we can create effective potentials or *pseudopotentials* that are not the actual potentials seen by the electrons but yield the same properties. To understand why we can do this, we first have to understand the difference between core and valence electrons.

Nearly everyone is aware of the typical picture of electrons "orbiting" around a central nuclei – but this picture is entirely wrong. The motion of electrons around nuclei is much more complex and unintuitive. Electrons in an atom occupy certain areas around the nuclei. These areas are *called* orbitals but again, they are not orbiting like planets in a solar system. Depending on the energy of the electron, the shape of these orbitals will be different.

Electrons at lower energies are "more tied" to their nuclei and don't want to move through a material at all. It turns out that we can often ignore these electrons, making the calculations simpler without sacrificing much accuracy. The way to do this is through the pseudopotential method.

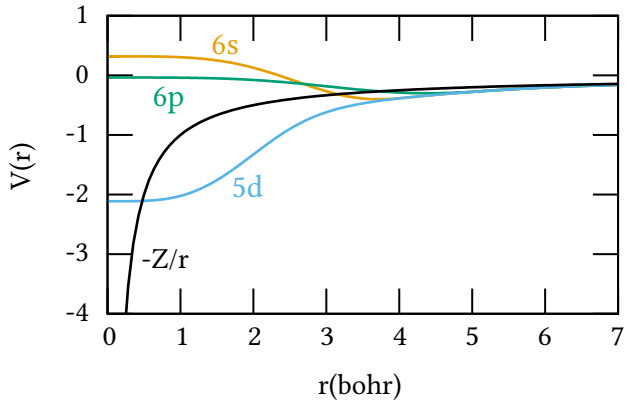

Figure 2.2 Pseudopotentials generated for Cesium. The actual potential is diverges to negative infinity, but the pseudopotentials are smoother and do not diverge, allowing for less demanding calculations with fewer planewave components while still accurately representing long range behavior of the potential.

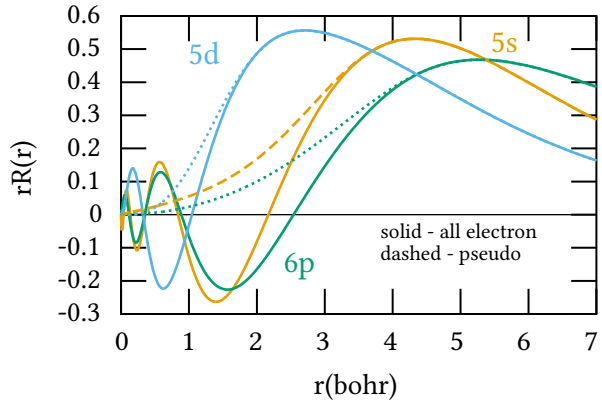

Figure 2.3 The corresponding radial wave functions for Cesium. When all electrons are taken into account the wave functions rapidly vary closet to the nucleus requiring many planewaves to accurately model. But by smoothing this behavior near the core and acccurately representing behavior away for the core we can reduce the number of planewaves needed to model the system.

Pseudopotentials separate the electrons in an atom into *core* and *valence* electrons, terms you might be familiar with from a chemistry class taken long ago. Core electrons are, unsurprisingly very close to the nuclei. Valence electrons are less tied to the nuclei and in solids, can often roam around from nuclei to nuclei, giving a material it's properties.[8]

The main question is: if we ignore the core electrons, how do we get accurate potentials for the remaining valence electrons. Without explicitly having the core electrons (which repel the outer electrons because of their same charge) we need to somehow simply simulate that repulsion.

Pseudopotentials provide the solution to this problem and a prescription for making potentials for the outer electrons without explicitly considering the core electrons.

To construct pseudopotentials we start by separating out a core and valence part of the electron wave functions .

$$|\psi_v\rangle = \overbrace{|\tilde{\psi}_v\rangle}^{\text{Smooth Part}} + \sum_c \alpha_{cv}|\psi_c\rangle$$

Step one in finding a pseudopotential is finding a model for the smooth part $\tilde{\psi}_v$. If we plug this smooth part into our original Hamiltonian, we'll find that it doesn't actually satisfy the original equation. This shouldn't be a surprise, since it's a totally different function. But what we can do instead is come up with a pseudo-Hamiltonian that the smooth part *does* satisfy:

$$\hat{H}_{\text{PS}} = \hat{H} + \sum_j (\epsilon_v - \epsilon_j^c)|\psi_c^j\rangle\langle\psi_c^j| \tag{2.8}$$

$$= \underbrace{\hat{T} + \hat{V} + \sum_j (\epsilon_v - \epsilon_j^c)|\psi_c^j\rangle\langle\psi_c^j|}_{\text{Pseudopotential}} \tag{2.9}$$

Our pseudopotential will always be weaker than V, since our difference in energies is greater than zero because valence electrons have higher energy than core electrons. The pseudopotential is also "semi-local", meaning it acts differently on states with different angular momentum. This is more explicit if we write in terms of spherical harmonics:

$$V_{\text{PS}} = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} \overbrace{V_{\text{PS}(r)}}^{\text{Radial}} \underbrace{|\ell m\rangle\langle\ell m|}_{\text{Spherical Harmonics}}$$

[8] This is not a hard distinction. The fuzzy are between core and valence electrons is called semi-core electrons. These electrons may be important to include for some properties or to get more accurate calculations but not for others.

We've shown that we can satisfy the Schrödinger equation in a modified way with a smooth valence ave function. This is useful because we can use this to find a smootherpotential (the pseudopotential). This is an inverse problem, which we can write:

So we find the pseudopotential and then we can find the pseudo wave function. There is a lot of flexibility in choosing this potential, but there are many physical characteristics that the potential must satisfy, which lets us find ones that are appropriate for our material.

1. Outside of a certain radius (the cutoff, r_c) the pseudo-radial wave function must match the true radial wave function.
2. At this cutoff radius, we must enforce the continuity of the radial wave function

Another property that we use is *norm-conservation*, which isn't necessary but aids in the transferability of the pseudopotential. By making sure the norm of our pseudo-wave functions are conserved we make sure that these wave functions are valid at a range of energies rather than at one particular energy. One of the best examples of norm-conserving pseudopotentials is the Troullier-Martin's method Troullier and Martins [7], (which is what we use for our pseudopotentials throughout this thesis). This method enforces continuity up to four derivatives, and enforces zero curvature at the origin in addition to the norm-conservation property.

2.3 Quasiparticle Corrections with GW

In theory, DFT can be used to find any ground or excited state property we want. In practice, without being able to explicitly write down V_{xc} , DFT is limited to ground state properties of the material. This is limiting because one of the key properties for understanding LEDs is determining the band gap which is an excited state property. However, DFT often provides a very good *starting point* which can be corrected by other methods.One of the main methods used in this thesis is the GW method. This method uses Green's function (G) and the screened Coulomb interaction (W) to find accurate excited state properties of the systems that we are interested in.

This method views electrons moving through the system as *quasiparticles*. As an electron moves it is always pushing other electrons nearby and therefore has a trail of positive charge that follows it around. This combination of electron with a cloud of positive charge is called a quasiparticle and has an energy, effective mass and a lifetime. Using Green's functions, we can describe these quasiparticles which *include* the interactions that we had hidden in DFT. +fig. 2.4 shows the improvements that are made on the band gap when corrected with the GW method, one of the central excited state properties we are interested in. The tradeoff is that these calculations are much more expensive. A thorough review by Aryasetiawan and Gunnarsson [8] is a good place to start with the methodology behind this method.

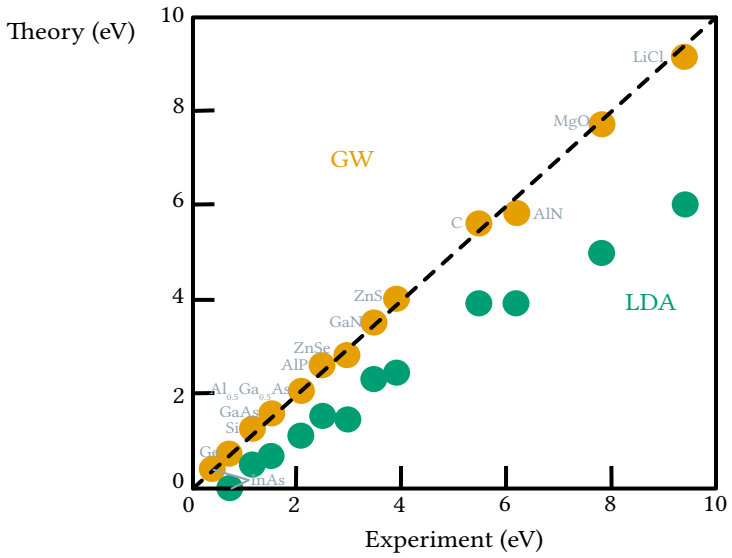

Figure 2.4 A comparison of how band gaps compare from DFT using LDA and from GW. LDA consistenly underestimates the band gap, and in some cases does not find a band gap at all. GW on the other hand is able to get very good agreement with experiment. The data in this figure is from Aryasetiawan and Gunnarsson [8].

2.4 Maximally Localized Wannier Functions

The idea behind Wannier functions is to take the extended, Bloch like states that we use in our first principles code and reduce them to localized states in real space. You can loosely think about this in terms of Fourier transforms between wave functions in k and real space. In k space, wave functions are spread out over entire supercells which will correspond to localized wave functions in real space.

Wannier functions can be used in a technique called Wannier interpolation.[9] Starting with a DFT calculation on a coarse mesh, we can find the associated Wannier functions. These Wannier functions can serve as a basis set to find desired properties on any k point by transforming the localized functions back to Bloch states at intermediate k points. This technique allows us to calculate band energies on very fine k-meshes which are required for calculating Auger rates.

References

[1] M Hilbert and P López. "The world's technological capacity to store, communicate, and compute information." In: *Science* 332.6025 (2011), pp. 60–65. DOI: 10.1126/science.1200970.

[2] Pierre Hohenberg and Walter Kohn. "Inhomogeneous electron gas." In: *Physical review* 136.3B (1964), B864. URL: https://link.aps.org/pdf/10.1103/PhysRev.136.B864.

[3] Walter Kohn and Lu Jeu Sham. "Self-consistent equations including exchange and correlation effects." In: *Physical review* 140.4A (1965), A1133. URL: https://link.aps.org/pdf/10.1103/PhysRev.140.A1133.

[4] Richard M. Martin. *Electronic Structure*. Cambridge University Press, 2004. URL: http://books.google.com/books?id=dmRTFLpSGNsC%5C&hl=%5C&source=gbs_api.

[5] Jorge Kohanoff. *Electronic Structure Calculations for Solids and Molecules*. Cambridge University Press, 2006. URL: https://play.google.com/store/books/details?id=v2ElZFCKeBsC%5C&source=gbs_api.

[6] Vladimir I. Anisimov, F. Aryasetiawan, and A. I. Lichtenstein. "First-principles calculations of the electronic structure and spectra of strongly correlated systems: the LDA+U method." In: *Journal of Physics: Condensed Matter* 9.4 (1997), pp. 767–808. DOI: 10.1088/0953-8984/9/4/002.

[7] N. Troullier and José Luriaas Martins. "Efficient pseudopotentials for plane-wave calculations." In: *Physical Review B* 43.3 (1991), pp. 1993–2006. DOI: 10.1103/physrevb.43.1993.

[8] F. Aryasetiawan and O. Gunnarsson. "The GW Method." In: *Reports on Progress in Physics* 61.3 (1998), pp. 237–312. DOI: 10.1088/0034-4885/61/3/002.

[9] Nicola Marzari et al. "Maximally localized Wannier functions: Theory and applications." In: *Reviews of Modern Physics* 84.4 (2012), pp. 1419–1475. DOI: 10.1103/revmodphys.84.1419.

CHAPTER 3

The Auger Effect

The Auger effect describes an electron transitioning to a lower energy state and giving the transition energy to another carrier, which is excited into a higher energy state. This process was initially described and studied in atoms, but eventually was seen to be an important recombination mechanism for free carriers in semiconducting materials. This chapter begins with a short history of the discovery of Auger effect in atoms and the realization of its importance in semiconductors. After discussing early theoretical and experimental efforts to measure Auger recombination in semiconducting materials, I review the relevant theory and derive the equations used in our first-principles code. I then discuss recent novel experiments related to Auger and other early first-principles efforts to understand this process. I conclude by emphasizing the importance of the Auger effect in semiconducting materials, in particular with LEDs (efficiency droop problems).

3.1 Discovery of the Atomic Auger Effect

Explaining the Auger effect was primarily done by Lise Meitner and Pierre Auger in the early 1920's. Meitner was interested in understanding β–radiation. At the time, there was intense disagreement between over why β–radiation appeared to be both continuous (having any energy) and discrete (having very specific energy values). Meitner's group observed that β–radiation was discrete while Ernest Rutherford showed that the β–radiation was continuous with discrete peaks superimposed.

One confusing aspect of this discussion is that *all* electrons are referred to as β–radiation, even though electrons ejected from the atomic shells are *just* electrons (and are not accompanied by the additional antineutrino that we associate with β–radiation today). This was one of the main issues in explaining the phenomenon and was not resolved theoretically until the 1930's when Wolfgang Pauli proposed the antineutrino particle, and finally put to rest in 1956 when the

antineutrino was observed experimentally).

The Auger effect was first described by Lise Meitner in 1922 while studying β–decay of radioactive nuclei.[1] At the time, there was an argument going on over why there was both discrete energies for β~ particles as well as a continuous spectrum of energies[1]. In describing the discrete β~ energies, Meitner wrote (translation from German in[2]):

> The primary β–rays eject K-electrons, thus exciting the K_α-radiation, which in turn disperses L-, M-, or N-electrons, the whole process from the emission of the primary β–ray related to the nucleus disintegration to the dispersion of an L-, M-, or N-electron, occurring within the same atom.

The references to L, M, and N electrons are in *X-ray notation* and correspond to the principle quantum number of orbitals (n). This description is the first description of the Auger effect – in this case, an excited electron relaxes causes another electron to be emitted from the atom.

At the same time as Meitner's writing on the β–decay process, Pierre Victor Auger was using cloud chambers to study atomic photoexcitation. In an article in 1923,[3] he wrote (translation from French in[2]):

> When the first electron leaves, as a secondary β–ray, there is a vacancy left in the electronic system of the excited atom. The drop of a more peripheral electron on that level is accompanied by the emission of a characteristic radiation quantum. This quantum may be absorbed in the atom itself, and produce, at the expense of the peripheral levels, a tertiary β–ray (...). The repetition of that process must lead to the production of a fourth order ray; and I indeed believe I have observed such rays in the case of iodine.

In this case, Auger explicitly described what eventually became known as the Auger effect.

3.2 Auger Effect in Semiconductor Materials

The Auger process described above is based on observations of atoms. Auger recombination in semiconductors wasn't recognized as an important process at first. Other recombination processes were more likely and occurred more often. Recombination because of defects was what early devices struggled with. In addition, radiative recombination, where an electron and hole recombine to emit a photon, was another major recombination process. As growth techniques improved, defects had less of an impact on device operation, and devices began to be run at higher

[1] Ultimately, this discrepancy was resolved when antineutrinos were experimentally confirmed in 1956.

current densities. This is where lifetimes of carriers were observed to be shorter experimentally than what was being predicted by theory.

First explanations were non-radiative processes that lead to large simultaneous (10 to 100) phonon emissions.[4] However, this simultaneous emission was found to be improbable. Early research also focused on trapping carriers and often data could be fit using these models, but they often provided no additional understanding of the underlying physics. Instead, a process similar to the Auger effect in atoms was proposed, first in metals by Skinner [5] and then in semiconductors by Fröhlich and O'Dwyer [6]. A few years later, Pincherle [7], in a short letter to the editor discusses trap assisted Auger, while Bess [4] gave a more thorough calculation of Auger transitions with defects. Sclar and Burstein [8] did calculations and showed that in some cases Auger could be the dominant recombination mechanism at higher temperatures.

Building on the early work mentioned above, Beattie and Landsberg [9] were the first to put a more rigorous understanding of band-to-band Auger recombination processes and this lead many others toward getting more and more accurate calculations of the Auger recombination process (see Beattie [10] for an extended list). In the 1950's, computational materials science had not yet been born, and any sorts of calculations of Auger recombination lifetimes were done by hand and required many simplifications of the band structures involved. These approximations included only using Boltzmann statistics, neglecting electron-electron screening and assuming parabolic bands. Nevertheless, Auger lifetimes and were obtained in reasonable agreement with experimentalists at the time.

3.3 Describing Recombination Processes Using the ABC Model

There are three main types of recombination that take place in LEDs. Shockley-Read-Hall (SRH) is recombination that is mediated by defects in the material. Radiative recombination is the desired process, and involves an electron and hole recombining and emitting a photon. Finally, Auger recombination, which is a three-carrier process where an electron and hole recombine and instead of emitting light, give that energy to a third carrier. That carrier an be an electron being excited to a higher band or a hole being excited into a lower valence band creating two types of Auger processes, *eeh* and *hhe* respectively. Auger can also take place in a direct or indirect way. Direct Auger is simply the process described above, when the recombined electron and hole *directly* give their energy to a third particle. Indirect Auger is when there is a mediating mechanism for this energy transfer. This can be through effects like phonons or alloy disorder which increase the amount of final states available to the excited particle by making it easier to conserve momentum

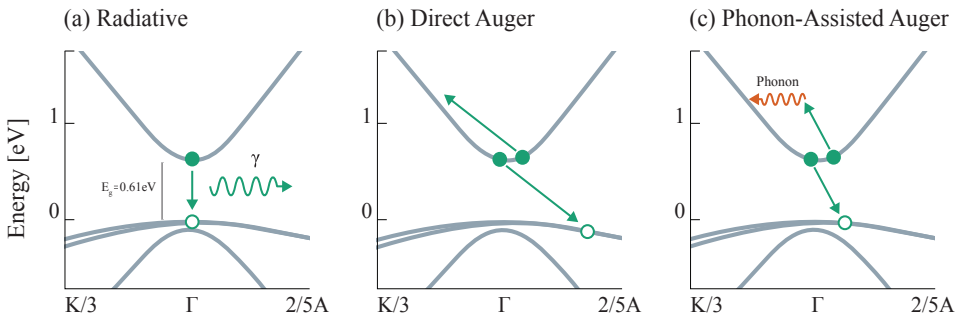

Figure 3.1 (a) is the radiative recombination process, where an electron and hole recombine to emit a photon. This is the process that we desire in LED. (b) shows a schematic of direct Auger recombination, where an electron and hole recombine and excite a third carrier. In this case the excited carrier is an electron and this is *eeh* direct Auger recombination. Finally in (c), is phonon-assisted Auger recombination. Without the extra momentum from the phonon, this Auger transition would not be possible.

during the Auger process.

To quantify these recombination rates and their effects LEDs in particular) the *ABC Model* is often used (as in eq. 3.1).[11] In the ABC model, the total recombination rate is written as the sum of Shockley-Read-Hall (SRH) or defect, radiative, and Auger recombination rates. The rates in turn are written in terms of the carrier densities involved and *recombination coefficients* which are proportionality factors for each rate.

$$\frac{dn}{dt} = R = \overbrace{An}^{\text{Defects}} + \underbrace{Bn^2}_{\text{Radiative}} + \overbrace{Cn^3}^{\text{Total Auger}} \tag{3.1}$$

This model explains in a simple way the recombination processes inside of semiconductor materials, and gives us a way of understanding the internal quantum efficiency (IQE) of the material.

$$\text{IQE} = \frac{\text{Emitted Photons}}{\text{Total Recombination}} = \frac{Bn^2}{An + Bn^2 + Cn^3} \tag{3.2}$$

For a simple model, the ABC model does a good job of fitting the IQE curves of LED materials. However its simplicity is both an advantage and a disadvantage. One of the main assumptions made is that the recombination processes do not depend on the carrier density. At high carrier densities, this assumption fails in an effect known as *phase-space filling*.[12] As carrier density increases there are fewer states for excited carriers to transition to through the Auger process. At higher carrier densities screening also become more important and causes the rate coefficients to vary. This has been experimentally shown by Aurelien et al. [13] to occur in bulk-GaN. Occa-

sionally additional terms are added to the ABC model to account for "higher order" processes.[14] However, adding extra terms will improve the model's fit to experimental data without those processes actually occurring in the material itself.

3.4 Auger and the Golden Rule of Time Dependent Perturbation Theory

Using time-dependent perturbation theory we can find analytical expressions for the C recombination parameter which we will then use to calculate these numbers from first principles. For a more in depth description of the methodology, see Kioupakis et al. [15]. For calculating how carriers change state with respect to time, we turn to the *Golden Rule* of time dependent perturbation theory[16] which allows us to find the probability of a transition per unit time as:

$$\frac{R}{V} = \frac{4\pi}{h}\frac{1}{V_{\text{cell}}^3}\frac{1}{N_{\mathbf{k}}^3}\sum_{\mathbf{1234}} P|VM_{\mathbf{1234}}|^2\delta(\epsilon_{\mathbf{1}}+\epsilon_{\mathbf{2}}-\epsilon_{\mathbf{3}}-\epsilon_{\mathbf{4}}) \tag{3.3}$$

Here P is an occupation number factor using Fermi-Dirac statistics: $P = f_{\mathbf{1}}f_{\mathbf{2}}(1-f_{\mathbf{3}})(1-f_{\mathbf{4}})$ which makes sure that transitions occur from occupied to unoccupied states. M is the matrix element for the perturbing Hamiltonian, in this case the screened Coulomb potential:

$$W = \frac{1}{V_{\text{cell}}}\frac{1}{\epsilon(\mathbf{q})}\sum_{\mathbf{q}}\frac{4\pi e^2}{q^2+\lambda^2}e^{i\mathbf{q}\cdot(\mathbf{r}_1\mathbf{r}_2)} \tag{3.4}$$

Here written in k-space. We consider both direct and exchange terms for the matrix elements. The sums are over both k-point and band indicies $\mathbf{1} = (n_1, k_1)$.

3.4.1 Phonon-Assisted

Adding phonons requires us to use the 2nd Order Equation for Fermi's Golden Rule and makes some other modifications. The full formalism can again be found in Kioupakis et al. [15].

$$R = \frac{4\pi}{h}\sum_{\mathbf{1234},\mathrm{v}\mathbf{q}}\tilde{P}|\tilde{M}_{\mathbf{1234},\mathrm{v}\mathbf{q}}|^2\delta(\epsilon_{\mathbf{1}}+\epsilon_{\mathbf{2}}-\epsilon_{\mathbf{3}}-\epsilon_{\mathbf{4}}\mp h\omega_p) \tag{3.5}$$

Where we have added the energy of the phonon to the energy delta function. $\tilde{P}$ now includes the phonon occupations and the $\tilde{M}$ matrix elements now include the intermediate phonon state and consider the eight different possible phonon-assisted Auger processes.

3.5 Experimental Measurement Techniques

Most experimental measurements rely on getting carrier lifetime data from photoluminescence experiments and using the ABC model to fit this data (with A known from other techniques and B from theoretical calculation). However, as mentioned before, the ABC model does not include certain effects and these fitting techniques can lead to a wide range of possible Auger recombination coefficients (see sec. 7.3 and sec. 5.2). In addition, experiment is often only able to determine the *total* Auger rate and not the contributions from *eeh* /*hhe* or direct/indirect processes. In designing new materials with higher IQE it is important to know which Auger process is dominant to better understand methods of reducing it's impact.

3.6 Experimentally Verifying Auger in LED Materials

While directly measuring the recombination rates is difficult, there are other experimental methods that have at least measured Auger electrons being generated in LED materals and helped explain this difficult to measure process. While these experiments do not determine the Auger recombination *rate*, they provide evidence that the Auger process is taking place and justifies studying the process with first-principles methods.

Strong support for Auger as the primary mechanism in LEDs came from Iveland et al. [17]'s 2013 paper. Using an InGaN LED, they measured hot electrons in vacuum during device operation that coincided with the onset of efficiency droop of the device. In the same year, researchers at OSRAM used a novel quantum well setup to determine whether *eeh* or *hhe* Auger was dominant in InGaN.[18] They created quantum wells that alternated between emitting green and UV light. Using a laser, they optically excited only the green wells to carrier densities that the device would see under electrical carrier injection. They observed some UV emission, confirming that carriers were being excited out of the green wells and into the nearby UV wells through an Auger process. While they were able to measure Auger electrons, the microscopic details of Auger are still unclear.

References

[1] Lise Meitner. "Über die Entstehung der ?-Strahl-Spektren radioaktiver Substanzen." In: *Zeitschrift für Physik* 9.1 (1922), pp. 131–144. DOI: 10.1007/bf01326962.

[2] Olivier Hardouin Duparc. "Pierre Auger — Lise Meitner: Comparative contributions to the Auger Effect." In: *International Journal of Materials Research* 100 (2009), pp. 1162–1166.

[3] Pierre Auger. "On secondary beta-rays produced in a gas by x-rays." In: *C.R.A.S.: Comptes Rendus des séances hebdomadaires de l'Académie des Sciences* 177 (1923), pp. 169–171.

[4] Leon Bess. "Possible Mechanism for Radiationless Recombination in Semiconductors." In: *Physical Review* 105.5 (1957), pp. 1469–1475. DOI: 10.1103/physrev.105.1469.

[5] H. W. B. Skinner. "The Soft X-Ray Spectroscopy of Solids. I. K- and L-Emission Spectra from Elements of the First Two Groups." In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 239.801 (1940), pp. 95–134. DOI: 10.1098/rsta.1940.0009.

[6] H. Fröhlich and J. O'Dwyer. "Time Dependence of Electronic Processes in Dielectrics." In: *Proceedings of the Physical Society. Section A* 63.2 (1950), p. 81. DOI: 10.1088/0370-1298/63/2/301.

[7] L Pincherle. "Auger Effect in Semiconductors." In: *Proceedings of the Physical Society. Section B* 68.5 (1955), pp. 319–320. DOI: 10.1088/0370-1301/68/5/108.

[8] N Sclar and E Burstein. "Optical and Impact Recombination in Impurity Photoconductivity in Germanium and Silicon." In: *Phys. Rev.* 98.6 (1955), pp. 1757–1760. DOI: 10.1103/physrev.98.1757.

[9] A. R. Beattie and P. T. Landsberg. "Auger effect in semiconductors." In: *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 249.1256 (1959), pp. 16–29. URL: http://rspa.royalsocietypublishing.org/content/249/1256/16.short%202014-02-20.

[10] Alan R Beattie. "A lifetime in Auger transitions." In: *Journal of Physics and Chemistry of Solids* 49.6 (1988), pp. 589–597. URL: http://scholar.google.com/scholar?output=instlink%5C&nossl=1%5C&q=info:CEktiwWtlu0J:scholar.google.com/%5C&hl=en%5C&as_sdt=0,23%5C&scillfp=14492993888134881147%5C&oi=lle.

[11] Sergey Karpov. "ABC-model for interpretation of internal quantum efficiency and its droop in III-nitride LEDs: a review." In: *Optical and Quantum Electronics* 47.6 (2015), pp. 1293–1303. DOI: 10.1007/s11082-014-0042-9.

[12] J Hader, JV Moloney, and SW Koch. "Supression of carrier recombination in semiconductor lasers by phase-space filling." In: *Applied Physics Letters* 87.20 (2005), p. 201112. URL: https://aip.scitation.org/doi/full/10.1063/1.2132524.

[13] David Aurelien et al. "All-optical measurements of carrier dynamics in bulk-GaN LEDs: Beyond the ABC approximation." In: *Applied Physics Letters* 110.25 (2017), p. 253504. DOI: 10.1063/1.4986908.

[14] Guan-Bo Lin et al. "Analytic model for the efficiency droop in semiconductors with asymmetric carrier-transport properties based on drift-induced reduction of injection efficiency." In: *Applied Physics Letters* 100.16 (2012), p. 161106. DOI: 10.1063/1.4704366.

[15] Emmanouil Kioupakis et al. "First-principles calculations of indirect Auger recombination in nitride semiconductors." In: *Physical Review B* 92.3 (2015). DOI: 10.1103/physrevb.92.035207.

[16] Eugen Merzbacher. *Quantum Mechanics*. Wiley, 1997, p. 672.

[17] Justin Iveland et al. "Direct Measurement of Auger Electrons Emitted from a Semiconductor Light-Emitting Diode under Electrical Injection: Identification of the Dominant Mechanism for Efficiency Droop." In: *Phys. Rev. Lett.* 110 (2013), p. 177406. DOI: 10.1103/PhysRevLett.110.177406.

[18] M. Binder et al. "Identification of nnp and npp Auger recombination as significant contributor to the efficiency droop in (GaIn)N quantum wells by visualization of hot carriers in photoluminescence." In: *Applied Physics Letters* 103.7 (2013), p. 071108. DOI: 10.1063/1.4818761.

CHAPTER 4

Auger Recombination Code Development

While there are many commercial and open-source high-performance computing codes to use for DFT calculations, calculating other more specific parameters often require being written yourself. As part of my thesis, I substantially modified and updated early code written to calculate Auger recombination coefficients. This included updating the code with modern programming practices, making the code more user-friendly and improving the parallelization.

4.1 Code Summary

Laks, Neumark, and Pantelides [1] first calculated direct Auger recombination coefficients in silicon using the empirical pseudopotential method. As DFT became more popular, it was used to calculate direct Auger recombination coefficients in GaAs, InGaAs[2] and InGaN.[3] Phonon-assisted Auger recombination was calculated using DFT for InGaN.[4]

At the end of the previous chapter, we wrote down the equations for calculating the direct and phonon-assisted Auger processes (eq. 3.3 and eq. 3.5). These equations require us to know the wave functions of the electrons, the energies of the system, and (for the phonon-assisted process) the phonon frequencies. This information can be obtained from the DFT calculations we discussed in chapter 2, but we still need tell the computer how to take that information and use it to calculate the Auger recombination coefficients. A part of this thesis has been on improving the code we use to do this.

The Auger recombination code is written in Fortran90 and uses the output from quantum espresso[5,6] and Wannier90[7,8] along with some of the libraries from FFTW[9] and LAPACK.[10] There are around 5000 lines of code separated into three directories: `modules` which consists of code to deal with input and output, parallelization, and the matrix elements and screening (which are common to both the direct and phonon codes). `direct` which consists of code to calculate

direct *eeh* and *hhe* Auger recombination coefficients and code that finds the reduced k-point lists[1]. Finally phonon, which calculates the phonon-assisted Auger recombination coefficients.

Before getting into specifics about the code, there are some approximations that make the Auger calculations more feasible that we need to go over first.

4.2 Approximations for Tractable Calculations

To calculate the Auger recombination coefficients we must split the unit-cell into discrete k-points and then sum over them. For each k-point we must calculate the energies and wave functions of the system using the Wannier interpolation method. While using Wannier functions helps avoid very costly non-self consistent field (NSCF) DFT calculations, it still is expensive to do these calculations for the entire Brillouin zone (BZ). A way to speed up the calculation with little loss of accuracy is to only consider the k-points *near the band extrema*. These are the most important k-points, and eliminating the rest of the BZ will not substantially affect the results.

For this reason, we begin the direct Auger calculations by calculating the quasi-Fermi energy of electrons and holes. We then determine a cutoff energy which limits the number of k-points in the calculation, $E_{\text{cut}} = \epsilon_F + NkT$, where N is an integer. We determine N by calculating the electron density and comparing it to the original desired electron density. If the difference is less than 1%, we keep that value for E_{cut}, otherwise we increase N and try again. Only the k-points that have bands within this cutoff energy are kept.

We also define the hole k-point grid to be half as fine as the grid for the electron calculations. The group-III nitrides in particular have heavier hole masses and do not need to be sampled as finely. This speeds up calculations slightly, and since we converge the Auger coefficients anyway should not impact the accuracy.

Phonon calculations are greatly simplified by only considering transitions from the Γ point. This assumption comes from the same justification as keeping only 99% of the electron density – most of the carriers in the conduction band are near Γ and we don't need to consider carriers too far away from this point. In the phonon case, we take a more extreme approach simply because of the increased computational complexity to adding more k-points.

The screened-Coulomb-interaction requires $\epsilon(q)$. This *could* be calculated using GW, but this has not been implemented yet (see chapter 8). Instead, a model dielectric function is used.[11] This model dielectric function is given by:

$$\epsilon(q) = 1 + \left[\frac{1}{\epsilon_\infty - 1} + \alpha \left(\frac{q}{q_{\text{TF}}} \right)^2 + \frac{h^2 q^4}{4 m^2 \omega_p^2} \right]^{-1}$$

[1] Usually referred to as *pre-direct* routines.

Where ϵ_∞ is the dielectric constant at high frequency, q_{TF} is the Thomas-Fermi wave vector, and ω_p is the plasma frequency. α is determined empirically, but has been shown to be accurate for the group-III nitrides.

Finally, a smaller point is converting the delta-functions in into a finite-width Gaussian function. For each Auger calculation we use a variety of widths for this Gaussian with larger widths being used to smooth out the results for under-converged calculations (to observe general behavior) and smaller widths used for highly converged calculations.

4.3 Refactoring and Modularization using Modern Programming Practices

Fortran is one of the first programming languages made, but development of the language continues to this day.[12] Fortran90 in particular introduced modular programming, allowing for code to be easily written in separate parts and interfaced together. One of the first goals of this thesis was to make the Auger code easier to make modifications to. The code initially consisted of many repeated sections. Meaning any change to one of those sections had to be remembered and copied back into all the other areas. This made changing the code prone to errors both from typos and from forgetting exactly where the copied code was throughout the various files. This situation is a classic example of code refactoring and modularization[13] – where the repetitive portions of the code are put into "modules" where they can be called from other parts of the code. This means there is only one place to go to make changes at the cost of needing to properly link the files together.

The `wannier_module` consolidates all uses of wannier information including: reading and writing eigenvalues, reading u_matricies, and performing the Wannier interpolation to fine sets of k points. These functions are used throughout the code and consolidating them makes changing them much simpler. `input_parameters` allows for the use of common user parameters in the code without having to manually change the source code and recompile it. These parameters are input as fortran namelists and read through the `read_namelists` module. The `screening` and `matrix_elements` modules simplify the repetitive calculation of matrix elements, a task that is common to both the direct and phonon code. Finally there are various minor modules like the `kinds` module, which ensures that the computational precision is uniform throughout the code, the `parameters` modules ensures that common parameters and conversion factors are the same and can be updated easily.

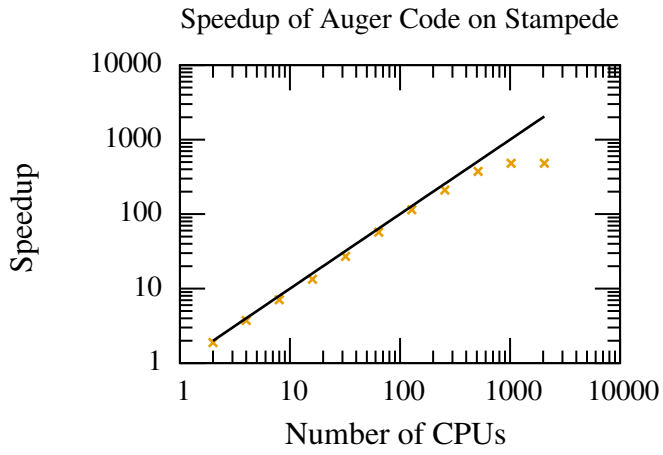

Figure 4.1 The speedup of the direct Auger code. We see approximately linear scaling until a few hundred CPUs.

4.4 Parallelism and Optimization

Currently, the pre_direct code is serial because it does not take a substantial amount of time to run. Parallelizing the direct code is somewhat tricky. The code must check the energies at each k point to see if they are within the cutoff generated in the pre_direct code. If they are not, the Auger coefficients should not be calculated for that case. Because of this, processors which are evaluating points far away from Γ wind up with less work than processors near Γ. A simple way to fix this involves running the code twice. The first run does no actually Auger calculations and only finds the total number of combinations of energies and k points are valid (band energies within the cutoff). This total number of combinations is easier to distribute among many processors evenly, since each combination will be a similar amount of work.

Rough tests of the parallelization done for a simple InN system on the XSEDE Stampede system (from chapter 5) are shown in fig. 4.1 and fig. 4.2.

The phonon-assisted code is simpler to parallelize over phonon points, because they are determined explicitly before running the calculation. Ensuring that the number of processors evenly divides the number of phonon points should ensure a balanced workload over the processors. A detailed study of the parallelization of the phonon code has not yet been done, since the major effort required is in the determination of the electron-phonon coupling rather than the explicit calculation of the phonon-assisted coefficients. # Running Auger Calculations The following is required to run a direct Auger recombination calculation:

- an eigenvalue file from a Wannier90 run (either DFT or GW)
- u matrix files (`u_matrix.dat` and `u_matrix_opt.dat`) from custom wannier-1.1 code

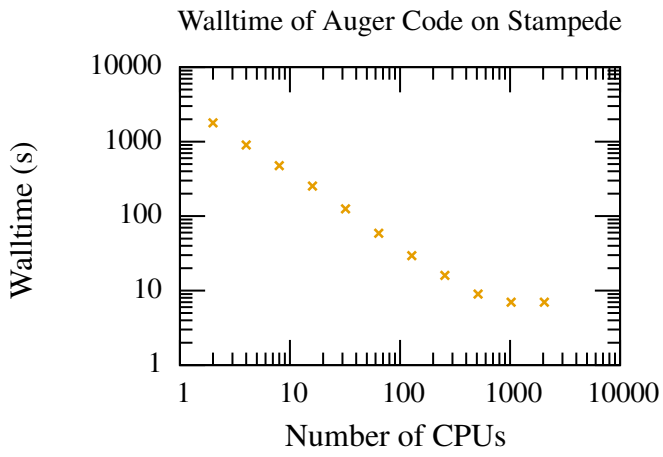

Figure 4.2 The improvement in walltime of the direct Auger code. Parallelism shortens the test run from nearly twenty minutes to ten seconds.

- a converged scf calculation from quantum espresso

In addition to the above, to run a phonon calculation:

- phonon modes
- electron-phonon coupling from custom quantum espresso code

And finally, each part of the code uses an input file defined as:

```
&CONTROL                  ! Define the type of run and information about the run
prefix=STRING             ! Name of run, must use the same as Quantum Espresso
calculation=STRING        ! Type of run which specifies namelists read:
                          ! options are: pre_direct, auger, phonon
density_e=REAL            ! Electron and Hole density in cm^-3
density_h=REAL
temperature=REAL          ! Electronic temperature in kelvin
nkint(:) = INTEGER ARRAY ! Kmesh
scratch_dir=STRING        ! Where to write output
pre_direct_dir=STRING     ! Where pre_direct files are located
/
&MATERIAL                 ! Material parameters
alat=REAL                 ! Lattice Parameter
Vcell=REAL                ! Volume of Unit Cell
nbnd=INTEGER              ! Number of bands
ivbm=INTEGER              ! Valence band index
icbm=INTEGER              ! Conduction band index
vbmdeg=INTEGER            ! Degeneracy of vbm
```

```
valence_electrons=INTEGER ! Number of valence electrons
nphonon=INTEGER           ! Number of phonon modes
epsilon_infty=REAL        ! high frequency dielectric constant
a(:,1) =                  ! Matricies defining unit cell in real and reciprocal space
a(:,2) =
a(:,3) =
b(:,1) =
b(:,2) =
b(:,3) =
/
&PRE_DIRECT               ! Parameters for pre_direct calculation
halfholegrid = BOOLEAN    ! Use nkint/2 for holes
density_conv_thr = REAL   ! Threshold for carrier density in determining E_cut
/
&WANNIER                  ! Parameters from Wannier90
nwan=                     ! Number of wannier functions
nwanband=                 ! Number of wannier bands
nkwan1=                   ! Wannier90 K mesh
nkwan2=
nkwan3=
/
&AUGER                    ! Auger parameters
ngap=INTEGER              ! Number of gaps to scissors shift to
Egap_min=REAL             ! Smallest gap
Egap_max=REAL             ! Largest gap
Egap_step=REAL            ! Step size for increasing gap
nk=INTEGER                ! Number of phonons
direct_first_run=BOOLEAN ! Calculate combinations or calculate Auger
screening_model=STRING    ! Choice of screening models:
                          ! TF (Thomas Fermi), zero (No screening)
                          ! and combined (Thomas Fermi and Debye Huckle
                          ! based on Fermi Energy)
/
```

References

[1] DB Laks, GF Neumark, and ST Pantelides. "Accurate interband-Auger-recombination rates in silicon." In: *Physical Review B* 42.8 (1990), p. 5176. URL: https://link.aps.org/pdf/10.1103/PhysRevB.42.5176.

[2] S Picozzi et al. "Accurate first-principles detailed-balance determination of Auger recombination and impact ionization rates in semiconductors." In: *Physical review letters* 89.19 (2002), p. 197601. URL: https://link.aps.org/pdf/10.1103/PhysRevLett.89.197601.

[3] Kris T Delaney, Patrick Rinke, and Chris G Van de Walle. "Auger recombination rates in nitrides from first principles." In: *Applied Physics Letters* 94.19 (2009), p. 191109. DOI: 10.1063/1.3133359.

[4] Emmanouil Kioupakis et al. "Indirect Auger recombination as a cause of efficiency droop in nitride light-emitting diodes." In: *Applied Physics Letters* 98.16 (2011), p. 161107. URL: https://aip.scitation.org/doi/full/10.1063/1.3570656.

[5] Paolo Giannozzi et al. "QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials." In: *Journal of Physics: Condensed Matter* 21.39 (2009), p. 395502. DOI: 10.1088/0953-8984/21/39/395502.

[6] P Giannozzi et al. "Advanced capabilities for materials modelling with Quantum ESPRESSO." In: *J Phys Condens Matter* 29.46 (2017), p. 465901. DOI: 10.1088/1361-648X/aa8f79.

[7] Arash A. Mostofi et al. "wannier90: A tool for obtaining maximally-localised Wannier functions." In: *Computer Physics Communications* 178.9 (2008), pp. 685–699. DOI: 10.1016/j.cpc.2007.11.016.

[8] Arash A. Mostofi et al. "An updated version of wannier90: A tool for obtaining maximally-localised Wannier functions." In: *Computer Physics Communications* 185.8 (2014), pp. 2309–2310. DOI: 10.1016/j.cpc.2014.05.003.

[9] M. Frigo and S.G. Johnson. "The Design and Implementation of FFTW3." In: *Proceedings of the IEEE* 93.2 (2005), pp. 216–231. DOI: 10.1109/jproc.2004.840301.

[10] E. Anderson et al. *LAPACK Users' Guide*. Philadelphia, PA: Society for Industrial and Applied Mathematics, 1999.

[11] G. Cappellini et al. "Model dielectric function for semiconductors." In: *Physical Review B* 47.15 (1993), pp. 9892–9895. DOI: 10.1103/physrevb.47.9892.

[12] International Organization for Standardization. *Fortran 2018*. URL: https://wg5-fortran.org/f2018.html.

[13] Robert C. Martin. *Clean Code*. Pearson Education, 2009. URL: http://books.google.com/books?id=hjEFCAAAQBAJ%5C&hl=%5C&source=gbs_api.

CHAPTER 5

Radiative and Auger Recombination Processes in Indium Nitride

InN and In-rich InGaN alloys emit in the infrared range desirable for telecommunications applications. However, the droop problem reduces their efficiency at high power. Nonradiative Auger recombination is a strong contributor to this efficiency loss. Here we investigate radiative and Auger recombination in InN and In-rich InGaN with first-principles calculations. We find that the direct *eeh* process dominates Auger recombination in these materials. In the degenerate carrier regime, the Auger and radiative rates are suppressed by different mechanisms: the radiative rate is affected by phase-space filling while Auger recombination is primarily reduced by free-carrier screening. The suppression of the radiative rate onsets at lower carrier densities than the Auger rate, which reduces the internal quantum efficiency of InN devices. Droop in InN can be mitigated by increasing the band gap through alloying with GaN. We demonstrate that the peak efficiency of $In_{0.93}Ga_{0.07}N$ alloys (which emit at 1550 nm) is 33% higher than InN and occurs at higher carrier densities.

5.1 Indium Nitride for Telecommunication

Growing high quality indium nitride proved to be a challenge until the early 2000's. Many fundamental properties were unknown or incorrect because of the low quality samples being studied. In 2002 Davydov et al. [1], first reported that the band gap of InN was much lower than commonly claimed.[1] This began revisions for the band gap from 1.8 - 2.1 eV eventually down to 0.6 - 0.8 eV and extending the emission range of the group-III nitrides into the IR regime.

IR wavelengths are used for telecommunication – transmitting information over short and long distances. For these applications, fast, secure and efficient light emission is necessary. There

[1]The issue being the Burnstein-Moss effect, where the doping is so high that the conduction band is completely filled and optical expedients are not seeing the band gap, but the band gap plus a shift[2]

is increasing interest specifically in single-photon emission light sources that can be included in existing systems. Single-photon light sources use the basic properties of quantum mechanics for highly secure and efficient communication communications.

Indium nitride or In-rich InGaN has many appealing properties for these applications. For single-photon emission, they have strong quantum confinement effects leading to bound excitons at room temperature and tunability through nanostructures. They also can can be alloyed with GaN and therefore have gaps tuned across a range of the infrared spectrum and covering both the major IR bands of interest for telecommunication (1300 nm and 1550 nm). In addition, they are resistant to radiation, are non-toxic, have fast switching speeds, and can handle high-frequency operation.[3–5]

For these light emission applications understanding the IQE is vital. Because InN has a much smaller band gap than GaN, we exepect the direct Auger process to be stronger because of Auger's exponential dependence on the band gap.[6]

5.2 Review of Experimental Literature

Chen et al. [7] studied three unintentionally doped samples of InN at 20K and 300K. The free-carrier densities vary over about one order of magnitude from 1.3×10^{18} to 1.2×10^{19} cm^{-3}. This free-carrier density dependence is used to study the density dependence of the recombination rates in the material using time-resolved differential transmission techniques. For both temperatures, the lifetimes of carriers are shown to decrease linearly with increasing carrier density. Because of the linear dependence, the authors rule out the possibility of Auger recombination. However, at degenerate carrier densities like those considered in the experiment, a linear dependence on Auger is expected due to phase-space filling effects.

This discrepancy comes up later in the paper as they calculate theoretical radiative recombination lifetimes of carriers for the given carrier densities and temperatures. They find good agreement at low temperatures (when non-radiative recombination is suppressed) but poor agreement at 300K. They attribute this discrepancy to nonradiative defect recombination rather than degenerate Auger recombination.

Using time-resolved reflectivity experiments, Ascázubi et al. [8] studied Auger in both unintentionally doped and Si-doped InN. Combining data from their own experiments with the data from Chen et al. [9], they also found that carrier lifetime depended on the inverse of carrier density from 1×10^{17} to 4×10^{20} cm^{-3}. However, they also assumed that Auger recombination cannot have a linear dependence on carrier density and attributed the nonradiative recombination to impurity trapping rather than degenerate Auger recombination.

Tsai, Chang, and Gwo [10] was the first to attribute a linear dependence on carrier lifetime to

Auger recombination with degenerate carriers. They used pump-probe, time-resolved reflectivity measurements to determine Auger rates in degenerate InN samples. Their data was fit using a biexponential function, establishing slow and fast relaxation times for the excited carriers. The authors justify hot electron relaxation for the fast recombination time and then calculate a theoretical radiative rate was calculated based on $k \cdot P$ theory.[11] This rate was found to be too slow to account for the slow relaxation time, and they attributed this relaxation time to the non-radiative Shockley-Reed-Hall and Auger recombination effects. By fitting their data using an ABC model for *degenerate* carriers, $1/\tau_{\text{slow}} = B_0 + B_2 * n$ they find both a defect recombination and Auger recombination lifetime (515 ps for the former and $1/\tau_{\text{Auger}} = 2.4 \times 10^{-10}\text{cm}^3\text{s}^{-1} * n$ for the latter.

The following year, Jang et al. [12] used time-resolved photoluminescence to study Auger recombination for a variety of temperatures in two degenerately doped InN samples. The total decay rate of carriers was fitted to an non-degenerate ABC model where all three rate coefficients were obtained from the fit at 35 K. At higher temperatures, the fit is semiempirical, with the radiative rate fitted using the theoretical $T^{3/2}$ dependence while the Auger and SRH coefficients are fit to experimental data. They fit the Auger temperature dependence with an exponential activation energy function $C = C_0 \exp -E_a/k_B T$, with different activation energies for the samples. However, they attribute the weak temperature dependence and n^2 dependence to mean that Auger is a phonon-assisted process rather than the weaker degenerate direct process. They find values for Auger recombination on the order of $10^{-29}\text{cm}^6\text{s}^{-1}$ for the Auger coefficients at 300K.

Nargelas et al. [13] used differential transmission methods to study recombination lifetimes in two unintentionally doped InN samples. Their experiment uses probe energies from below to above the band gap for the first time, which allows free-carrier absorption kinetics to be studied directly since only photoexcited carriers are observed and "optical bleaching"[^bleaching] does not occur. Their data show a linear dependence of lifetime on carrier density which they attribute to trap-assisted Auger recombination rather than degenerate Auger recombination. This is justified by the sample with higher defect density exhibiting faster recombination lifetimes.

Cho et al. [14] studied degenerately doped InN samples using photoluminescence spectra of the samples at room temperature. Their goal was to investigate how PL intensity changed with both dislocation density and with background carrier density. They found that PL intensity and dislocation density are unrelated but that PL intensity decreases with increasing carrier concentration – strongly pointing to the Auger effect being the determining factor in the efficiency. To find the Auger coefficient, they fit lifetimes using an ABC model where the fitting parameters were the Auger coefficient, C, and the density scaling. The defect density was measured, while the radiative lifetime was calculated from the dipole transition rate, which was originally calculated by Tsai, Chang, and Gwo [10]. However the original calculation assumed an absorption edge of 0.8 eV for the samples considered in that paper and Cho et al. [14]'s samples show a range from

0.67 - 0.77 eV, adding some uncertainty to the fitting process. Ultimately the fit gave a degenerate Auger recombination coefficient of 4.5×10^{-9} $cm^3s^{-1}$and a quadratic density dependence for the Auger rate – as expected for the degenerate Auger effect.

Finally, Seetoh et al. [15] used steady-state photoluminescence measurements at various temperatures to study Auger and SRH recombination. While this method does not allow an absolute value of the Auger recombination rate to be determined, it does point to a specific dominant mechanism in the material. For their samples (what are they, more than one?) they find that Auger is dominant at low T with an activation energy of 10-17 meV, but that SRH is dominant at higher temperatures as expected. However, because Auger is an intrinsic property of the material it would present an upper limit to the IQE.

5.3 Results

In this work, we performed first-principles calculations to understand the radiative and Auger recombination properties of bulk indium nitride as a function of free-carrier density. We found that the dominant Auger mechanism is the direct *eeh* process. We also uncovered that for degenerate carriers, screening between carriers is the primary mechanism of reducing the Auger rate power law, rather than phase-space filling. We further found that the radiative rate is suppressed by phase-space filling at lower carrier densities than the Auger rate is suppressed by screening or phase-space filling, which is detrimental to the efficiency of devices. Finally, alloying InN with GaN to increase the band gap reduces the Auger rate and increases the internal quantum efficiency.

5.3.1 Calculation Parameters

For these calculations, we performed plane-wave norm-conserving pseudopotential DFT calculations with the local density approximation (LDA)[16] for the exchange-correlation functional and the Quantum Espresso[17,18] code. We found relaxed lattice parameters of a = 6.69 and c = 10.83 Bohr, which are within (%%). For the band structure, we included a Hubbard U correction (LDA+U) for the In $4d$ and the N $2p$ orbitals to avoid the artificial closing of the gap of InN by LDA and the subsequent unphysical mixing of valence and conduction states near Γ. Our U parameters (U_p = 1.5 eV and U_d = 6.0 eV) were obtained in Ref..[19]

Even with the U correction, LDA does not correctly predict the band gap of InN, which is needed for recombination rate calculations. We therefore perform G_0W_0 calculations[20] with the BerkeleyGW code[21] to correct the LDA+U eigenvalues. Our calculated electron effective mass (0.07 m_e) agrees with experiment[22] Our calculated G_0W_0 band gap of InN is 0.61 eV, which

agrees with the experimental range of measurements (0.6-0.8 eV)[1] that depend on doping and temperature. Since the gap is very sensitive to small changes in environment, we use our rigid shift of the band gap to account for these sample-dependent gap variations on the recombination rates. Our gap varies over the 0.5-0.8 eV range. This adjustment of the gap value further allows us to assess the convergence of out rate calculations, and to simulate the effect of alloying with GaN.

For the recombination coefficient calculations, we again use the methods discussed in chapter 4. In this case, we use a $100 \times 100 \times 50$ grid for the *eeh* and an $80 \times 80 \times 40$ grid for the *hhe* process. The fineness of these grids is necessary primarily for low carrier densities, where many lower lying states are empty and many more Auger transitions are possible. Because the *hhe* process converges faster we do not need to use as fine grids in the *hhe* calculation. Gaussian functions with a width of 0.1 eV are used to evaluate the energy delta functions in the Auger rates. The experimental high frequency dielectric constant of InN ($\epsilon_\infty = 8.4$).[23] The lattice and electronic temperatures are set to 300K.

5.3.2 Dominant Auger Mechanism

We expect that because of the small band gap of InN the direct Auger process will be dominant. Auger depends on the band gap as $\propto \exp(-E_g/kT$.[6] However, it is not as clear from the band structure as in 7 whether or not the *eeh* or *hhe* process will be stronger.

These physical arguments are validated by our calculations for InN (Fig.5.1). The direct *eeh* Auger process is the largest contributor by a factor of 9 over the direct *hhe* process. The phonon-assisted *eeh* and *hhe* processes are both much weaker than the direct *eeh* one by approximately two orders of magnitude. Therefore, direct *eeh* Auger dominates in InN over the entire range of band-gap values we examined (0.5-0.8 eV), and should also dominate in high-In-content InGaN.

5.3.3 Carrier Density Dependence

After identifying the dominant Auger process, we examine how the radiative and Auger recombination coefficients vary with carrier density (Fig.5.2). At low carrier density, both recombination coefficients are constant with respect to density. But as the density increases (and the carriers become degenerate), both coefficients become decreasing functions of the density, but with different characteristic density values. The radiative rate declines at a carrier density of roughly one order of magnitude smaller than the Auger rate. The efficiency of devices in this regime would suffer because of high Auger recombination but suppressed radiative recombination.

Because the Auger rate is not suppressed at the same carrier density as the radiative one, a mechanism other than phase-space filling may be responsible. To understand whether phase-

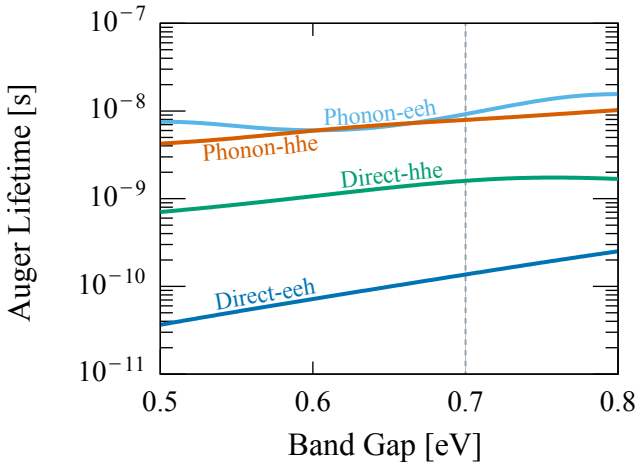

Figure 5.1 Auger lifetimes of InN as a function of band gap for the various Auger processes and for a carrier density of 10^{19}cm^{-3}. The rigid gap adjustment accounts for the experimental variation of the InN gap, as well as simulates alloying with GaN. Direct *eeh* Auger dominates throughout the 0.5-0.8 eV gap range, as expected for narrow-gap semiconductors.

space filling or screening impacts the Auger rate most, we performed calculations both with and without screening of the Coulomb interaction by free carriers. To understand the density dependence analytically, we fit our data using (Ref.[24]):

$$B(n) = \frac{B_0}{1 + (n/n_0)^b}, \tag{5.1}$$

where $B(n)$ is the density-dependent radiative coefficient, B_0 is the coefficient in the low-density limit, n_0 is the characteristic density for the onset of phase-space filling or screening, and the exponent b is a dimensionless exponent approximately equal to 1. A similar equation to Eq.5.1 is used to fit the density-dependent Auger $C(n)$ coefficients. In contrast to previous work,[25,26] here we assume the n_0 and b parameters to be different for the $B(n)$ and $C(n)$ fits. The fitted values %for the radiative and Auger coefficients are listed in Table 5.1.

The radiative coefficient deviates from the non-degenerate constant value at a characteristic density $n_0 = 2.1 \times 10^{18}$cm^{-3}, which is approximately one order of magnitude lower than the characteristic density of the *screened* Auger coefficient ($n_0 = 1.9 \times 10^{19}$cm^{-3}). In comparison, the *unscreened* Auger coefficient (which is limited at high densities only by phase-space filling) has a characteristic density that is one further order of magnitude higher (2.5×10^{20}cm^{-3}). Meaning that the primary cause for the decline of the Auger coefficient is the screening of the Coulomb interaction, while phase-space filling is a secondary effect. We note that a weaker density dependence for the C coefficient than the B coefficient has been reported in semipolar InGaN wells.[27] The unscreened Auger coefficient exhibits a small increase before the effects of phase-space fill-

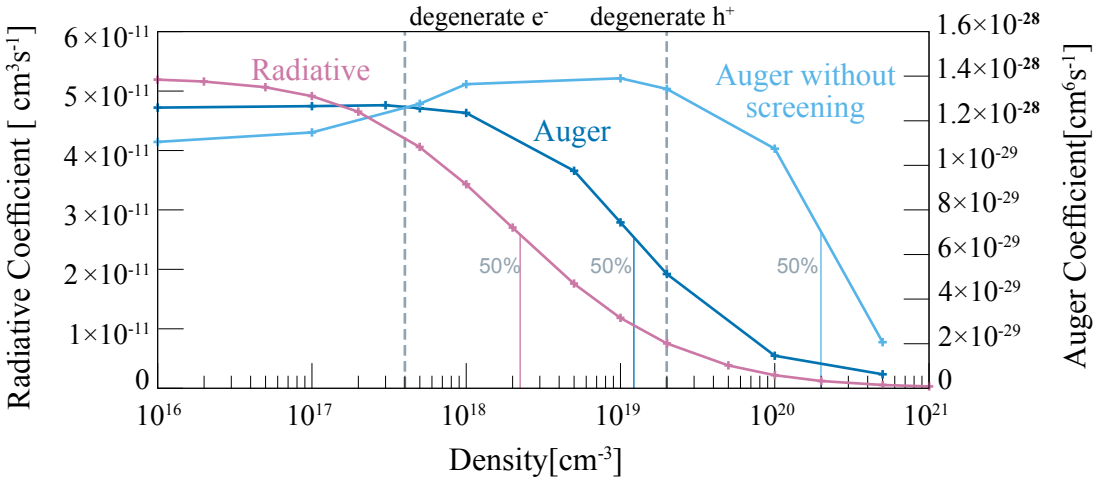

Figure 5.2 Variation of the radiative and Auger (both with and without free-carrier screening) coefficients as a functions of free-carrier density. The vertical dashed lines denote the densities at which carriers become degenerate, while the vertical solid lines indicate the densities at which the coefficients are reduced to 50% of their non-degenerate values. The onset of the reduction of the radiative coefficient occurs at lower carrier densities than the Auger coefficient, which is detrimental to the IQE of devices in this carrier-density range.

ing ultimately cause the rate to decline, a behavior which is also observed in the theory work of Ref.[28] for GaSb. We expect that our conclusions on the density-dependence of the coefficients, which were derived for bulk InN, also directly apply to nonpolar quantum wells. The physics is more complex in polar and semipolar wells, in which the strong polarization fields that separate carriers and reduce the recombination rates are also screened by free carriers. However, since the polarization fields affect both the radiative and the Auger rate proportionately,[29] we expect that our conclusions about the relative importance of phase-space filling and the screening of the Coulomb interaction should also apply to the radiative and Auger coefficients of polar and semipolar wells, once their values are corrected by the overlap of the electron and hole envelope functions.

Our fits are compared to experimental Auger lifetimes in Fig.5.3, which range from 20 to 1200 ps[7,10,12,14] for carrier densities in the 10^{18} to 10^{19} cm^{-3} range. Our values at a carrier density of 10^{19} cm^{-3} lie within this range (129 ps for unscreened Auger, 72 ps for screened Auger). Most experiments measured the Auger rate as having a n^2 dependence,[7,10,14] in agreement with our calculations. Moreover, although Chen et al. [7] ruled out the possibility of Auger recombination from their lifetime measurements, we find that their measured lifetimes are two orders of magnitude shorter than our calculated radiative values, and are more consistent with our Auger data.

Finally, we examine how the Auger and radiative rates change when InN is alloyed with small

Table 5.1 The parameters used to fit the density dependence of the radiative and Auger coefficients according to Eq.5.1.

	Prefactor	n_0 (cm^{-3})	b
Auger w/ Screening	1.1×10^{-28} cm^6s^{-1}	1.9×10^{19}	0.98
Auger w/o Screening	1.4×10^{-28} cm^6s^{-1}	2.5×10^{20}	1.5
Radiative	5.2×10^{-11} cm^3s^{-1}	2.1×10^{18}	0.83

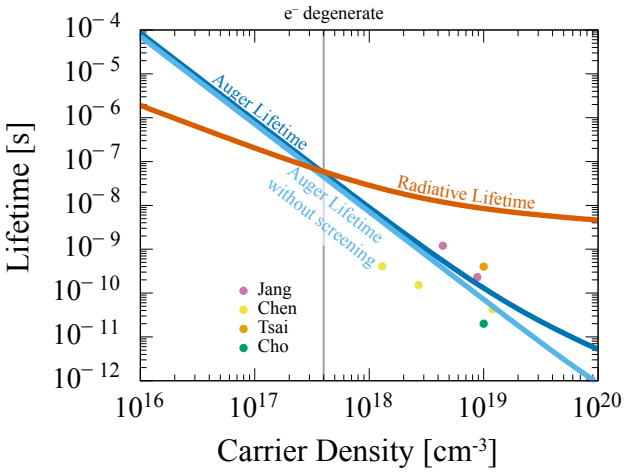

Figure 5.3 Comparison of calculated Auger and radiative lifetimes with experimental values, which vary from 20 to 1200 ps in the 10^{18} to 10^{19} cm^{-3} density range (Refs.[7,10,12,14]). Our calculated lifetimes for both the screened and unscreend Auger processes lie within the range of experimental measurements. Auger recombination becomes stronger than the radiative process at a carrier density of 4×10^{17} cm^{-3}, at which coincidentally electrons become degenerate.

amounts of GaN. Alloying InN with 7% GaN increases the gap to 0.8 eV and leads to emission at 1550 nm, an important optical fiber telecommunication wavelength. The Auger coefficient also decreases exponentially with increasing band gap. The internal quantum efficiency (IQE) is given by:

$$\eta = \frac{B(n)n^2}{An + B(n)n^2 + C(n)n^3}.$$

For binary InN we use our model Auger and radiative coefficients presented earlier, while for the $In_{0.93}Ga_{0.07}N$ alloy, we used Auger values obtained for InN with the band gap rigidly increased to 0.8 eV. At a carrier density of 10^{18} cm^{-3}, the C coefficient decreases from 1.2×10^{-28} cm^6s^{-1} for binary InN to 5.0×10^{-29} cm^6s^{-1} for the $In_{0.93}Ga_{0.07}N$ alloy, i.e., a nearly 50% reduction. For A we use an experimental value for InGaN ($A \approx 6 \times 10^7$ s^{-1}).[30] The two IQE curves are shown in Fig.5.4. The IQE peak of $In_{0.93}Ga_{0.07}N$ is 33% higher than InN and occurs at a higher carrier density, which results in overall more efficient optoelectronic devices. Although we did not consider the effects of other possible energy-loss mechanisms in our analysis (e.g., carrier leakage), our results point

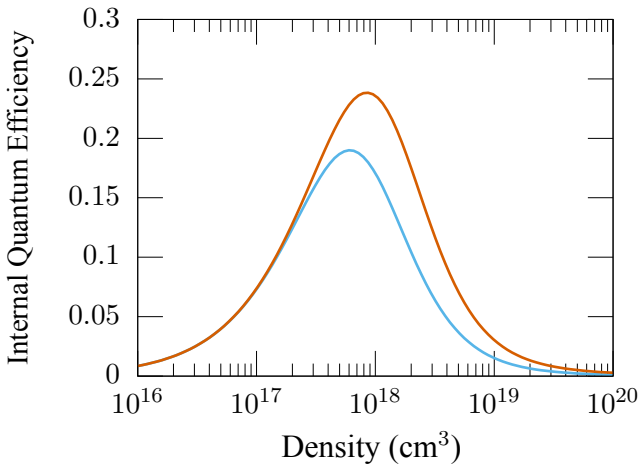

Figure 5.4 Simulated internal quantum efficiency versus carrier density for InN and $In_{0.93}Ga_{0.07}N$ devices. Alloying InN with 7% GaN increases the band gap to 0.8 eV (for light emission at 1550 nm), increases the maximum efficiency by 33%, and reduces the efficiency droop at high power.

to Auger recombination being a strong nonradiative loss mechanism in InN and In-rich InGaN optoelectronic devices and an important source of efficiency droop.

Although our analysis focused on bulk InN, our conclusions (dominant role of direct *eeh* Auger, relative importance of phase-space filling and carrier screening) must also hold for In-rich InGaN alloys and for quantum-well structures, with best quantitative accuracy for thick nonpolar wells and for high-In-content alloys. For polar or atomically thin wells, the polarization fields and quantum confinement need to also be considered.[31] On the other hand, the effects of confinement and alloying on Auger recombination are much stronger for GaN[31,32] than for InN, since Auger in GaN is weak and only enabled by lack of momentum conservation (due to phonons, alloy disorder, confinement etc.). For InN and In-rich InGaN, however, direct Auger is allowed and we anticipate that confinement and alloy disorder have only a minor quantitative effect on the Auger rates. Further work is needed to fully assess Auger recombination in quantum-confined InN wells and in In-rich InGaN alloys that include composition fluctuations and carrier localization.[33]

In conclusion, we found that direct *eeh* Auger is the dominant Auger process in bulk InN. Carrier degeneracy impacts Auger and radiative recombination differently. While the radiative coefficient declines because of phase-space filling, the Auger coefficient is primarily reduced by free-carrier screening at higher carrier densities. This effect suppresses the radiative rate at lower carrier densities than Auger recombination and reduces the IQE of devices. Alloying InN with GaN to increase the band gap is an effective method to reduce the Auger rate and increase the IQE, while also shifting the emission wavelength to the 1550 nm telecommunications range.

References

This chapter was originally published in part in Applied Physics Letters:
Andrew McAllister, Dylan Bayerl, and Emmanouil Kioupakis. "Radiative and Auger recombination processes in indium nitride." In: *Applied Physics Letters* 112.25 (2018), p. 251108. DOI: 10.1063/1.5038106

[1] V.Yu. Davydov et al. "Absorption and Emission of Hexagonal InN. Evidence of Narrow Fundamental Band Gap." In: *physica status solidi (b)* 229.3 (2002), r1–r3. DOI: 10.1002/1521-3951(200202)229:3<r1::aid-pssb99991>3.0.co;2-o.

[2] J. Wu et al. "Effects of the narrow band gap on the properties of InN." In: *Physical Review B* 66.20 (2002), p. 201403. DOI: 10.1103/physrevb.66.201403.

[3] Andreas Winden et al. "Spectral sensitivity tuning of vertical InN nanopyramid-based photodetectors." In: *Japanese journal of applied physics* 52.8S (2013), 08JF05. DOI: 10.7567/JJAP.52.08JF05.

[4] *InGaN mesoscopic structures for low energy consumption nano-opto-electronics*. IEEE, 2016, pp. 69–72. DOI: 10.1109/asdam.2016.7805897.

[5] Md Mobarak Hossain Polash, M Shah Alam, and Saumya Biswas. "Design and analysis of InN- In 0.25 Ga 0.75 N single quantum well laser for short distance communication wavelength." In: *Optical Engineering* 57.3 (2018), p. 036110. DOI: 10.1117/1.OE.57.3.036110.

[6] A. R. Beattie and P. T. Landsberg. "Auger effect in semiconductors." In: *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 249.1256 (1959), pp. 16–29. URL: http://rspa.royalsocietypublishing.org/content/249/1256/16.short%202014-02-20.

[7] Fei Chen et al. "Temperature dependence of carrier lifetimes in InN." In: *physica status solidi (a)* 202.5 (2005), pp. 768–772. DOI: 10.1002/pssa.200461501.

[8] Ricardo Ascázubi et al. "Ultrafast recombination in Si-doped InN." In: *Applied physics letters* 88.11 (2006), p. 112111.

[9] Fei Chen et al. "Ultrafast carrier dynamics in InN epilayers." In: *Journal of crystal growth* 269.1 (2004), pp. 10–14. DOI: 10.1016/j.jcrysgro.2004.05.028.

[10] Tsong-Ru Tsai, Chih-Fu Chang, and S Gwo. "Ultrafast hot electron relaxation time anomaly in InN epitaxial films." In: *Applied physics letters* 90.25 (2007), p. 252111. DOI: 10.1063/1.2751110.

[11] Matsusue Toshio and Sakaki Hiroyuki. "Radiative recombination coefficient of free carriers in GaAs-AlGaAs quantum wells and its dependence on temperature." In: *Applied Physics Letters* 50.20 (1987), pp. 1429–1431. DOI: 10.1063/1.97844.

[12] D-J Jang et al. "Auger recombination in InN thin films." In: *Applied Physics Letters* 92.4 (2008), p. 042101. DOI: 10.1063/1.2837537.

[13] S Nargelas et al. "Dynamics of free carrier absorption in InN layers." In: *Applied Physics Letters* 95.16 (2009), p. 162103. DOI: 10.1063/1.3251077.

[14] YongJin Cho et al. "Auger recombination as the dominant nonradiative recombination channel in InN." In: *Physical Review B* 87.15 (2013), p. 155203. DOI: 10.1103/physrevb.87.155203.

[15] IP Seetoh et al. "Auger recombination as the dominant recombination process in indium nitride at low temperatures during steady-state photoluminescence." In: *Applied Physics Letters* 102.10 (2013), p. 101112. DOI: 10.1063/1.4795793.

[16] John P Perdew and Alex Zunger. "Self-interaction correction to density-functional approximations for many-electron systems." In: *Physical Review B* 23.10 (1981), p. 5048. DOI: 10.1103/PhysRevB.23.5048.

[17] Paolo Giannozzi et al. "QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials." In: *Journal of Physics: Condensed Matter* 21.39 (2009), p. 395502. DOI: 10.1088/0953-8984/21/39/395502.

[18] P Giannozzi et al. "Advanced capabilities for materials modelling with Quantum ESPRESSO." In: *J Phys Condens Matter* 29.46 (2017), p. 465901. DOI: 10.1088/1361-648X/aa8f79.

[19] A. Terentjevs et al. "Importance of on-site corrections to the electronic and structural properties of InN in crystalline solid, nonpolar surface, and nanowire forms." In: *Physical Review B* 82.16 (2010), p. 165307. DOI: 10.1103/physrevb.82.165307.

[20] Dylan Bayerl and Emmanouil Kioupakis. "Visible-wavelength polarized-light emission with small-diameter InN nanowires." In: *Nano letters* 14.7 (2014), pp. 3709–3714. DOI: 10.1021/nl404414r.

[21] Jack Deslippe et al. "BerkeleyGW: A massively parallel computer package for the calculation of the quasiparticle and optical properties of materials and nanostructures." In: *Computer Physics Communications* 183.6 (2012), pp. 1269–1289. URL: https://arxiv.org/pdf/1111.4429.

[22] Jl Wu et al. "Unusual properties of the fundamental band gap of InN." In: *Applied Physics Letters* 80.21 (2002), pp. 3967–3969. DOI: 10.1063/1.1482786.

[23] W. L. Chin V., L. Tansley T., and Osotchan T. "Electron mobilities in gallium, indium, and aluminum nitrides." In: *Journal of Applied Physics* 75.11 (1994), pp. 7365–7372. DOI: 10.1063/1.356650.

[24] Emmanouil Kioupakis et al. "Temperature and carrier-density dependence of Auger and radiative recombination in nitride optoelectronic devices." In: *New Journal of Physics* 15 (2013), p. 125006. DOI: 10.1088/1367-2630/15/12/125006.

[25] David Aurelien et al. "All-optical measurements of carrier dynamics in bulk-GaN LEDs: Beyond the ABC approximation." In: *Applied Physics Letters* 110.25 (2017), p. 253504. DOI: 10.1063/1.4986908.

[26] David Aurélien and J. Grundmann Michael. "Droop in InGaN light-emitting diodes: A differential carrier lifetime analysis." In: *Applied Physics Letters* 96.10 (2010), p. 103504. DOI: 10.1063/1.3330870.

[27] M Monavarian et al. "Explanation of low efficiency droop in semipolar (202?1?) InGaN/GaN LEDs through evaluation of carrier recombination coefficients." In: *Opt Express* 25.16 (2017), pp. 19343–19353. DOI: 10.1364/OE.25.019343.

[28] A Haug. "Auger recombination in direct-gap semiconductors: band-structure effects." In: *Journal of Physics C: Solid State Physics* 16.21 (1983), p. 4159. DOI: 10.1088/0022-3719/16/21/017.

[29] Emmanouil Kioupakis, Qimin Yan, and Chris G Van de Walle. "Interplay of polarization fields and Auger recombination in the efficiency droop of nitride light-emitting diodes." In: *Applied Physics Letters* 101.23 (2012), p. 231107. DOI: 10.1063/1.4769374.

[30] YC Shen et al. "Auger recombination in InGaN measured by photoluminescence." In: *Applied Physics Letters* 91.14 (2007), p. 141101. DOI: 10.1063/1.2785135.

[31] Francesco Bertazzi et al. "Auger recombination in InGaN/GaN quantum wells: A full-Brillouin-zone study." In: *Applied Physics Letters* 103.8 (2013), p. 081106. DOI: 10.1063/1.4819129.

[32] Emmanouil Kioupakis et al. "First-principles calculations of indirect Auger recombination in nitride semiconductors." In: *Physical Review B* 92.3 (2015). DOI: 10.1103/physrevb.92.035207.

[33] Christina M Jones et al. "Impact of carrier localization on recombination in InGaN quantum wells and the efficiency of nitride light-emitting diodes: Insights from theory and numerical simulations." In: *Applied Physics Letters* 111.11 (2017), p. 113501. DOI: 10.1063/1.5002104.

[34] Andrew McAllister, Dylan Bayerl, and Emmanouil Kioupakis. "Radiative and Auger recombination processes in indium nitride." In: *Applied Physics Letters* 112.25 (2018), p. 251108. DOI: 10.1063/1.5038106.

CHAPTER 6

Auger Recombination in AlGaN Alloys

We apply density functional theory and many-body perturbation theory to study Auger and radiative rates in the AlGaN alloys. Previous results have shown that in GaN (gap 3.4 eV), Auger primarily occurs through the assistance of phonons. While in pure InN (gap 0.7 eV), Auger occurs directly, with no assistance from other mechanisms. I present results of my first-principles Auger calculations on special quasirandom structures of AlGaN. Using these results I can suggest approaches to reduce the impact of Auger recombination on the efficiency of group-III nitride devices.

6.1 Alloys

Alloys are combination of two materials which creates some mixture of the properties of both materials. Metal alloys, like steel (carbon and iron) are seen frequently in everyday life. As mentioned in the introduction and in fig. 1.4, the group-III nitride materials have band gaps that span the UV to the infrared. Meaning we should be able to combine these materials to make LEDs that emit light of any visible color or in the UV and infrared.

AlGaN alloys with high AlN content emit in the deep-UV and have optical applications including sterilization, medical diagnostics/treatment, and sensing.[1] There are still major materials hurdles that need to be overcome before deep-UV AlGaN devices reach maturity. For optical applications in particular, an understanding of how to maximize the IQE is vital. Just like for InN in the previous chapter, AlGaN suffers from a low IQE at high current densities that may be from Auger recombination. In this case it is not *direct* Auger that we are interested in, since the band gap of these alloys is so large and the direct process reduces exponentially with the band gap. Instead, we are interested in understanding the assisted Auger process.

The phonon-assisted Auger process is one example, and was shown to be dominant in GaN.[2] In addition to phonons, there are two mechanisms specific to alloys that can enhance the Auger recombination process. The first is alloy scattering, where the folding of bands allows Auger tran-

sitions to occur because of the additional states carriers can be excited to. The second is carrier localization from alloy fluctuations. These nanometer scale, statistically significant variations in the alloy composition can lead to carriers becoming confined in areas of disparate alloy composition. This effect has have been shown to occur in AlGaN.[3] It is not well understood how this localization affects the recombination rates. Both radiative and Auger recombination may be enhanced by localized carriers, therefore the ratio of this enhancement will determine if the IQE is increased or decreased by this mechanism.[4]

6.2 Modeling Alloys With First-Principles Calculations

Modeling an alloy is more difficult because we must take into account more atoms to obtain the properties that we want. There is an inherent randomness in the ways alloys form – the crystal structure is a given but which atom ends up on which spot is a probability question. Since our calculations necessarily repeat the unit cell over and over again they cannot be truly random even if we pick elements to sit at sites in a random way they will be repeated by the code.

We can get around this problem by making bigger and bigger unit cells that hold more and more atoms. These are called *supercells*. The challenge with this is that computation is increased in two ways. First, because the supercell is made randomly, we have to take a *few* supercells and average them to get properties that we are interested in. Second, because these cells contain more atoms, the calculations are inherently more expensive.

A way that helps a bit with this problem is to use *special quasi-random supercells*: rather than just randomly deciding to have an atom at each spot, we can use the correlation function of the true alloy as a guide. The correlation function of an alloy is a way to see the neighbors of each atom. By using this as a guide, we can get a supercell that closely resembles the correlation function of the true alloy and base our calculations on this. We still have to consider more atoms than before, but at least we know our one supercell is a closer approximation.

In this work, we use the special quasi-random structure (SQS) from Shin et al. [5] for the 25%, 50% and 75% compositions. We relaxed these structures using plane wave DFT with norm conserving pseudopotentials. Our band structures are also calculated with DFT. Since we rigidly shift the band gap in later steps, the additional cost of doing GW calculations was unnecessary.

6.3 Auger Recombination in AlGaN Alloys

Applying the methodology from chapter 4, we calculated the alloy-assisted Auger recombination rates for the three AlGaN compositions modeled with the SQS. We used $18 \times 18 \times 18$ kmeshes for both the eeh and hhe processes. Gaussian functions with a width of 0.1 eV are used for evaluating

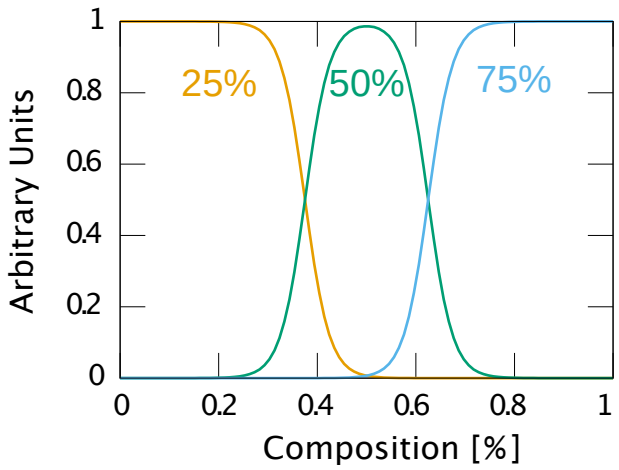

Figure 6.1 Our model for interpolating Auger results between the actual calculated compositions. The functions have maximums for their respective alloy and mix results to 100% between alloys.

energy delta functions in the Auger rate equations. The screened Coulomb matrix elements use a model dielectric function[6] with a high-frequency dielectric constant that is interpolated between GaN ϵ_∞ = 5.35[7] and AlN ϵ_∞ = 4.8.[8] We also calculate radiative recombination rates for the three alloys.

6.3.1 Interpolating Coefficients to Arbitrary Composition

From these three points for the 25%, 50% and 75% alloys we then use a simple model to obtain Auger rates across the entire composition range. First, for each alloy we rigidly shift the band gap of the material across the entire composition range, roughly simulating how Auger would change with alloying. Second, we correct the Auger coefficients to account for the band-bowing of AlGaN alloys.[9]

$$C_{Al_xGa_{1-x}N}(E_g) = C_{Al_{25}Ga_{75}N}(E_g)\frac{x(1-x)}{0.25.75}$$

Finally, we expect the Auger coefficient at the SQS compositions (25%, 50%, and 75%) to be fully determined by only their respective calculations. Between the outside compositions (25% and 75%), we combine the Auger coefficients with the 50% composition using Fermi functions that weight the contributions from each calculation to be higher near that composition, lower further away and split evenly at the middle compositions (37.5% and 62.5%). We also assume that at compositions lower (higher) than 25% (75%), there are only alloy contributions from the nearest composition calculation. This method of merging results from different compositions is shown in fig. 6.1.

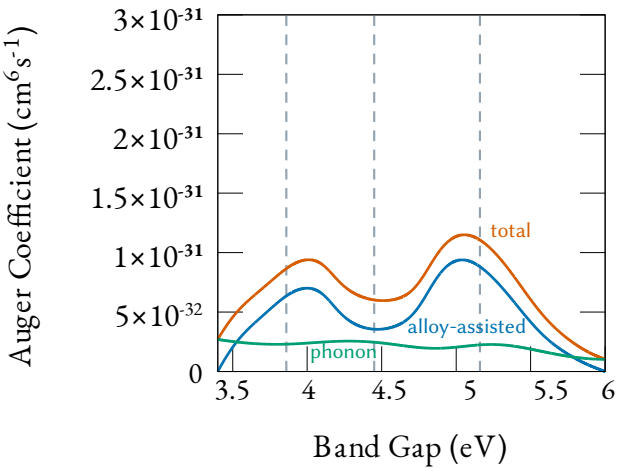

Figure 6.2 The *eeh* Auger recombination rates across AlGaN alloy compositions. The alloy-assisted process is the dominant process except at the extreme alloy compositions (which is expected, since there is minimal alloying for those compositions.). Note the y-axis scale crosses an order of magnitude.

6.3.2 Alloy-Assisted Auger

The results of our calculations are shown in figs. 6.2, 6.3, 6.4. Our results generally show that for the *eeh* process, alloy-assisted Auger is the dominant process but for the *hhe* process the phonon assisted Auger process is dominant throughout all compositions. Both processes are of the same order of magnitude for most of the composition range, although lower composition GaN alloys have a peak where the *eeh* alloy-assisted process becomes stronger than the *hhe* phonon process.

We would expect to find the largest alloy-assisted Auger rate at 50%, but instead our calculations show the alloy-assisted Auger rates are peaked at 25% and 75% and are at a minimum at 50%fig. 6.2. Because this is not intuitive, we are not sure if this represents the actual effect of alloy-assisted Auger recombination. It could be an artifact of the particular SQS that is suppressing Auger, an issue in the code, or the fact that the SQS is capturing a bit of the carrier localization. Work is ongoing to determine if this is the actual behavior or an error in the calculations.

The magnitude of the total rates is ~1×10^{-31}cm^{6}s^{-1}. This is smaller than other values reported in the literature. Hao et al. [10] found an Auger coefficient of 1.52×10^{-30}cm^{6}s^{-1}by fitting electroluminescene data with a fourth order polynomial. In 2018, Nippert et al. [11] found a similar coefficient studied AlGaN quantum well heterostructures using photoluminescence techniques and found an Auger recombination coefficient of 2.3×10^{-30}cm^{6}s^{-1}. Both of these experiments were done with AlGaN multi-quantum-well structures ($Al_{0.75}Ga_{0.25}N$ in Hao et al. [10] and $Al_{0.45}Ga_{0.55}N$ in Nippert et al. [11]). These structures are not analogous to our simulation, which is a bulk alloy. Differences could be caused effects like carrier localization that occurs

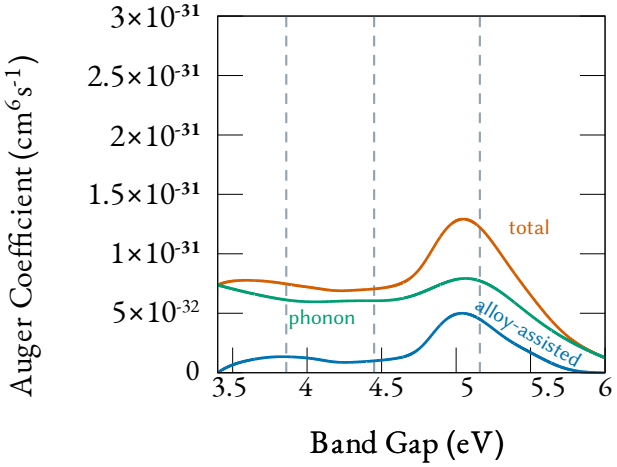

Figure 6.3 The *hhe* Auger recombination rates across AlGaN alloy compositions. The phonon process is the dominant process throughout all compositions. Note the y-axis scale crosses an order of magnitude.

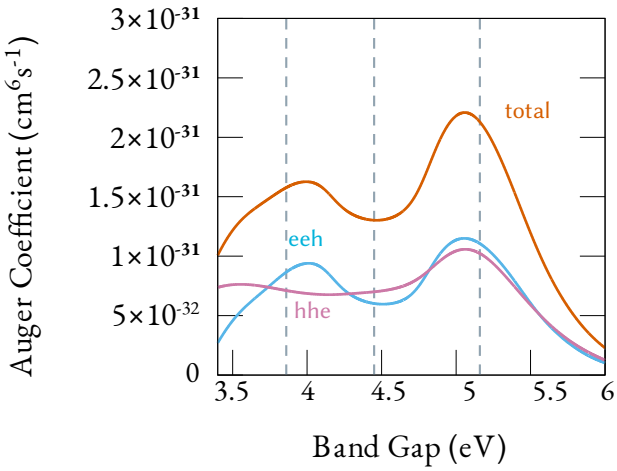

Figure 6.4 The combined (alloy and phonon-assisted) *eeh* and *hhe* Auger processes. Note the y-axis scale crosses an order of magnitude.

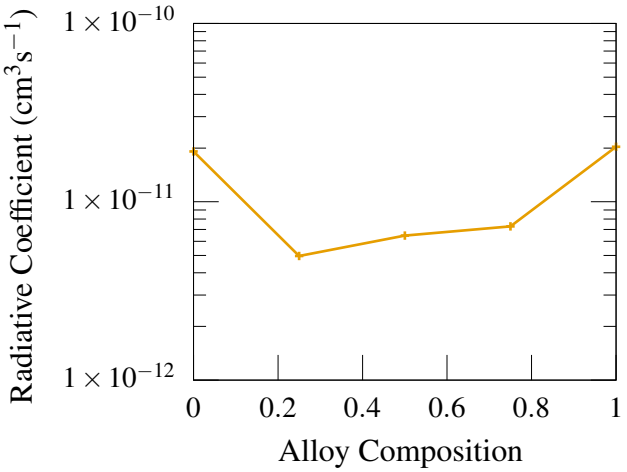

Figure 6.5 Calculated radiative rates for the three alloy compositions and the pure materials. While both GaN and AlN have similar values, the alloys are smaller by a factor of ~4. Lines are a guide for the eye.

specifically in the quantum well structures. Still, the lack of well determined Auger coefficients for this material and a recent interest in it (both papers are from the last two years) show that understanding Auger in AlGaN alloys is necessary.

We also can compare these results to experimental and theoretical results for InGaN alloys. Earlier calculations show values on the same order of magnitude in $In_{0.25}Ga_{.75}N$ alloy.[2] Early experimental work by[12] first called attention to Auger in InGaN with values on the order of 1×10^{-30} $cm^6s^{-1}$from photoluminescence measurements. Other experimental work has found Auger coefficients varying slightly around 1×10^{-31} cm^6s^{-1}[13] for bulk InGaN of various compositions. Because the Auger coefficients are similar, we expect that making highly efficient AlGaN devices will also be limited by Auger recombination.

This is further supported by radiative recombination coefficient calculations shown in fig. 6.5. These calculations show that the radiative rates are smaller in the alloy structures by a factor of ~4 compared to the bulk structures. This may be due to the decreased overlap of valence and conduction band wave functions in the alloy structure.

It has been shown by Jones et al. [4] that composition fluctuations increase both radiative and Auger recombination rates, but that Auger is increased by an order of magnitude more than the radiative rate. If composition fluctuations in AlGaN alloys cause a similar ratio of Auger to radiative recombination enhancement, the IQE will be significantly decreased. However as mentioned earlier, there is the possibility that some of the localization effects are being captured by the SQS used in our calculations. This would mean that we cannot explicitly combine calculations the Auger rates calculated here with calculations similar to Jones et al. [4] because we would be double counting the effects. Further investigation is needed to separate the two effects and their

individual impact on the recombination rates.

References

[1] JY Tsao et al. "Ultrawide-Bandgap Semiconductors: Research Opportunities and Challenges." In: *Adv. Electron. Mater.* 4.1 (2017), p. 1600501. DOI: 10.1002/aelm.201600501.

[2] Emmanouil Kioupakis et al. "First-principles calculations of indirect Auger recombination in nitride semiconductors." In: *Physical Review B* 92.3 (2015). DOI: 10.1103/physrevb.92.035207.

[3] David A Browne et al. "Vertical transport through AlGaN barriers in heterostructures grown by ammonia molecular beam epitaxy and metalorganic chemical vapor deposition." In: *Semiconductor Science and Technology* 32.2 (2017), p. 025010. DOI: 10.1088/1361-6641/32/2/025010.

[4] Christina M. Jones et al. "Impact of carrier localization on recombination in InGaN quantum wells and the efficiency of nitride light-emitting diodes: Insights from theory and numerical simulations." In: *Applied Physics Letters* 111.11 (2017), p. 113501. DOI: 10.1063/1.5002104.

[5] Dongwon Shin et al. "Thermodynamic properties of binary hcp solution phases from special quasirandom structures." In: *Physical Review B* 74.2 (2006). DOI: 10.1103/physrevb.74.024204.

[6] G. Cappellini et al. "Model dielectric function for semiconductors." In: *Physical Review B* 47.15 (1993), pp. 9892–9895. DOI: 10.1103/physrevb.47.9892.

[7] Michael E. Levinshtein, Sergey L. Rumyantsev, and Michael S. Shur. *Properties of Advanced Semiconductor Materials*. John Wiley & Sons, 2001. URL: http://books.google.com/books?id=u26CpULkD_wC%5C&hl=%5C&source=gbs_api.

[8] A. T. Collins, E. C. Lightowlers, and P. J. Dean. "Lattice Vibration Spectra of Aluminum Nitride." In: *Physical Review* 158.3 (1967), pp. 833–838. DOI: 10.1103/physrev.158.833.

[9] Stephen Pearton. *GaN and ZnO-based materials and devices*. Vol. 156. Springer Science & Business Media, 2012. URL: https://books.google.com/books?hl=en%5C&lr=%5C&id=VwlDtqc1c7oC%5C&oi=fnd%5C&pg=PR3%5C&dq=GaN+and+ZnO-based+Materials+and+Devices%5C&ots=s21TYwICEC%5C&sig=LjX8BamZY2x2rWxZG3nUkkfAD1w.

[10] GD Hao et al. "Electrical determination of current injection and internal quantum efficiencies in AlGaN-based deep-ultraviolet light-emitting diodes." In: *Opt Express* 25.16 (2017), A639–A648. DOI: 10.1364/OE.25.00A639.

[11] Felix Nippert et al. “Auger recombination in AlGaN quantum wells for UV light-emitting diodes.” In: *Applied Physics Letters* 113.7 (2018), p. 071107. DOI: 10.1063/1.5044383.

[12] YC Shen et al. “Auger recombination in InGaN measured by photoluminescence.” In: *Applied Physics Letters* 91.14 (2007), p. 141101. DOI: 10.1063/1.2785135.

[13] Joachim Piprek, Friedhard Römer, and Bernd Witzigmann. “On the uncertainty of the Auger recombination coefficient extracted from InGaN/GaN light-emitting diode efficiency droop measurements.” In: *Applied Physics Letters* 106.10 (2015), p. 101101. DOI: 10.1063/1.4914833.

CHAPTER 7

Auger Recombination Contributions to Scintillator Non-Proportionality

Scintillator radiation detectors suffer from low energy resolution that has been attributed to non- linear light yield response to the energy of the incident gamma rays. Auger recombination is a key non-radiative recombination channel that scales with the third power of the excitation density and may play a role in the non-proportionality problem of scintillators. In this work, we study direct and phonon-assisted Auger recombination in NaI using first-principles calculations. Our results show that phonon-assisted Auger recombination, mediated primarily by short-range phonon scattering, dominates at room temperature. We discuss our findings in light of the much larger values obtained by numerical fits to z-scan experiments.

7.1 "Seeing Through" Shipping Containers

Over 20 million shipping containers go through United States ports each year by sea, truck and rail.[1,2] Physically opening and inspecting each of these containers is impossible, but we would like to at least make sure that there is nothing inside of these containers that would be harmful to the country or its citizens. One specific concern for the Department of Homeland Security is radioactive materials. These materials could be then used in an improvised nuclear device or in a "dirty" radioactive bomb that could cause thousands of deaths, immense damage to buildings and infrastructure, expensive decontamination efforts and the need to relocate thousands of Americans.

This possibility lead the Department of Homeland Security to installing radiation portal monitors (RPMs) at ports of entry to scan for radioactive materials (both within the United States and at major allied ports around the world). These detectors are large enough that shipping containers or trucks can be moved or driven through easily. RPM's use a combination of materials to detect different types of ionizing and indirectly ionizing radiation. Helium-3 is used

to detect neutrons,[3] while polyvinyl toulene (PVT) is used to detect gamma radiation.[4] But because of the (necessary) sensitivity, the detectors suffer from many "nuisance" alarms caused by benign radioactive materials[1]. These range from radioisotopes that are used in industry[5] and medicine[6] to commercial products like granite and kitty litter.[7] Nuisance alarms create the need for secondary inspections through more sophisticated (and expensive) scanners or by smaller hand scanners which require more time (and people) to operate. Although these nuisance alarms are rare (about 2%,[8]), this still amounts to hundreds of thousands of containers.

To reduce the nuisance alarm rate, the Department of Homeland security funded research into alternatives that could identify not just that there was radioactive material inside of a shipping container, but also identify the elemental source of radioactivity.[3,7] Scintillating materials are one candidate for both neutron detection and gamma ray detection. There were early efforts in 2007 to implement these materials in new RPMs called Advanced Spectroscopic Portal Monitors (ASPs). But after extensive testing this program was discontinued due to excessive costs and no substantial decrease in nuisance alarms compared to the first-generation of RPMs.

Although the current "fleet" of RPM's based on Helium-3 and PVT is still operational and even expected to last until 2030,[8] there are still major efforts to find new materials for radiation detection. These could be applicable to more accurate RPMs at our borders, but also could have applications in medicine and basic science.

7.2 Scintillators and Non-Proportionality

As mentioned above, one promising candidate for detecting neutrons and gamma radiation are scintillating materials. Scintillators convert high energy radiation (or neutrons) into excited carriers: electrons, holes and excitons. These excited carriers then recombine to emit lower-energy light, which is easier to detect and analyze than the initial high energy radiation (see fig. 7.1). The ultimate goal is to identify the energy of the original high energy particle from the spectrum of lower energy light emitted from the scintillating material.[9] This would allow us to know what element that the original particle came from, and whether it is dangerous or not.

For the whole scintillator detector to work, the carriers must recombine radiatively. However, this is not the *only* way that carriers can recombine. Free carriers can can lose energy non-radiatively through impurities and Auger recombination. These non-radiative recombination mechanisms essentially trap energy in the scintillating material rather than allow it's detection through the photomultiplier[2]. This limits the detector's resolution and causes an elemental

[1]Nuisance alarms differ from false alarms that are caused by outside natural sources like cosmic background radiation. False alarms do not pose as much of a problem for RPMs.

[2]I'm glossing over how complex this is. There are many other difficulties in modeling this including differing carrier masses contributing to different carrier distributions, dE/dx...

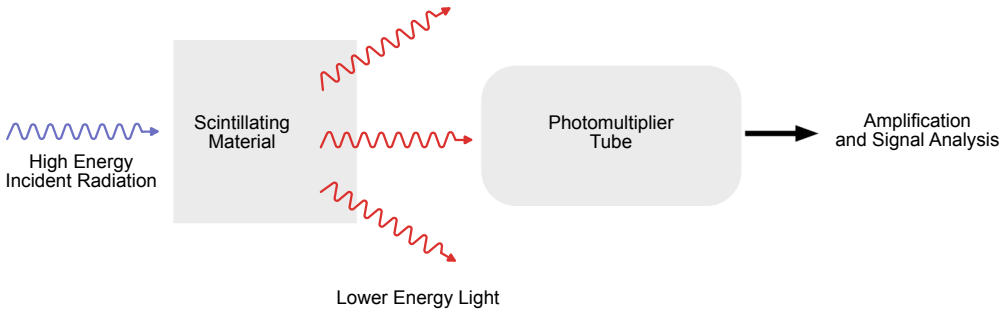

Figure 7.1 Diagram illustrating how scintillating crystals can be used to detect radiation. An incident neutron or photon with high energy excites carriers in the scintillating material. These carriers will relax to the ground state, emitting photons at lower energy than the incident radiation. Those photons are picked up by a photomultiplier tube, detected and analyzed to determine the original energy of the incident particle.

fingerprint to not be as clear.

This contributes to a problem known as scintillator *non-proportionality*.

One method used to better understand how the ionized carriers behave is through kinetic monte carlo simulations. These simulations model the "excitation tracks" created by incident radiation and then evolve them with time considering separately densities of free carriers and excitons. These kinds of simulations can help understand how much energy is typically lost inside of the scintillating material and how to translate that into a reliable detector. But to do this, they need to have the rates of various recombination processes including Auger recombination. As mentioned in the chapter 3, this can prove difficult to obtain from both theory and experiment and is where first-principles calculations can be used.

7.3 Why Sodium Iodide?

Sodium Iodide is a simple, cubic crystal that is easy to grow in large quantities.[10] It can be used in scintillators in single and polycrystalline form[11] and is already widely used. It was one of the first materials to be shown to have scintillating properties when and lead to the first PET scanners used in medicine.[12] Unfortunately, sodium iodide is also one of the most *non-proportional* scintillating materials.[13] While many properties are well-studied in NaI, Auger recombination has prove difficult to understand because of the theoretical and experimental hurdles described in chapter 3.

Experimental measurements of Auger recombination in sodium iodide have widely varying results. Using rate equations for the temporal and spatial evolution of excitation densities, the work in Bizarri et al. [14] found an AR coefficient of 1.07×10^{-20} $cm^6s^{-1}$when modeling Compton coincidence data from SLYNCI (see Choong et al. [15]). But a much smaller coefficient of 3.2×10^{-29} $cm^6s^{-1}$is found when modeling z-scan experiments using a pulsed laser.[16,17]

Because of the need for accurate Auger recombination numbers for KMC modeling and because of the large variance among experimental measurements, sodium iodide is an ideal candidate for study using our first-principles methods.

7.4 Auger Recombination in Sodium Iodide from First Principles

We used the methodology discussed in chapter 4 to study the Auger processes in NaI. This includes using plane wave DFT via Quantum Espresso,[18,19] wannier interpolation using Wannier90[20,21] and our own Auger recombination code. The pseudopotentials are norm conserving[22] and gen-

erated using the Fritz-Haber code[23] in the GGA approximation.[24]

For the Auger rates, Gaussian functions with a width of 0.1 eV were used to evaluate the energy delta functions. We used a model dielectric function[25] to calculate the screened Coulomb interaction matrix elements, including the screening by free carriers using the Debye-Hückel equation for non-degenerate carriers and the Thomas-Fermi model for degenerate carrier concentrations.[26] The lattice temperature is taken to be 300 K while the electron temperature is set to 500 K. We assume a carrier density of 1×10^{19} cm^{-3}. The structure of the material is relaxed and our lattice constant is 12.58 Bohr, within 3% of the experimental lattice constant. Finally, we used the experimental high frequency dielectric constant $\epsilon_\infty = 2.98$.[27]

Our calculated DFT band-structure (rigidly-shifted to 5.8 eV, the experimental band gap[28]) is in +fig. 7.2. Spin-orbit interaction effects were not included for this material. Using this band structure, we can make some predictions about Auger in the material. Auger excites carriers to states that are at energies approximately equal to the band gap from the conduction and valence-band edges. As is evident from the band structure, there are numerous conduction bands at the energy ($2E_g$ = 11.6 eV) needed to accommodate hot electrons excited by the *eeh* Auger process. Thus we conclude that direct and phonon-assisted *eeh* AR are possible in NaI. However, there are no valence-band states around −5.8 eV to accommodate holes excited by the *hhe* Auger process and therefore *hhe* AR cannot occur in NaI. In the following, we focus our attention to *eeh* Auger processes only. Finally, we note that phonon-assisted Auger will likely be dominant because excited electrons will need additional momentum to be excited into higher energy states #eqn:augerscaling.

7.4.1 Direct Auger

The *eeh* direct Auger vales calculated for NaI are shown in +fig. 7.3. Recall that we vary the gap to estimate how sensitive the coefficients are to the band gap value (here from 4.5 - 6.5 eV). The calculated value at the experimental band gap is $(1.17 \pm 0.01) \times 10^{-33}$ cm^{6}s^{-1}. *fig. 7.3 shows that a BZ sampling grid of $40 \times 40 \times 40$ is sufficient to converge the Auger coefficients. The same figure also shows that the direct Auger coefficient does not depend sensitively on the band gap. For example, if the band gap changes within ±0.3 eV from the experimental value (5.8 eV) the direct Auger coefficient changes by at most 50%.

The dependence of the direct Auger coefficient on the band gap follows the Arrhenius activation law derived for *intraband* Auger processes ($C \propto \exp(-E_A/k_BT)$)[29,30] even though direct AR in NaI occurs through *interband* processes to higher conduction bands. Here, E_A is the activation energy which is proportional to the band gap. We show below that E_A = 0.18 eV, which is a small fraction (3.1 %) of the experimental gap.

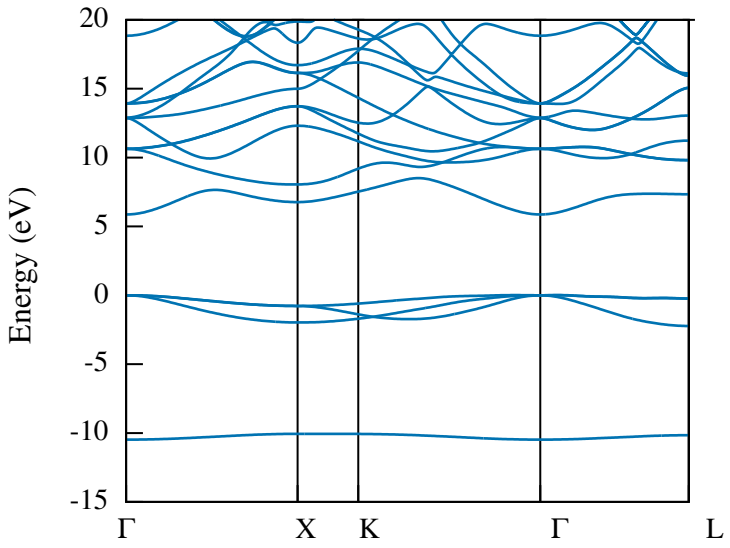

Figure 7.2 The calculated band structure of NaI. Energies are referenced to the valence band maximum at Γ. The gap has been rigidly adjusted to the experimental value (5.8 eV) to account for the band gap problem of density functional theory. The absence of valence bands around − 5.8 eV indicates that the *hhe* Auger process is not possible in this material.

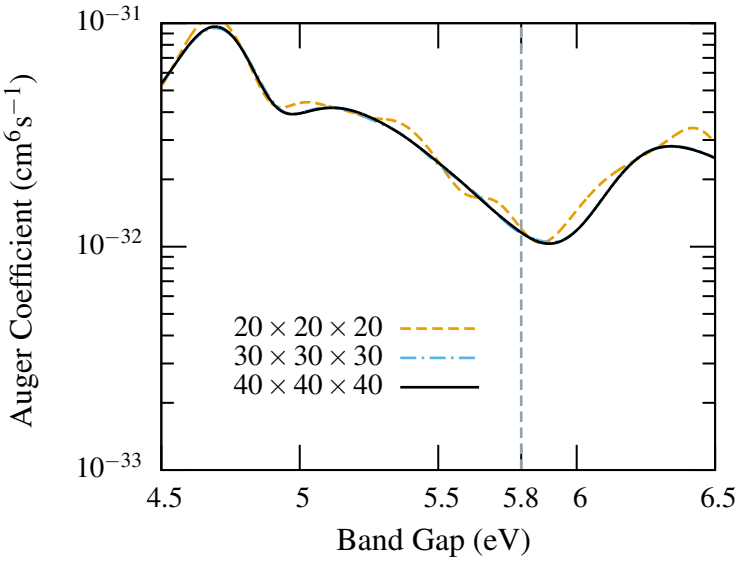

Figure 7.3 Calculated values for the direct *eeh* Auger recombination coefficient of NaI as a function of the grid spacing used to sample the BZ and the (rigidly adjusted) band gap of the material. The converged value for the experimental band gap is $(1.17 \pm 0.01) \times 10^{-33}$ cm^6s^{-1}.

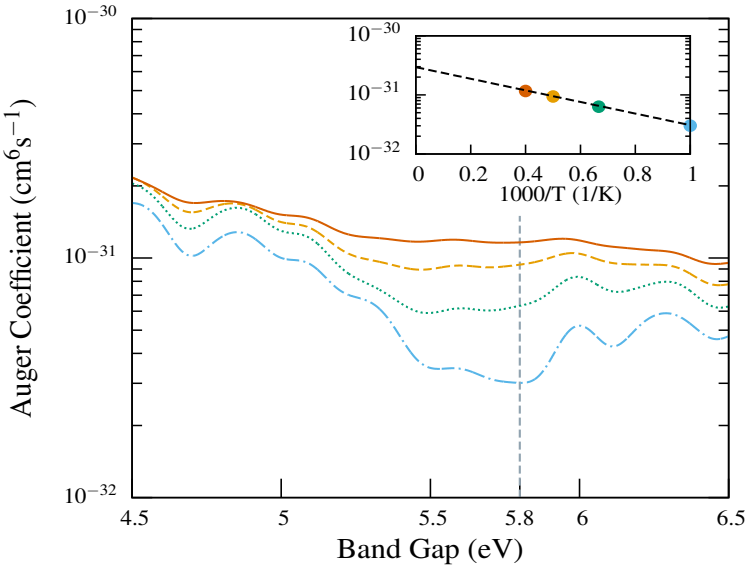

Figure 7.4 The direct Auger recombination coefficient for varying electronic temperatures. The direct AR coefficient follows an Arrhenius activation-law model (inset), which predicts a maximum direct AR coefficient of 2.95×10^{-32}cm^{6}s^{-1} that is less than the phonon-assisted coefficient.

7.4.2 Phonon-Assisted Auger

The phonon-assisted AR coefficients are shown in +fig. 7.5 as a function of the band gap. The value of the phonon-assisted Auger coefficient at the experimental band gap, $(5.6 \pm 0.3) \times 10^{-32}$cm^{6}s^{-1}is approximately two orders of magnitude larger than the direct one and therefore phonon-assisted processes dominate AR in NaI. This value is also comparable to phonon-assisted *eeh* AR values in semiconductors such as GaN[31] (4×10^{-32} cm^{6}s^{-1}) and GaAs[32] (1.1×10^{-31} cm^{6}s^{-1}).

We can investigate which kinds of phonons are important to assisting the Auger process. *fig. 7.6 shows the contribution of the various phonon modes to the phonon-assisted Auger coefficient. AR in NaI is mediated primarily by the acoustic and the longitudinal optical (LO) phonons, while the contribution by the transverse optical (TO) modes is approximately one order of magnitude smaller. *fig. 7.7 shows the phonon-assisted Auger results analyzed in terms of the contribution by the various phonon wave vectors. The results indicate that the dominant contributions occur at wave vectors comparable to the BZ dimensions that correspond to phonon wavelengths comparable to the lattice constant. Therefore, AR in NaI is primarily assisted by short-range phonon scattering.

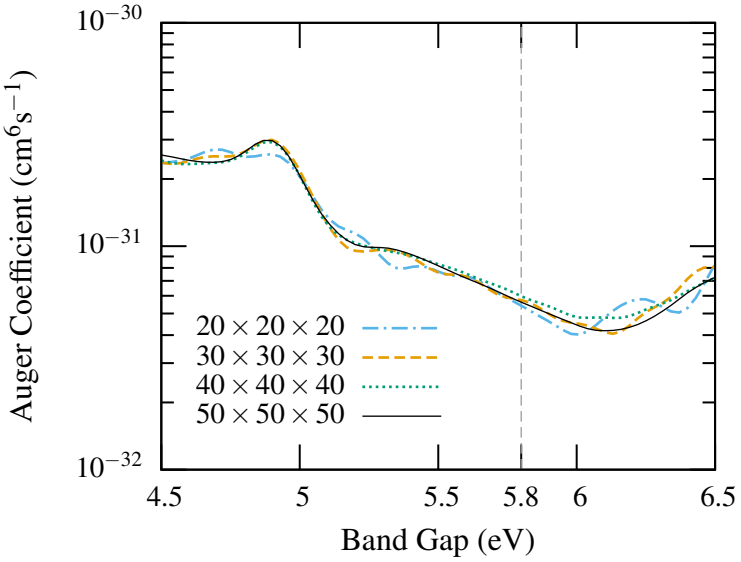

Figure 7.5 Calculated phonon-assisted *eeh* Auger recombination coefficients of NaI as a function of the adjusted band gap and the BZ sampling. The phonon-assisted Auger coefficient at the experimental band gap value is $(5.6 \pm 0.3) \times 10^{-32}$ cm^6s^{-1}, which is approximately two orders of magnitude larger than the direct Auger coefficient (see +fig. 7.3).

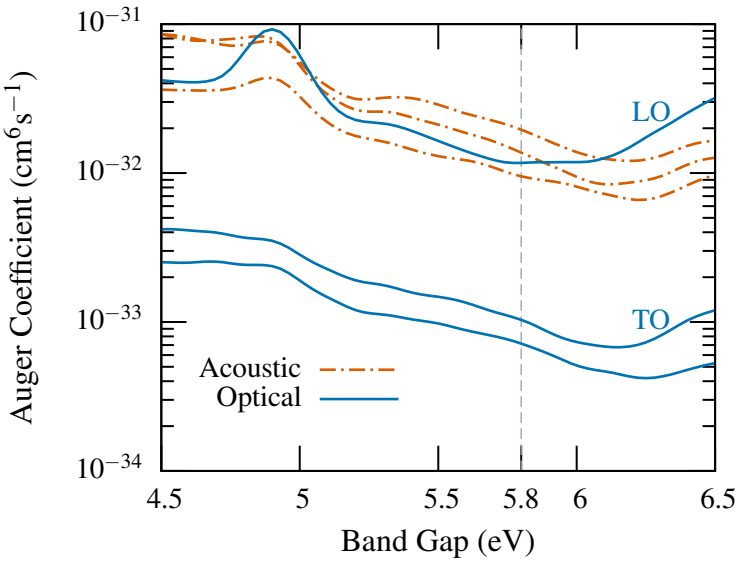

Figure 7.6 The contribution of the various phonon modes to the phonon-assisted Auger coefficient of NaI. The phonon-assisted processes are dominated by the acoustic and the longitudinal optical (LO) phonon modes, while the contribution by the transverse optical (TO) phonons is approximately one order of magnitude smaller.

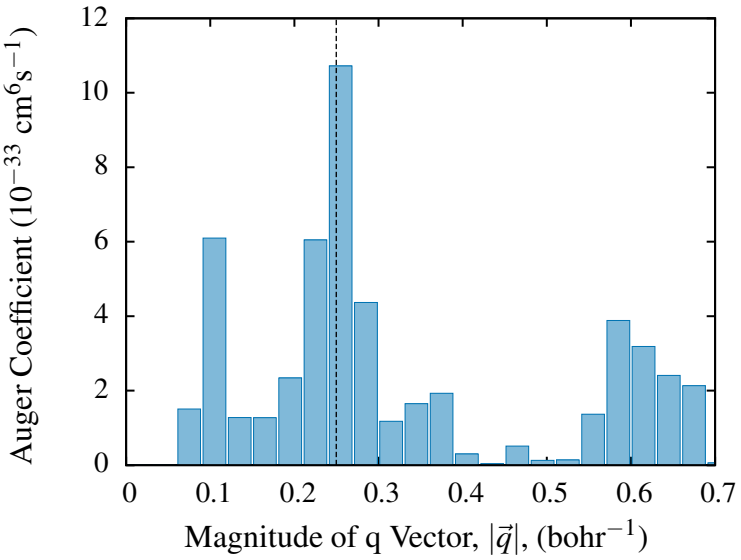

Figure 7.7 The distribution of the contribution by the various phonon wave vectors to the phonon-assisted Auger coefficient of NaI. The vertical line represents the edge of the BZ.

7.5 Conclusions

Comparing with the coefficients fitted to experimental data, we note that our calculated rate is several orders of magnitudes smaller than the value of 1.07×10^{-20} cm^{6}s^{-1}obtained by Bizarri *et al.*[14] We note however, that recently more improved sets of rate equations and boundary conditions have been developed.[16,17,33] Comparing to more recent results from fits ($C = 3.2 \times 10^{-29}$ cm^{6}s^{-1}) to z-scan experiments, our calculated values are off by a few decades.[16,17] This illustrates the difficulty in fitting a rate equation model to a very complex set of events and interdependent mechanisms. While one possible cause for the difference between our results and experimental data is thallium doping present in NaI samples that were studied in experiment, we note that Williams et al. [16] and Grim et al. [17] have considered both pristine and Tl-doped CsI with only a small variation (a factor of about 1.5) of the AR coefficients.We therefore expect that the inclusion of dopants would not have a large effect on the AR rate.

Interestingly, for a wide range of materials, Williams *et al.*, noted an empirical band-gap rule.[16,17] This trend was however not followed by the halide systems under consideration, which instead displayed values that exceeded the expected values by three to four orders of magnitude. The plausible explanation by the authors was the following: *(i)* the underlying model and assumptions are wrong or insufficient, or *(ii)* the excitation energy exceeds the band gap for at least NaI by about 0.3 eV (3500 K electronic temperature) and might leave the carriers with very high electronic temperatures, or *(iii)* the Auger recombination might involve self-trapped holes (V_k-centers) with localized hole states in the forbidden gap. We note that our predicted value actually

obeys the observed empirical band gap rule. To assess the effect of high electronic temperature on AR coefficients, we performed direct AR calculations for increasing electronic temperatures up to 2500 K +fig. 7.4.

The high-temperature calculations were performed with a $14 \times 14 \times 14$ k point grid since the convergence with respect to mesh size is faster for higher temperatures. We further fitted the data for the experimental band gap of 5.8 eV with an Arrhenius activation law $C(T) = C_0 \exp(-E_A/k_B T)$ (inset of +fig. 7.4). The fit parameters are $C_0 = 2.95 \times 10^{-32}$ $\mathrm{cm}^6\mathrm{s}^{-1}$and $E_A = 0.18$ eV $= 0.031 E_\mathrm{g}$. Using this model, we obtained changes in the direct AR coefficient of at most 50% if the band gap is changed by ±0.3 eV as was found in the calculated data. We can extrapolate this model to infinite temperature and find that the maximum direct AR coefficient is C_0, which is still a factor of 2 lower than the phonon-assisted AR coefficient. Therefore phonon-assisted AR dominates over direct AR even at high electronic temperatures.

The third possibility of a self-trapped hole to participate in the Auger process is indeed very interesting, but it is beyond the scope of the present work. We note however that the localized nature of the hole level relaxes the momentum conservation selection rules and may lead to larger rates. Furthermore, it is also possible that the localized empty levels provide more routes for direct AR.

In summary, we performed first-principles calculations to find the direct and phonon-assisted AR coefficients for the basic scintillating material, sodium iodide. We found that the phonon-assisted process is dominant in NaI, and the magnitude of the AR coefficient is smaller than values derived in previous work. Understanding AR and resolving the discrepancy with previous modeling work is necessary to successfully model the scintillating process in sodium iodide. Being able to better predict how this and other scintillating materials react during the high-energy impact would allow us to guide future materials design aimed at dramatically improving the device performance and developing detectors that easily allow radiation sources to be identified without error.

References

[1] U.S. Customs and Border Protection. *Performance and Accountability Report: Fiscal Year 2016*. 2017. URL: https://www.cbp.gov/sites/default/files/assets/documents/2017-Mar/FY-2016-CBP-PAR-508C.pdf.

[2] U.S. Customs and Border Protection. "Snapshot: A Summary of CBP Facts and Figures." In: (2017). URL: https://www.cbp.gov/sites/default/files/assets/documents/2018-Aug/cbp-snapshot-20180823.pdf.

[3] Government Accountability Office. "Neutron detectors: Alternatives to using helium-3." In: 11-753 (2011). URL: https://www.gao.gov/products/GAO-11-753.

[4] Government Accountability Office. "Lessons Learned from DHS Testing of Advanced Radiation Detection Portal Monitors." In: 09-804T (2009).

[5] World Nuclear Association. "Radioisotopes in Industry." In: (2017). URL: http://world-nuclear.org/information-library/non-power-nuclear-applications/radioisotopes-research/radioisotopes-in-industry.aspx.

[6] United States Nuclear Regulatory Comission. "Backgrounder on Medical Use of Radioactive Materials." In: (2011). URL: https://www.nrc.gov/reading-rm/doc-collections/fact-sheets/med-use-radioisotopes-bg.html.

[7] Government Accountability Office. "Lessons Learned from Cancelled Radiation Portal Monitor Program Could Help Future Acquisitions." In: 13-256 (2013).

[8] Government Accountability Office. "DHS's Fleet Is Lasting Longer than Expected, and Future Acquisitions Focus on Operational Efficiencies." In: GAO-17-57 (2016). URL: https://www.gao.gov/products/GAO-17-57.

[9] Nicholas Tsoulfanidis. *Measurement and detection of radiation*. CRC press, 2010. URL: https://books.google.com/books?hl=en%5C&lr=%5C&id=zRjSBQAAQBAJ%5C&oi=fnd%5C&pg=PP1%5C&dq=Measurement+and+detection+of+radiation.%5C&ots=VbiMixS8ki%5C&sig=ruUMmXZ7guC_SKga6idSj0iV6lM.

[10] LG Eidelman et al. "Automated pulling from the melt—an effective method for growing large alkali halide single crystals for optical and scintillation applications." In: *Journal of crystal growth* 128.1-4 (1993), pp. 1059–1061. URL: https://www.sciencedirect.com/science/article/pii/S0022024807800972.

[11] V. Taranyuk et al. "NaI(Tl) and CsI(Tl) scintillation crystal growth by skull method." In: *Journal of Crystal Growth* 318.1 (2011), pp. 820–822. DOI: 10.1016/j.jcrysgro.2010.11.089.

[12] DW Townsend. "Physical principles and technology of clinical PET imaging." In: *Annals-Academy of Medicine Singapore* 33.2 (2004), pp. 133–145. URL: http://www.annals.edu.sg/pdf200403/V33N2p133.pdf.

[13] Stephen A Payne et al. "Nonproportionality of scintillator detectors: Theory and Experiment. II." In: *IEEE Transactions on Nuclear Science* 58.6 (2011), pp. 3392–3402. URL: https://ieeexplore.ieee.org/abstract/document/6043891.

[14] G. Bizarri et al. "Progress in Studying Scintillator Proportionality: Phenomenological Model." In: *IEEE Transactions on Nuclear Science* 56.4 (2009), pp. 2313–2320. DOI: 10.1109/tns.2009.2022625.

[15] Woon-Seng Choong et al. "Performance of a Facility for Measuring Scintillator Non-Proportionality." In: *IEEE Transactions on Nuclear Science* 55.3 (2008), pp. 1073–1078. DOI: 10.1109/tns.2008.922824.

[16] Michael Fiederle et al., eds. *Experimental and computational results on exciton/free-carrier ratio, hot/thermalized carrier diffusion, and linear/nonlinear rate constants affecting scintillator proportionality*. SPIE, 2013. DOI: 10.1117/12.2027716.

[17] Joel Q. Grim et al. "Nonlinear quenching of densely excited states in wide-gap solids." In: *Physical Review B* 87.12 (2013). DOI: 10.1103/physrevb.87.125117.

[18] Paolo Giannozzi et al. "QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials." In: *Journal of Physics: Condensed Matter* 21.39 (2009), p. 395502. DOI: 10.1088/0953-8984/21/39/395502.

[19] P Giannozzi et al. "Advanced capabilities for materials modelling with Quantum ESPRESSO." In: *J Phys Condens Matter* 29.46 (2017), p. 465901. DOI: 10.1088/1361-648X/aa8f79.

[20] Arash A. Mostofi et al. "wannier90: A tool for obtaining maximally-localised Wannier functions." In: *Computer Physics Communications* 178.9 (2008), pp. 685–699. DOI: 10.1016/j.cpc.2007.11.016.

[21] Arash A. Mostofi et al. "An updated version of wannier90: A tool for obtaining maximally-localised Wannier functions." In: *Computer Physics Communications* 185.8 (2014), pp. 2309–2310. DOI: 10.1016/j.cpc.2014.05.003.

[22] J Ihm, A Zunger, and M L Cohen. "Momentum-space formalism for the total energy of solids." In: *Journal of Physics C: Solid State Physics* 12.21 (1979), pp. 4409–4422. DOI: 10.1088/0022-3719/12/21/009.

[23] N. Troullier and José Luriaas Martins. "Efficient pseudopotentials for plane-wave calculations." In: *Physical Review B* 43.3 (1991), pp. 1993–2006. DOI: 10.1103/physrevb.43.1993.

[24] John P. Perdew, Kieron Burke, and Matthias Ernzerhof. "Generalized Gradient Approximation Made Simple." In: *Physical Review Letters* 77.18 (1996), pp. 3865–3868. DOI: 10.1103/physrevlett.77.3865.

[25] G. Cappellini et al. "Erratum: Model dielectric function for semiconductors." In: *Physical Review B* 48.15 (1993), pp. 11520–11520. DOI: 10.1103/physrevb.48.11520.2.

[26] BK Ridley. *Quantum Processes in Semiconductors*. Oxford: Clarendon, 1982.

[27] M. E. Lines. "Bond-orbital theory of linear and nonlinear electronic response in ionic crystals. II. Nonlinear response." In: *Physical Review B* 41.6 (1990), pp. 3383–3390. DOI: 10.1103/physrevb.41.3383.

[28] Piotr A Rodnyi. *Physical processes in inorganic scintillators*. Vol. 14. CRC press, 1997. URL: https://books.google.com/books?hl=en%5C&lr=%5C&id=BxgNoWH6aPAC%5C&oi=fnd%5C&pg=PA1%5C&dq=Physical+Processes+in+Inorganic+Scintillators%5C&ots=a3Vxf30LCY%5C&sig=5oTAJtJM97xhAm2H2F-g4F_pIuM.

[29] H. Fröhlich and J. O'Dwyer. "Time Dependence of Electronic Processes in Dielectrics." In: *Proceedings of the Physical Society. Section A* 63.2 (1950), p. 81. DOI: 10.1088/0370-1298/63/2/301.

[30] A. R. Beattie and P. T. Landsberg. "Auger effect in semiconductors." In: *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 249.1256 (1959), pp. 16–29. URL: http://rspa.royalsocietypublishing.org/content/249/1256/16.short%202014-02-20.

[31] Emmanouil Kioupakis et al. "Indirect Auger recombination as a cause of efficiency droop in nitride light-emitting diodes." In: *Applied Physics Letters* 98.16 (2011), p. 161107. URL: https://aip.scitation.org/doi/full/10.1063/1.3570656.

[32] Daniel Steiauf, Emmanouil Kioupakis, and Chris G. Van de Walle. "Auger Recombination in GaAs from First Principles." In: *ACS Photonics* 1.8 (2014), pp. 643–646. DOI: 10.1021/ph500119q.

[33] Arnold Burger et al., eds. *Toward a user's toolkit for modeling scintillator non-proportionality and light yield*. SPIE, 2014. DOI: 10.1117/12.2063468.

CHAPTER 8

Outlook and Future Work

8.1 The Auger Code

There are three categories of improvements that could be made with the Auger code: (i) parallelization improvements to allow the study of larger systems, (ii) expanding the methodology to investigate other recombination mechanisms, and (iii) use/quality of life improvements that continue to make the code more usable and easier to expand.

Currently the code is parallelized using message-passing interface (MPI) and it is decided at the beginning how many tasks each processor will be given. However, there can be some variability with the actual *work* or time spent computing that is given to each processor. One way to manage this would be to use a worker-manager method where processors are given chunks of tasks and when finished request more. This could lead to more balanced load for each processor and could speed up the program. Additional speedups could be obtained by using shared memory (openMP) parallelization. This may be particularly useful when the wave functions become very large and reading them over and over again by separate processors causes a bottleneck. In addition, taking advantage of modern many-core architectures (through graphics processing units (GPUs) or many-core processors) could also allow the code to run faster.

There are other improvements to the methodology that could be made as well. One of the major limitations of the code is that we only consider phonon transitions from Γ. While this intuitively makes sense, because most carriers are near the band extrema, it would be helpful to consider more transitions to quantify the error of only including transitions from Γ. This would also give us the ability to consider indirect band gap materials like Silicon, where phonon-assisted Auger has not been studied using first-principles methods.

The current screening model is based on experimental data. We could utilize screening calculated through the Bethe-Salpeter equation. Our InN results show that screening is the dominant cause of Auger declining at very high carrier densities and including first-principles screening could confirm and continue to shed light on how screening affects Auger.

There are also two similar processes to Auger recombination that would be useful to calculate and could use existing structures in the code. Excitonic Auger, Auger between two electron-hole pairs, is not considered in this method at all and could be a substantial contributer to the overall Auger rate. Using exciton wave functions from the GW method would be an important contribution. In addition, the inverse Auger process (impact ionization) where an excited carrier creates an electron-hole pair, could be added to the code as well.

Finally, there are certain improvements that could be made to using the code. While this list could be very extensive there are three important considerations that come to mind. The first would be to add checkpointing to the code, which would allow for jobs to restart if stopped before completion. The second would be to allow users to specify units in the input files, which would allow for less confusion about what units parameters are in. The final is to update the code to use the most recent versions of QuantumESPRESSO and Wannier90. Currently the code requires modifications to these programs to have them output intermediate files that are not normally output. However, more recent versions of both programs have implemented ways to output these files. The code should be updated to use these so that older, modified versions of these codes don't have to be used.

8.2 The Group-III Nitrides

Work on the AlGaN Auger recombination is ongoing because of the unintuitive results. Currently we are investigating if the SQS itself has an impact on the calculation of the Auger rates. This could explain the unexpected drop in Auger rate. We also are interested in whether or not the SQS is large enough to include some of the localization effects from alloying. This would also affect our Auger rates.

Other work on the group-III nitrides includes understanding how Auger is affected by confinement in quantum well or short-period super lattice structures. In these structures, carrier confinement and polarization fields affect the recombination rates, but there is debate over whether this will increase or decrease the total IQE. First-principles calculations could help resolve this debate.

8.3 Auger in Scintillator Materials

There are many potential directions that this project can go in. The first is to continue studying Auger in different scintillator materials to determine if there is any trend with Auger and the different scintillation crystal types. Cesium Iodide is very similar to NaI, and we would expect it's Auger recombination coefficient to be similar. But Strontium Iodide is a strongly non-

proportional scintillating material, and it would be helpful to know if its Auger recombination coefficient is much smaller than those of non-proportional materials. This would demonstrate whether Auger is an important piece of the non-proportionality puzzle. These could lead to more high-throughput type calculations where many materials are analyzed with the hope of uncovering a trend in how Auger is related to the scintillator non-proportionality.

Another direction would be to refine the approximations made in the calculations of Auger in NaI. We did non consider the normal doping of NaI with Thallium, which would affect the Auger recombination rate. In addition, we did non consider excitons which are also generated when radiation excites carriers in the material. Finally, we did not consider polarons, which are closely related to excitons and could enhance or reduce the Auger recombination rate in the material.

The challenges of these calculations is that they will necessarily require more supercomputer time to account for the larger supercells that we must use to take into account the more complex structure. But these are essential pieces to understanding how Auger affects scintillator non-proportionality.