ELECTRO-THERMAL-MECHANICAL MODELING OF GAN HFETS AND MOSHFETS

A Thesis
Presented to
The Academic Faculty

by

William T. James

In Partial Fulfillment
of the Requirements for the Degree
Masters of Science in the
G.W. Woodruff School of Mechanical Engineering

Georgia Institute of Technology
August 2011

ELECTRO-THERMAL-MECHANICAL MODELING OF GAN HFETS AND MOSHFETS

Approved by:

Dr. Samuel Graham, Advisor
School of Mechanical Engineering
Georgia Institute of Technology

Dr. Baratunde Cola
School of Mechanical Engineering
Georgia Institute of Technology

Dr. Yogendra Joshi
School of Mechanical Engineering
Georgia Institute of Technology

Date Approved: 6/24/2011

To Mum, Dad, Chyd, Faith, Fiona, Fuzz, Greg, Ezz and Eddie

ACKNOWLEDGEMENTS

With this document being so long already it seems a crime to add any more words to it. However, I will make a few acknowledgements.

I'd of course like to thank Dr. Samuel Graham as he has allowed me to pursue my ideas on this project and has provided guidance and support through it all. His professional manner and quick insights set a high standard. Also, Dr. Anusha Venkatachalam's teaching has been invaluable in my understanding of these devices. I'll thank Dr. Adam Christensen because this work was based so extensively on what he did. His help along the way has been critical.

Of course, I wouldn't want to forget my group members, Hossein Sojoudi who sat behind me and provided the occasional hilarious comment or piece of fruit, which ever was useful at the time, Dr. Jimmy Granstrom who was always seeking to build friendship over at the buffet, and Fernando Reiter whose thesis was even more enviously long than mine. Also thanks to the other members of my group who supported me with suggestions on presentations and help at different times.

I must thank the sponsors of this work, the Office of Naval Research, for providing generous support to Dr. Graham and EMRL so that this research could be completed.

Furthermore, I'd like to thank all my friends here in Atlanta who have made this place amazing while I've been here. I wouldn't attempt to list all of you (for fear of this acknowledgements section dominating the entire document), but know that I love this place because I love all of you.

Lastly, I'll thank the Creator of all stories and the Architect of the incredible physics we're attempting to understand here. It is His work that has led me here and He must receive the credit. While His ways are difficult (or even impossible) to understand, it is His sacrifice, character, and story that captivate me. What better god to follow and praise than one who is willing to not only befriend those who are far below him, but who is willing to die for them.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF COMMONLY USED SYMBOLS AND ABBREVIATIONS

$A_{\boldsymbol{qq'q''}}^{ss's''}$	Anharmonic scattering strength for a given phonon process
N_0	Number of unit cells
e^0	Equilibrium phonon energy
k_B	Boltzmann constant
y_k	Weight or test function for finite element method variable
$\boldsymbol{c}^i$	Thermal gradient induced term in the matrix form of the BTE
$\boldsymbol{e}^i$	Energies in direction i in matrix expression of BTE
$\vec{\boldsymbol{n}}$	Boundary normal vector
α_k	Boundary directional indicator or direction cosine correction
τ_r	Relative tolerance
$\hbar$	Reduced Planck's constant
Kn	Knudsen number
Γ	Diagonal portion of the scattering matrix
Λ	Off-diagonal portion of the interaction matrix
Ω	Solid angle, or domain of integration for FEM
B	Reflection coefficient
C	Volumetric specific heat capacity
G	Volumetric heat generation rate
I	Phonon intensity
K	Linearized BTE scattering matrix
P	Scattering matrix
R	Residual for finite element method
T	Lattice temperature
V	Volume of the unit cell

a	Coefficient of fully discretized BTE
e	Phonon energy
g	Group or band subscript
k	Bulk thermal conductivity
n	Phonon occupation number
s	Phonon polarization number
u	Dependent variable of a partial differential equation system
w	Weighting term for quadrature integration
$\boldsymbol{A}$	Scattering matrix in the matrix expression of the BTE, coefficient matrix of fully discretized BTE
$\boldsymbol{L}$	Lower triangular matrix in LU decomposition of $\boldsymbol{A}$
$\boldsymbol{M}$	Preconditioning matrix for full system solution via iterative methods
$\boldsymbol{Q}$	Heat flux vector
$\boldsymbol{R}$	Full reflection matrix
$\boldsymbol{U}$	Upper triangular matrix in LU decomposition of $\boldsymbol{A}$
$\boldsymbol{W}$	Conservation condition matrix
$\boldsymbol{b}$	Equilibrium term in the matrix form of the BTE
$\boldsymbol{b}$	Full right hand side for fully discretized BTE
$\boldsymbol{e}$	Full energy vector for fully discretized BTE
$\boldsymbol{q}$	q-point representing phonon wave-vector
$\boldsymbol{s}$	Direction cosine
$\boldsymbol{v}$	Phonon group velocity
$\boldsymbol{x}$	Stacked columns of the group wise scattering matrix *P*
$\mathbb{W}$	Linear basis for the null space of $\boldsymbol{W}$
δ	Delta function, or identity tensor
θ	Polar angle of direction

σ	Stefan-Boltzmann constant for phonons
τ	Phonon relaxation time
φ	Shape function for finite element method, or azimuthal angle of direction
ψ	Perturbation or disequilibrium term
ω	Phonon frequency

SUMMARY

High power Gallium Nitride (GaN) based field effect transistors are used in many high power applications from RADARs to communications. These devices dissipate a large amount of power and sustain high electric fields during operation. High power dissipation occurs in the form of heat generation through Joule heating which also results in localized hot spot formation that induces thermal stresses. In addition, because GaN is strongly piezoelectric, high electric fields result in large inverse piezoelectric stresses. Combined with residual stresses due to growth conditions, these effects are believed to lead to device degradation and reliability issues.

This work focuses on studying these effects in detail through modeling of Heterostructure Field Effect Transistors (HFETs) and metal oxide semiconductor hetero-structure field effect transistor (MOSHFETs) under various operational conditions. The goal is to develop a thorough understanding of device operation in order to better predict device failure and eventually aid in device design through modeling.

The first portion of this work covers the development of a continuum scale model which couples temperature and thermal stress to find peak temperatures and stresses in the device. The second portion of this work focuses on development of a micro-scale model which captures phonon-interactions at the device scale and can resolve local perturbations in phonon population due to electron-phonon interactions combined with ballistic transport. This portion also includes development of phonon relaxation times for GaN. The model provides a framework to understand the ballistic diffusive phonon transport near the hotspot in GaN transistors which leads to thermally related degradation in these devices.

CHAPTER 1

INTRODUCTION

1.1 Importance of Gallium Nitride HFETs

Continual demand for improvement in amplifier technology is driving the recent explosion in Gallium Nitride (GaN) Heterostructure Field Effect Transistors (HFETs). Today's communications devices must operate over longer ranges, maintain high bandwidth, and facilitate small directional antenna size. Long range communications devices require high enough power to overcome losses in transmission. High bandwidth and small antenna devices are dependent on low wavelength, high frequency electromagnetic waves, such as microwaves, which are broadly defined as having frequencies between 300 MHz and 300 GHz.

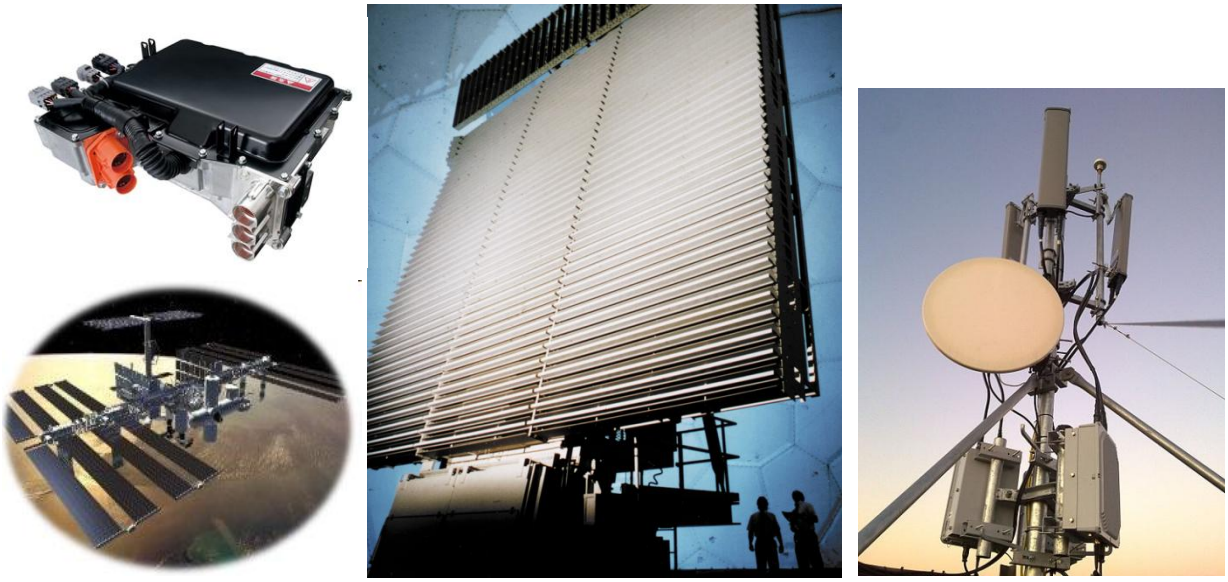

Figure 1.1. AlGaN/GaN HFET applications include wireless and satellite communications, solid-state radar, and solid state power electronics for hybrid or fuel-cell vehicles.

Current high bandwidth communications technologies that use high power amplifiers in the microwave spectrum are mobile phone or data networks like GSM (operating usually from 0.85 to about 2.3 GHz), WiMAX data networks (between 2 and 11 GHz), and satellite communications systems (about 4 to 40 GHz). In addition to

communications, modern radar systems use amplifiers generating microwaves at high powers, with applications in air traffic control, weather, and the military. Another application of GaN based HFETs is in power electronics used in hybrid electric vehicles or power distribution as the focus shifts towards achieving higher efficiencies. GaN based HFETs used as amplifiers can operate at high enough powers and frequencies to replace some non-solid state components in these technologies and promise to afford better efficiencies, smaller size and better integration. While for lower power applications other materials are more appropriate, no other materials are capable of operating in the kW range for a wide range of frequencies, see Figure 1.2.

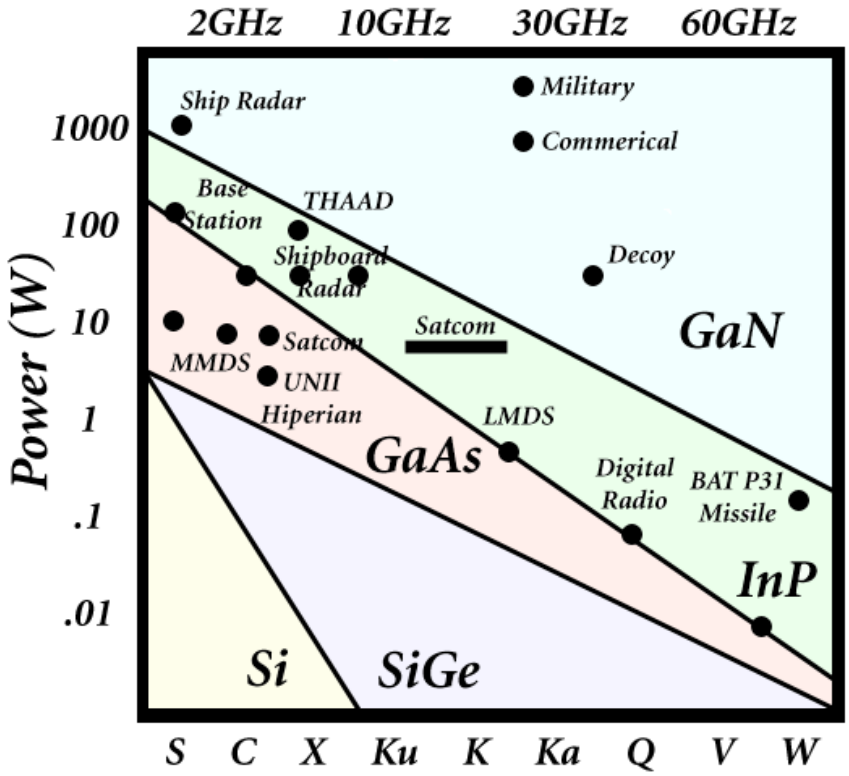

Figure 1.2. Power vs. frequency plot showing applicability of various material systems. GaN is capable of operating at higher powers and frequencies than any other material system.

1.2 Introduction to Gallium Nitride Heterostructure Field Effect Transistors

GaN HFETs are transistors used as the active component in amplifiers. The effectiveness of amplifiers based on GaN lies in a combination of features which make

this material uniquely suited to these applications. GaN has a wide electronic band-gap (3.4 eV), which leads to high breakdown fields on the order of 3.3 MV/cm allowing devices based on GaN to run at high voltages before impact ionization occurs. In addition, GaN is highly piezoelectric material and forms spontaneous polarization charges which facilitate the formation of an HFET. HFETs are FETs that operate without using doping to form a current channel. Instead the current channel is formed by taking advantage of band-shifting at the interface between two dissimilar materials, AlGaN and GaN. The orientation of the layers in these devices causes a buildup of electrons at the AlGaN/GaN interface forming a thin sheet of electrons known as the two-dimensional electron gas (2DEG) with electron densities on the order of $10^{13}cm^{-2}$[1] which are five times higher than comparable GaAs devices. The channel has high electron mobilities because there are no dopants present as there are in standard FETs, and therefore there is less ionized impurity scattering. Both the high electron densities and carrier mobilities contribute to high current and power throughput in GaN HFETs. GaN also has a high thermal conductivity (250 W/m/K [2]), which helps to dissipate heat, and it is thermally stable at high temperatures (up to 600 ºC [3]). All of these features are critical to obtaining high efficiencies in amplification[4]. A typical device structure is shown with a representation of the 2DEG in Figure 1.3.

When a voltage is applied to the source and drain, current flows through the open channel. Applying an electric field to the gate causes changes in the electron concentration in the 2DEG, which closes or opens the channel, thereby shutting off or encouraging electron flow. The effectiveness of the gate in controlling the current flow in the channel is quantified by the transconductance which is the ratio of drain current to effective gate voltage.

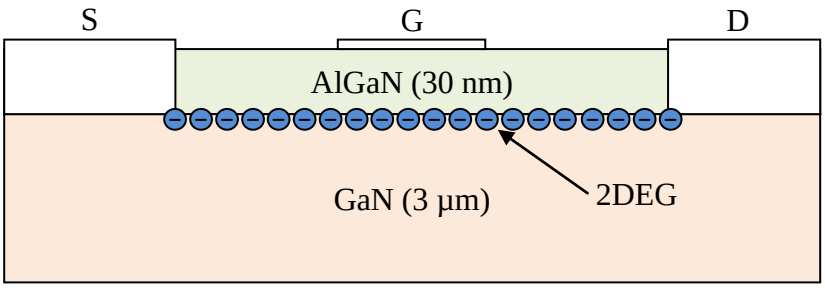

Figure 1.3. Typical schematic cross-section of an AlGaN/GaN HEMT device showing the 2DEG, AlGaN layer, GaN layer, source (S), gate (G) and drain (D).

GaN HFETs are typically larger than corresponding silicon devices, with gate to drain spacing on the order of 5 to 20 µm, as compared to 50-100nm[5]. In addition they are typically constructed with alternating source gate drain contacts into multi-finger devices with finger widths around 100 to 150 µm (width represents the longer dimension, counter-intuitively, while length is parallel to the direction of electron flow). Figure 1.4 shows a top view of a 6 finger HFET under a microscope with the source, gate, and drain connections labeled.

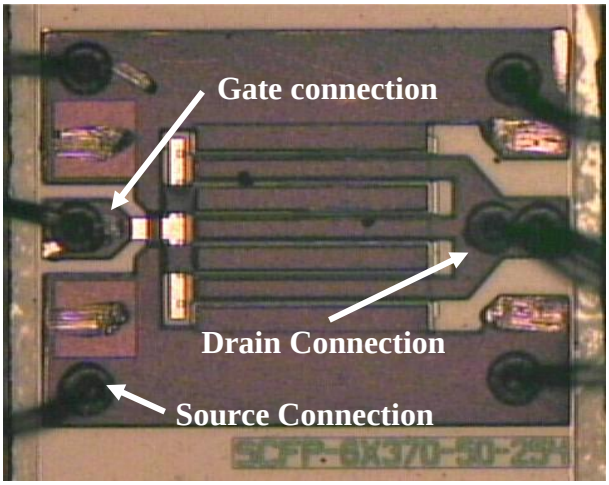

Figure 1.4. Microscope image of a 6 finger GaN HFET with source, gate and drain connections labeled.

During operation GaN HFETs can dissipate considerable amounts of power, beyond 40 W/mm at 4 GHz[6], which specifies power per unit width of channel in the device. Most of this power is dissipated by high speed electrons interacting with the GaN crystal lattice and causing generation of high energy longitudinal optical (LO) phonons[7]. This interaction occurs most where electron velocities are highest, which

occurs where electric fields are highest. Peak electric fields in GaN devices occur in the channel at the drain side of the gate, and most of the voltage is dropped over a small area, causing high peak temperatures. More recent device designs include a field plate, which spreads the drop in voltage over a larger area, reducing peak electric fields and thereby increasing the size of the hotspot[6, 8]. This improves many of the electrical and thermal characteristics of the device, leading to the possibility of higher powers[9]. Figure 1.5 shows an SEM image of a field plated HFET with the hot spot identified on the drain side of the gate by a red dot.

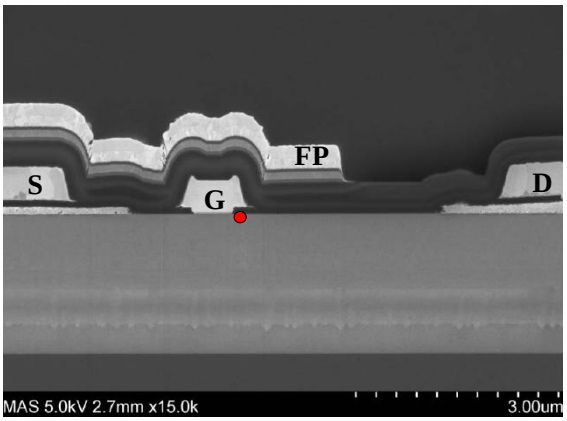

Figure 1.5. Cross-section SEM image of a GaN HFET, with field plate (FP) above the source and gate; the red dot shows the location of the hotspot[5].

In addition to having a large amount of power dissipated in a small region, non-equilibrium phonon effects are believed to affect device temperatures. The LO phonons generated by fast moving electrons decay relatively slowly and have relatively low velocities, and so they build up around where they are generated causing what is known as a phonon hotspot in the active region of the device. The slow decay of LO phonons can be attributed to the phonon band structure of GaN. In order for a phonon to decay, it must either combine with other phonons, or split into other phonons, respectively called fusion and fission processes. These processes obey simple momentum and energy conservation rules. Because Gallium and Nitrogen have a relatively large mass ratio of 4.98, there is a gap in the energy spectrum of phonons in GaN, meaning that intermediate

energy phonons which would contribute to decay of high energy phonons are not present. This hinders the decay of high energy modes and leads to their longer lifetimes.

The phonon hotspot is small but is likely to have much higher lattice temperatures than the surrounding area during device operation. In fact some groups have performed experiments which indicate optical phonon temperatures are as high as 1000 K in operational AlGaN/GaN HFETs[10]. The slow decay of LO phonons into other phonon modes that have higher velocities that can carry heat more effectively leads to what is known as a phonon bottleneck[2, 11]. Quantifying the size and temperature of the hot spot as well as the magnitude of the phonon bottleneck is difficult by measurements since the dimensions of the hotspot are small and in any case the hot spot is usually blocked from optical pathways by the gate and/or field plate in the device. None of these non-equilibrium effects are captured with a traditional Fourier diffusion model.

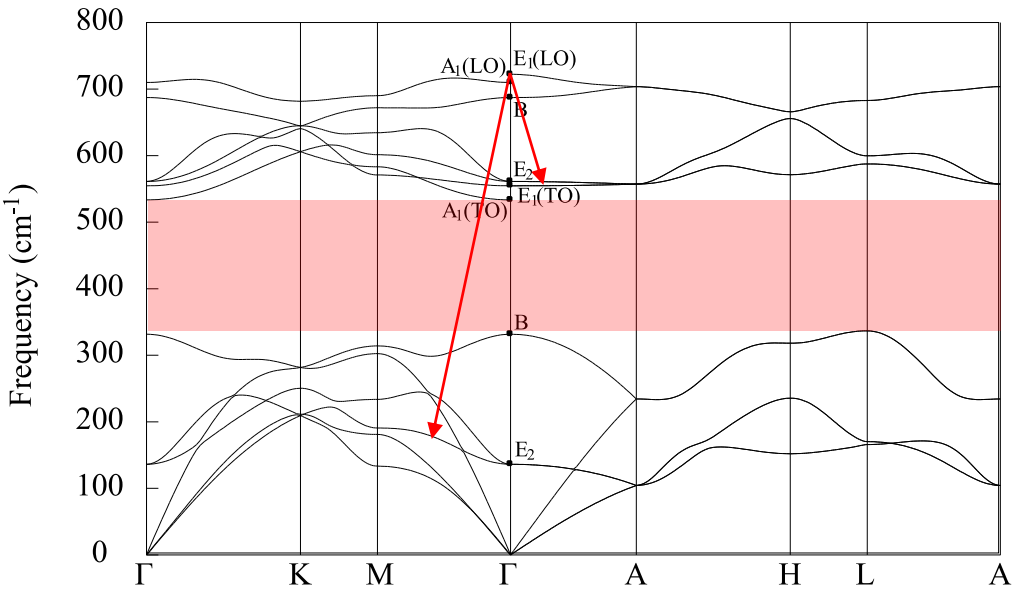

Figure 1.6. Phonon Dispersion curves for Gallium Nitride as calculated by ABINIT; also showing a representation of the decay of LO phonons into lower energy modes (red arrows), and the phononic bandgap (shaded red area).

1.3 Reliability Issues in GaN HFETs

As yet, GaN HFETs have not seen extremely widespread application at high powers mostly due to reliability concerns[12]. Devices degrade relatively quickly after

fabrication under heavy electrical stress. Kim et. al. looked at the degradation of AlGaN/GaN HFETs under DC and RF stresses of different magnitudes[13]. Figure 1.7 shows the result of running three devices under different stress conditions. Current appears to degrade significantly over a time scale of operation of only a matter of hours in some cases, especially for high bias conditions.

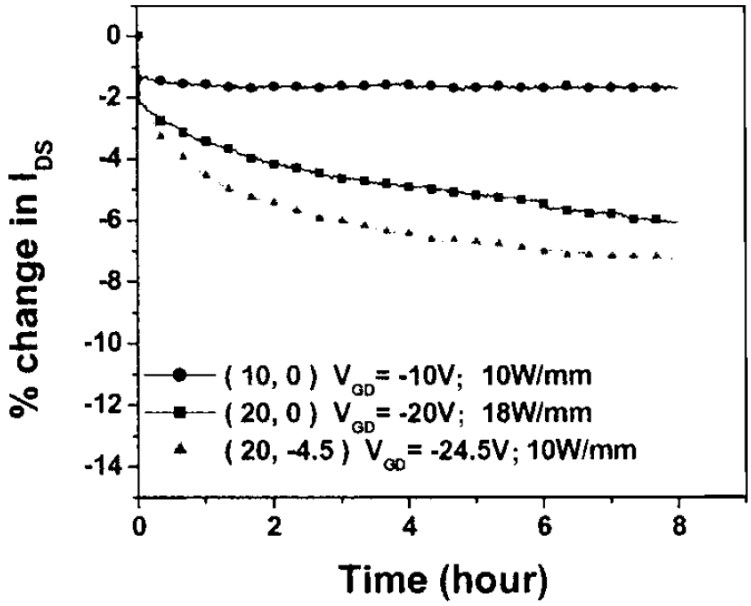

Figure 1.7. Current degradation in AlGaN/GaN HFETs under DC operation[13].

The reliability issues in GaN HFETs are multi-faceted and not well understood. It is likely that thermal degradation, thermal stress, inverse piezoelectric stress, and residual stress all play a role in device reliability.

1.3.1 Electrical Degradation

One of the key electrical degradation mechanisms that has been proposed is hot-electron degradation; that is, rapidly moving electrons can gain high enough kinetic energy to be ejected from the 2DEG into the AlGaN or passivation layers and cause permanent damage and change device operating parameters. This effect leads to what are known as traps which open up undesired current pathways. However, it is not clear that

this is actually being observed in AlGaN/GaN HFETs in the same way that it is being observed in other III-V FETs such as GaAs FETs[14].

1.3.2 Thermal Degradation

The reliability of many devices and materials systems can be fairly well understood by looking at an activation energy of failure and tying the operating temperature to a Mean Time to Failure (MTTF) to estimate device lifetime via the Arrhenius equation. Testing devices at elevated temperatures until failure and then extrapolating the results to lower temperatures through use of the Arrhenius equation is a commonly accepted way to predict failure time. However, this approach gives mixed results in GaN HFETs with some studies concluding that device lifetime is related to operating temperature, while others conclude it is not. Figure 1.8 shows a log-log plot of failure rate of GaN HFETs running at 6 W/mm for three different channel temperatures from Sozza et. al.[15], which does not show any clear trends in the temperature dependence of failure rate.

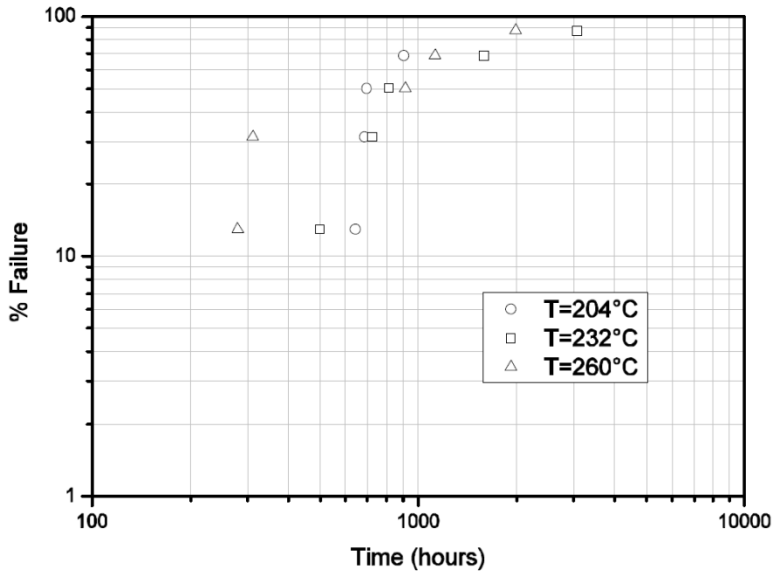

Figure 1.8. Failure rates GaN HFETs run at 6 W/mm for 3000 hrs[15]. There is not an obvious temperature trend.

It is possible that some of these effects are derived from the method some experimentalists use to derive the channel temperature. Most are unable to measure the channel temperature directly, and must measure the temperature of the device away from the channel and back-calculate the channel temperature via some kind of thermal simulation[15-17]. Deviations in the simulation results from the actual channel temperature will lead to unpredictable failure rates. In addition the models they use to predict peak temperatures do not account for any non-equilibrium effects and it is not clear how hot phonons contribute to device failure. What is clear is that simple Fourier type thermal effects are not the only drivers in device failure.

There are likely several ways in which increased temperature leads to device failure. These phenomena include channel relaxation and channel donor passivation[3], gate-drain surface charge accumulation (which can be greatly reduced by surface passivation)[12, 17], and metal contact diffusion (which seems to be resolved)[12].

1.3.3 Mechanical Degradation

In addition to degradation related to device temperature, there are mechanically related degradation phenomena. Mismatched thermal expansion coefficients between layers lead to build up of stress with changes in temperature. This effect combines with compressive stresses resulting from localized heating. In addition, since the AlGaN and GaN layers are strongly piezoelectric, when an electric field is applied, these materials strain significantly. The peak of the electric field occurs vertically at the drain side of the gate, and can be on the order of 7.3 MV/cm[18] due to the extreme thinness of the AlGaN layer. This leads to lateral strains on the order of 0.0015 which could be enough to relax the lattice and introduce electron traps, thereby reducing the sheet carrier charge and the device performance[18-19]. Makaram et. al.[20], among others, have verified the formation of pits at the drain side of the gate through the use of TEM imaging (see Figure 1.9) after operation under pinch-off, which suggests that this effect is a major cause of

device failure. This effect is likely much more pronounced when the devices are in a pinch-off condition (highly negative gate voltage, highly positive drain-source voltage), since the electric fields are higher.

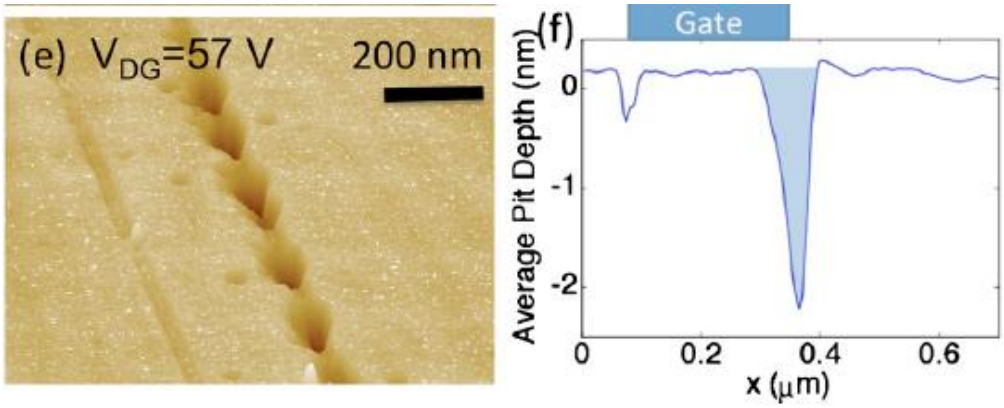

Figure 1.9. AFM scan showing pitting at the drain side of the gate (after it has been etched off) in a GaN HFET after stressing in a highly pinched off state[20].

Finally, residual stress contributes significantly to the background stress level within GaN HFETs. These residual stresses come from a number of sources: thermal expansion mismatch between adjacent layers during cooling after high temperature growth and lattice mismatch of interfacing materials. There is inherent epitaxial strain throughout the AlGaN layer since it is thin and has a slightly smaller lattice constant (3.15 Å) than GaN (3.169 Å) and is therefore under tensile strain when epitaxially grown on GaN[21]. This effect is actually at least partially responsible for the formation of the 2DEG since AlGaN and GaN are piezoelectric materials[22]. Stress related degradation phenomena include tensile failure like brittle fracture and changes in the 2D electron gas due to changes in the piezoelectric polarization charge at the interface of AlGaN and GaN in heterostructured devices. At this point, very few studies have investigated the connection between stress, electrical performance, and device degradation.

1.4 Research Overview

This work seeks to expand the understanding of coupled GaN HFET physics in order to improve their reliability through the use of computer simulation. With

continuum scale temperature and stress, modeling methods have been previously developed, but the strongly coupled nature of the physics and the small scales and large aspect ratios at hand demand better modeling techniques than have previously been employed. In other cases, the physics are fairly well understood but modeling methods have not yet been fully developed for the GaN material system.

Chapter 2 of this thesis further introduces GaN HFETs and discusses much more about the materials, structure, and measurement and modeling techniques used to study them.

Chapter 3 focuses on studying continuum level effects such as the peak temperatures and stresses within the devices while they operate under different conditions with the software package COMSOL combined with an electrical device simulator Sentaurus, which solves the drift diffusion and hydrodynamic electron models. The main outcomes of the continuum scale modeling include the following items.

- The effect of different bias conditions on the shape of the heat generation region, which has not been studied before, is determined accurately. Previous studies have assumed a constant heat generation region of varying size that is fitted to experiments on the devices and cannot predict some effects.
- The effect of inclusion of temperature dependent thermal conductivities is studied.
- The effect of using different oxide materials for gate insulation in MOSHFETs is investigated. Limited modeling has been performed on this in the past.
- The thermal expansion mismatch and thermal gradient stress contributions to device stress are studied, forming part of the picture of the device stress state during operation.

Chapter 4 focuses on quantifying further the effect of the phonon bottleneck discussed above through micro-scale thermal modeling that accounts for ballistic phonon transport and phonon dispersion effects. In GaN devices the peak equivalent phonon temperatures of the hotspot are not known with a high degree of certainty. There are three commonly used general methodologies for modeling phonon transport. The first is Molecular Dynamics simulation, wherein individual atoms and their interactions are modeled in a time dependent system that calculates atomic coordinates vs. time. This type of modeling is prohibitively computationally expensive for larger domains the size of GaN HFETs, although the accuracy of the results are unrivaled. The second method is Monte Carlo simulation which aims to solve the full scattering Boltzmann Transport Equation by modeling only a random small sample of the full phonon interactions and extrapolating the results to determine temperature and other quantities. This method is also computationally very expensive since a huge number of interactions must be modeled in order to extract reasonable results. The third method is to approximate the scattering term in the BTE and then discretize and solve it as one would any partial differential equation. Particularly, the directional propagation of the phonons is discretized, see Figure 1.10.

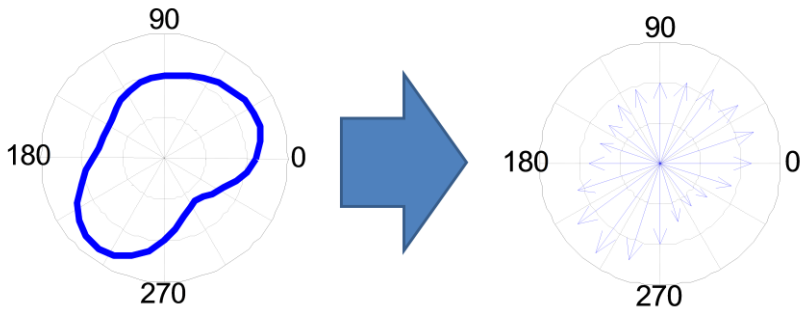

Figure 1.10. Representation of the discretization of directional phonon propagation.

The different directional discretization schemes are really methods of performing an integral over the entire solid angle with different quadrature schemes. These quadrature schemses are important in the accuracy of the BTE solution. Furthermore, the

in the case of models that account for some level of phonon dispersion, the phonon spectrum must also be discretized, commonly into bands or ranges of frequency, see Figure 1.11.

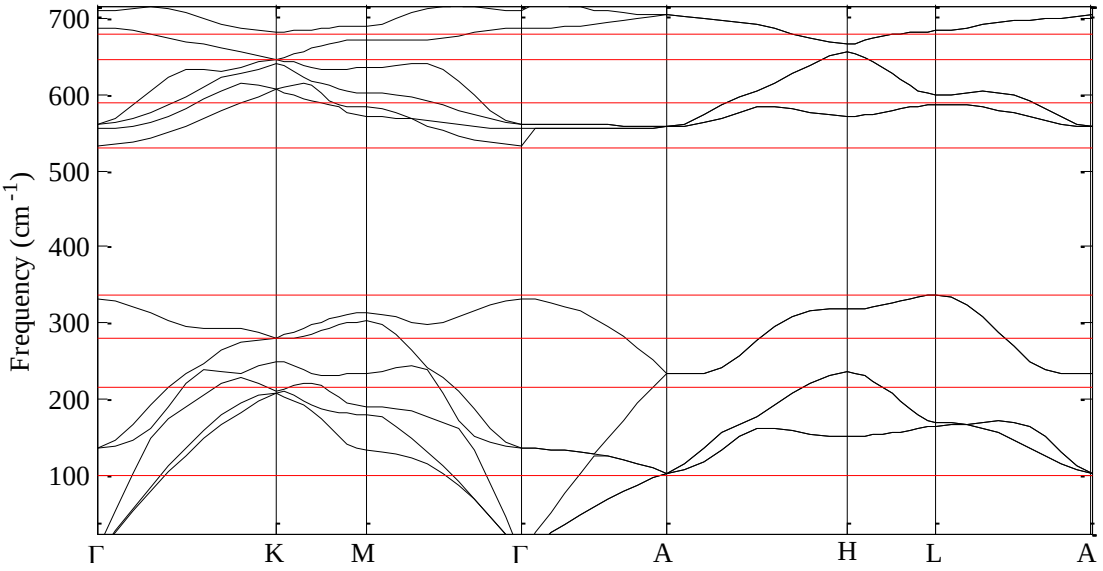

Figure 1.11. Representation of the discretization of the phonon spectrum into frequency ranges or bands. In this case the phonon spectrum is divided into 9 bands, one of which falls entirely in the bandgap, resulting in an empty band (between approximately 337 and 530 cm^{-1})

Christensen constructed a two fluid Lattice Boltzmann Method (LBM) model which captured some dispersion effects and some ballistic effects and offered a first insight into the nature of phonon transport within GaN[2, 23]. The two fluid model represented optical phonon modes as a non-propagating reservoir mode, and the acoustic modes as the propagating modes that were allowed to propagate in 8 discrete directions. Both of these are big assumptions; optical modes do have some measure of mobility in GaN, there are as many interacting phonon modes as there are atoms in a crystal, and phonons propagate in many more than 8 directions within a material. This model therefore resulted in a large overestimation of peak temperatures within the devices[2]. In addition to developing the first micro-scale model in GaN, Christensen developed a multi-scale coupling procedure for the LBM which greatly reduced computational time for large domains. This procedure coupled a LBM domain to a Fourier diffusive domain

so that the majority of a device could be modeled with traditional heat transport, while the hot-spot region where non-equilibrium and ballistic effects are important could be modeled with LBM.

A large amount of effort has been expended trying to capture dispersion effects within Silicon devices using a variety of sub-continuum thermal modeling techniques, including Lattice Boltzmann Method (LBM), Finite Volume Discrete Ordinates Method (FVDOM), Monte Carlo, and Molecular Dynamics (MD) simulations. Mazumdar and Majumdar investigated phonon transport in Si including dispersion and polarization effects with Monte Carlo modeling[24]. This work used relaxation time information derived from the work of Klemens and Holland[25-27], and matched well with experimental measurements of heat transport within thin Si films, although in that study the contribution of the optical phonon modes was neglected. Extensive work has been done by Narumanchi, Murthy, and Amon in the area of sub-micron heat transport in Silicon accounting for phonon relaxation information. Narumanchi et. al.[28] developed a fully dispersive model for Silicon again based on relaxation time information from Klemens and Holland. They compared the diffusive, grey, two fluid, and their full dispersion model in simulations of Silicon transistors and conclude that diffusive and grey models under predict, while the two fluid model greatly over predicts the hotspot temperatures[29]. Chapter 2 includes more discussion of the micro-scale thermal modeling that has been done in the past.

Chapter 4 of this thesis focuses on determining relaxation times and using them to calculate full phonon interactions in combination with the use of FVDOM in order to quantify with more accuracy the peak temperatures and deviation from equilibrium of higher energy phonon modes. COMSOL Multiphysics is used in one part of the study to obtain gray phonon solutions with quick development time and the benefit of using preexisting meshing code which should make it possible to solve the BTE in complicated geometries. An in-house code is also constructed in MATLAB which solves a form of

the BTE which includes dispersion effects using the FVDOM. This code is capable of resolving temperatures of different phonon modes within a device given the generation within any phonon modes. The main outcomes of the micro-scale thermal modeling portion of the work include the following items.

- A micro-scale heat transport model is developed using COMSOL Multiphysics which includes multi-scale coupling. This model is the first of its kind coupling DOM to a Fourier domain.
- A band-to-band model is developed whereby information about complicated three phonon interactions are compressed into a smaller, more tractable form for use in a computer model that can make some predictions about the effect that the phonon bottleneck has on peak device temperatures.
- The high energy optical mode dispersion information is accurately determined for input into the calculation of phonon interaction terms. This is achieved through use of an *ab. initio* software code ABINIT.
- Various advances are made in implicit numerical solution methods for the FVDOM including use of the Generalized Minimum Residual Method (GMRES) for improved speed and smaller memory usage. This is an improvement over existing methods which are based on inefficient explicit Gauss-Seidel iteration.
- The effect of selection of different quadrature schemes is analyzed and discussed at some length, specifically their accuracy at different scales.

CHAPTER 2

GALLIUM NITRIDE HFETS

2.1 Introduction

Having introduced the topic of study, it is necessary now to understand some more details about these GaN HFET devices. This chapter further introduces GaN HFETs and highlights some important points about their structure, materials, measurement, and modeling analysis.

2.2 GaN HEMT Structure and Materials

As has already been discussed, GaN is ideal for high current and high power applications because it has a high breakdown field, electron mobilities, and saturation velocities[4]. In addition, AlGaN/GaN junctions facilitate the formation of a two dimensional electron gas due to polarization fields and a conduction band discontinuity[1], and GaN has a relatively high thermal conductivity of around 250 W/m/K[2]. The large sheet charge density combined with the high mobility of GaN (2000 cm^2/Vs[4]) allows for high speed and high frequency operation. Combined with the large breakdown fields, large voltages and large powers can be obtain in AlGaN/GaN heterostructures. These properties make AlGaN/GaN device systems ideal for high power RF amplifiers, with applications such as RADAR and communications.

A typical HEMT device structure is shown in Figure 2.1. There are many variations of this configuration, including different substrate materials (such as Si[30], Sapphire[31], and Diamond[32]), different stacking orders of GaN and AlGaN layers including various device capping structures, addition of other layers for further control of the 2DEG density, and different Ohmic and Schottky contact materials.

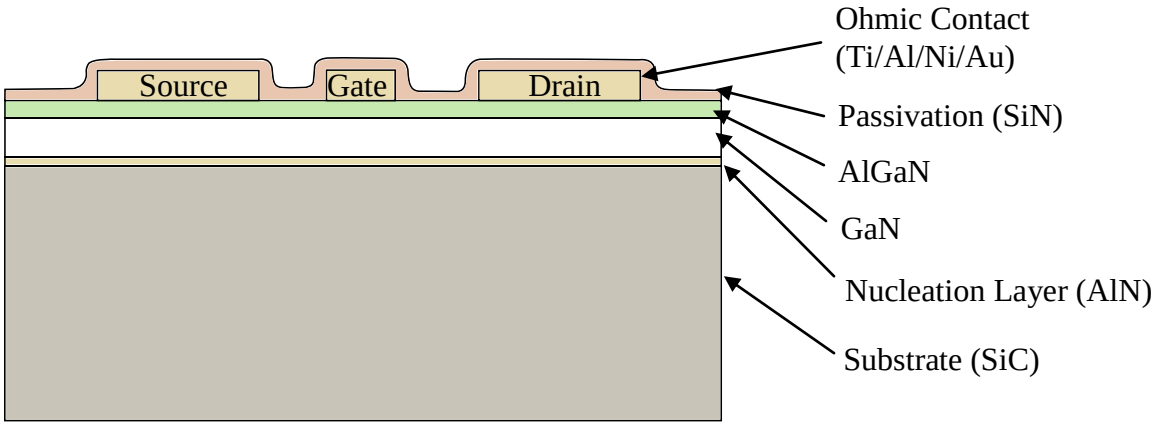

Figure 2.1. Typical AlGaN/GaN HEMT device structure (not to scale).

The substrate material is extremely important. It must facilitate growth of GaN, have a high thermal conductivity, and be stable at high temperatures. Silicon Carbide (SiC) is the most common substrate material, having a fairly similar lattice constant to GaN (3.07 Å to 3.169 Å), allowing for growth with a reasonable number of defects, and a high thermal conductivity (on the order of 330 W/m/K at room temperature) which allows heat to dissipate quickly and efficiently out of the device. A lower cost alternative to SiC is Sapphire, which has been used extensively, but it has poor thermal properties (k=32 W/m/K) and a larger lattice mismatch (a=4.785). Some groups are experimenting with growing GaN HFETs directly on Silicon, which has the advantage of CMOS integration. Silicon is also lower cost, but has a lower thermal conductivity than SiC (142 W/m/K) and a poor lattice mismatch of up to 20% and significantly different thermal expansion coefficients[33]. More recently, GaN on diamond has been investigated because of diamond's excellent thermal properties (k>1000 W/m/K). This leads to excellent heat dissipation and can lower peak device temperatures which leads to improved performance and reliability[34]. The substrate layer is typically on the order of 80 to 300 microns thick, and is bonded to a packaging layer (typically Copper Tungsten or Copper Molybdenum) with solder or thermal epoxy.

The rest of the device is typically grown with Metal Oxide Chemical Vapor Deposition (MOCVD) for the AlN, GaN and AlGaN layers, while Plasma Enhanced CVD (PECVD) is used to overlay the passivation. First, an extremely thin nucleation or

buffer layer of AlN is grown on top of the substrate to facilitate growth of the GaN layer with thickness typically on the order of 5-20 nm[35]. This nucleation layer produces higher quality, thinner GaN layers, which improve device performance by reducing substrate leakage current. However, there are some questions as to the thermal effects of nucleation layers, whether they play a large role in device temperatures or not. Interface mismatch and defects around the interface can cause phonon scattering and greatly decrease thermal conductivity in those regions. Thermal Boundary Resistance (TBR) for GaN on SiC is on the order of $3.3 \times 10^{-8} \mathrm{m}^2\mathrm{K/W}$ [36] to $1.7 \times 10^{-7} \mathrm{m}^2\mathrm{K/W}$ [37], and studies conclude that further work needs to be done in order to improve performance of GaN-SiC interfaces by decreasing defect density in the interface region[38].

The GaN and AlGaN layers are then grown to thicknesses of 1-3 microns and approximately 30nm [12] respectively. Both spontaneous and piezoelectric polarization play a role in formation of the two dimensional electron gas (2DEG). GaN and AlGaN are highly asymmetric crystals with inherent polarity. In addition, they are highly piezoelectric, meaning an electric field is induced within them upon application of a strain. When a thin AlGaN layer is grown on GaN, because of the lattice mismatch, a strain is induced in the AlGaN which along with the spontaneous polarization of both materials leads to the formation of a layer of electrons at the interface of the GaN and AlGaN[21-22, 39]. This layer of electrons becomes a current carrying sheet with extremely high electron mobilities. The mobilities are significantly higher than in FET devices because no doping of the GaN is required and therefore there is less ionized impurity scattering[4]. The piezoelectric nature of the GaN and AlGaN layers also leads to the development of stresses during device operation.

There are two types of metal contacts, Ohmic contacts (at the source and drain), where current is designed to flow in and out of the device, and Schottky contacts (at the gate), which applies the controlling electric field to switch on and off the device. Ohmic contacts are typically a stack of two to four metals of different thicknesses annealed

together. Metal configurations used in Ohmic contacts include Ti/Al/Pt/Au, Ti/Al/Ni/Au, Al/Ti, and Al/Ti/Ta, among many others[40-43]. Schottky contacts are designed to apply an electric field but prevent any current from flowing, and are typically also formed from a stack of metals which interacts with the semi-conductor so that there is no current flow due to a large difference in inherent work functions of the metal and semiconductor. Metal configurations used to form Schottky contacts include Ni/Au, Pt/Au, and Mo/Au[16, 42]. In some configurations, instead or in addition to a Schottky contact, gate insulation is used to reduce leakage current and allow operation under higher gate biases[44]. These additionally insulated devices are referred to as Metal Oxide Semiconductor HFETs (MOSHFETs). Not much effort has been expended in understanding the effect of residual stresses caused by deposition of the metal contacts, although it is likely that they do induce stresses on the AlGaN and GaN and possibly create stress concentrations that can determine the location of defect formation.

Finally, most devices are passivated in order to eliminate surface effects that limit the RF current and breakdown voltages in these devices. This passivation is typically a layer of Si_3N_4 on the order of 350 nm thick deposited on the surface of the device[45]. Passivation is also likely to contribute to device stress[46].

In some devices, in order to reduce the peak electric field at the drain side of the gate, one or more field plates are added, which smooth out the electric field profile and allow for higher voltages before breakdown. This is advantageous for driving devices at higher powers. This field plate(s) can be connected to the gate or to the source. Unfortunately, field plated devices make device measurement even more difficult, since they cover a larger portion of the active region of the device. Figure 2.2 shows gate and source connected field plated devices[6, 8-9].

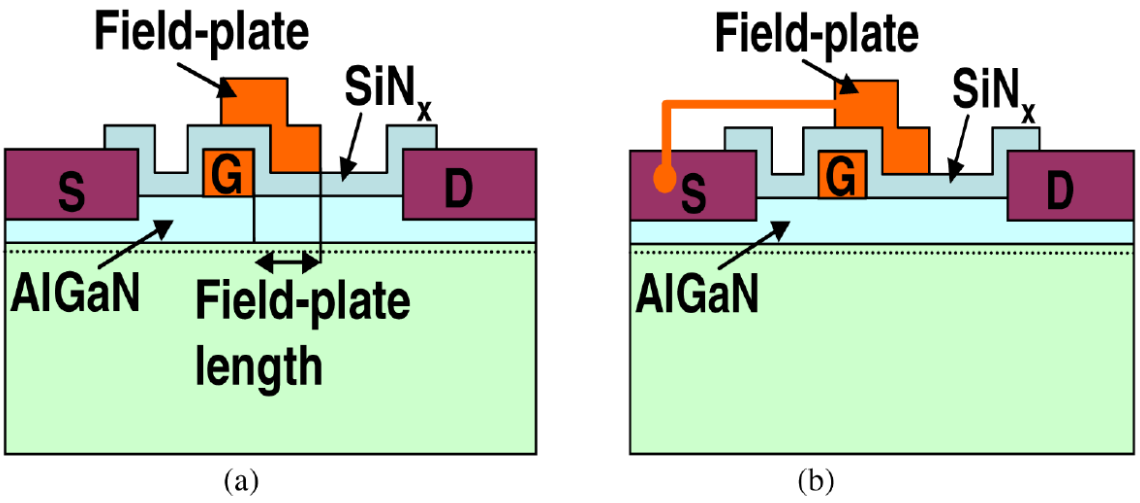

Figure 2.2. Schematics of a) gate connected and b) source connected field plates in AlGaN/GaN HFETs[4].

2.3 GaN HFET Measurement

For the purposes of device design and diagnostics as well as for predicting device lifetimes it is highly desirable to know the internal operating conditions of these devices, such as temperature and stress. Direct measurement of these properties is extremely important in understanding how AlGaN/GaN HFETs operate. These parameters affect reliability, as shall be described in the next sections.

2.3.1 Thermal Measurement

Some techniques being used to measure temperature in AlGaN/GaN HEMTs are: Raman Spectroscopy[47], Photoluminescence[48], electrical techniques such as forward voltage[49], and more recently AFM techniques are being developed that back channel temperature out from surface thermal expansion (unpublished work). These methods all have their merits and issues. The most utilized method is Raman Spectroscopy which can measure temperature within a few degrees Celsius with a resolution on the order of one micron. In addition it is capable of extraction of phonon information and stress magnitudes[50-51].

Raman works by probing the sample with a laser (e.g., 488 nm for GaN) and observing the scattering response of the crystal. When the energetic photon interacts with the crystal it typically excites an electron to a higher energy state. This electron then either immediately relaxes back to its equilibrium state by emitting a photon of energy equal to that of the incident laser (referred to as Rayleigh scattering), or it decays into or absorbs a phonon and then emits a photon of different energy (stokes and anti-stokes scattering). The change in energy between the probing photon and the returning photon can reveal valuable information about the energy and population of phonons within a material. With a change in device temperature, the energy of phonon modes shifts due to anharmonic effects. This dependence allows temperature to be extracted from Raman spectroscopy results. The change in phonon energy is also dependent on other factors such as stress and electron occupation, but by making a few assumptions and observing the shift in energy as well as scatter it is possible to deconvolute these effects and predict temperature quite accurately under a wide range of conditions.

Some reports in literature of temperature measurement of AlGaN/GaN HFETs with Raman include Kuball et. al. demonstrating the viability of the method with spatial resolutions of around 1μm and temperature resolution of less than 10 ºC[47] and application of the technique to measuring a multi-finger AlGaN/GaN HFET with temperatures on the order of 220 ºC for an 8 finger device operating at 7 W/mm and 120 W/mm for a 4 finger device operating at 9 W/mm[52]. Beechem et. al. demonstrate the necessity of accounting for the local stress state in Raman measurements[51].

However, Raman, like other methods, still cannot probe the active region of the device during normal device operation, due to the presence of the field plate and gate directly above the hot spot. Some experimentalists get around this by using Raman spectroscopy from the bottom of the device through the SiC substrate[53], which has many issues in itself, the first being that the devices cannot be operated under realistic conditions. In addition to being unable to probe the most active region of the device

under realistic operating conditions, the probing Raman laser penetrates a certain distance into the material being measured and therefore the information extracted is a volumetric average of where the laser is incident.

2.3.2 Stress Measurement

Stress is more difficult to measure than temperature, and measurement of mechanical stresses is less common, although stresses can easily have as much impact on device performance as temperature. Jungwoo et. al. look at degradation while mechanically bending a device in order to investigate the dependence of degradation on mechanical strain, and show strain is an important contributing factor to device failure[54]. Other studies were performed by Kuball's group in combination with thermal and stress modeling, which will be discussed in Section 2.4. Peak thermal and piezoelectric stresses are on the order of -0.35 GPa and 0.2 GPa respectively for a 20 W/mm device[55].

Raman is also the most common method for measuring stress. The phonon modes of a material are dependent on the deformation of the crystal lattice which occurs under most stress conditions. Raman can detect these changes and predict the relative change in stress of a crystal. The main difficulty with this technique is that Raman only allows one to determine the relative change in stress, but unless accurate values for phonon frequencies of unstressed GaN are available, absolute stresses are difficult to predict. This is still useful for determining the change in stress as a result of temperature shifts or inverse piezoelectric effects. Sarua et. al. [56] and Beechem et. al.[51] among others use Raman to study the piezoelectric stress and strain in AlGaN/GaN HFETs.

2.4 GaN HFET Modeling

2.4.1 Thermal Modeling

A fair amount of effort has been expended on thermal modeling of GaN/AlGaN HEMT devices. Initial efforts began with construction of analytical models of HEMT devices resulting in reasonably accurate estimations of device temperatures for very specific device configurations[57-58]. These studies have the advantage of being parameterized and provide some insight into the effect of heat generation and material parameters on peak temperatures and thermal resistances. However, temperature dependent thermal conductivities, full device simulations, and some changes in geometric configuration are difficult to capture in these models.

Some further studies improve upon these works by including more advanced geometries and temperature dependent thermal conductivities. Park et. al. compare results from a 3-dimensional analytical model that includes effects from 5 different layers using an in house code PAMICE to results from liquid crystal thermography measurements[59]. Eastman et. al. carry out nonlinear three-dimensional heat spreading simulations to aid in electrical device simulation to aid in experimental design[35]. Heller includes thermal effects with temperature dependent thermal conductivities in a 3D ANSYS model which predicts temperature evolution and degradation over time within a GaN/AlGaN HEMT[60]. Heller et. al. also construct a 2D/3D device model that does include effects of changing heat generation size on peak temperatures[61]. This model splits the domain into two regions, one of which is an electro-thermal model including the GaN epitaxy modeled in two dimensions using DESSIS, and the other a three dimensional model including just the substrate. The effect of multiple fingers with varying heat generation along the fingers and temperature dependent thermal conductivities are investigated. This results in the temperature profile shown in Figure 2.3 for a device running at an average of 6.79 W/mm.

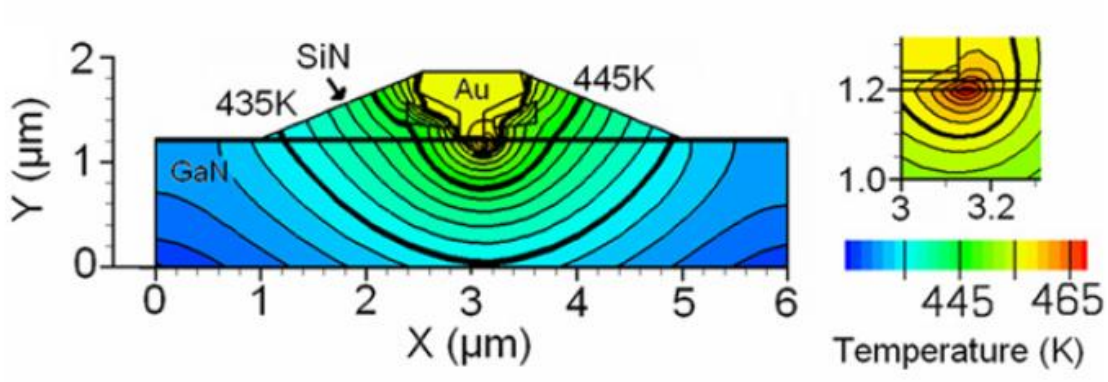

Figure 2.3. Temperature distribution for typical AlGaN/GaN HEMT showing close up of hotspot for a device running at 6.79 W/mm from [61].

Less effort has been expended on capturing micro-scale effects in AlGaN/GaN HFETs. Wang et. al. look at the size dependence of thermal conductivity of GaN nanowires using Molecular Dynamics (MD) with a Stillinger-Weber potential, seeing good agreement with experimental data[62], however this type of analysis is not particularly useful for determining peak temperatures in AlGaN/GaN HFETs because the computational demands are far too large given the device sizes in consideration. Christensen et. al. develop a multiscale model based on the Lattice Boltzmann Method (LBM) that can capture some dispersion effects with a two fluid model by assuming that heat in AlGaN/GaN transistors is generated in a high energy non-propagating optical phonon reservoir mode and then these phonons relax into heat carrying acoustic phonons that transport heat away from the optical mode[23]. This model has the disadvantage of significantly overestimating the peak temperatures because it overestimates the phonon bottleneck by assuming optical modes are non transporting and also that they have the same relaxation time as the acoustic modes. In addition, the LBM has been known to suffer from a lack of directional resolution in more ballistic transport regimes.

As little work as has been done in GaN, much more has been done in Silicon on micro-scale heat transport and many of the methods are applicable within the GaN material system. Majumdar et. al. introduced the Equation of Phonon Radiative Transport (EPRT) that connected the methods of radiation heat transport to those of

mico-scale heat transport[63]. In addition Mazumder and Majumdar[24] began exploring heat transport that included dispersive effects with a Monte Carlo model that captured dispersion and polarization information in Silicon with relaxation time information being derived from the analytical expressions developed by Klemens and later expanded upon by Holland[25-27]. Building upon this work, Narumanchi et. al. developed a Finite Volume Discrete Ordinates Method (FVDOM) code that captured energy conserving heat transport at the micro-scale[28]. They again used relaxation time expressions from Klemens[64] and were able to fit the thermal conductivity of Si with minimal fitting parameters. Figure 2.4 shows the temperature distributions obtained from a Fourier, grey, two-fluid and full dispersion model from [29] showing how grossly the two-fluid model overestimates and how much hotter a full dispersion model is than a grey or Fourier model. The full dispersion model developed by Narumanchi et. al. enforces energy conservation by assuming that modes interact at a given interaction temperature, and band-to-band relaxation times are symmetric, meaning that all the energy that passes out of one mode is exactly captured by energy passing into other modes.

While these models seem to represent with good accuracy the effect of non-equilibrium phonons in peak temperatures in transistors, the expressions developed by Klemens, Han, and Holland are not applicable in GaN where there is significantly more anisotropy. This thesis work develops a framework for finding band-to-band relaxation times or interaction strengths that do not depend on specific dispersion shapes and that should be truly representative of underlying three phonon processes.

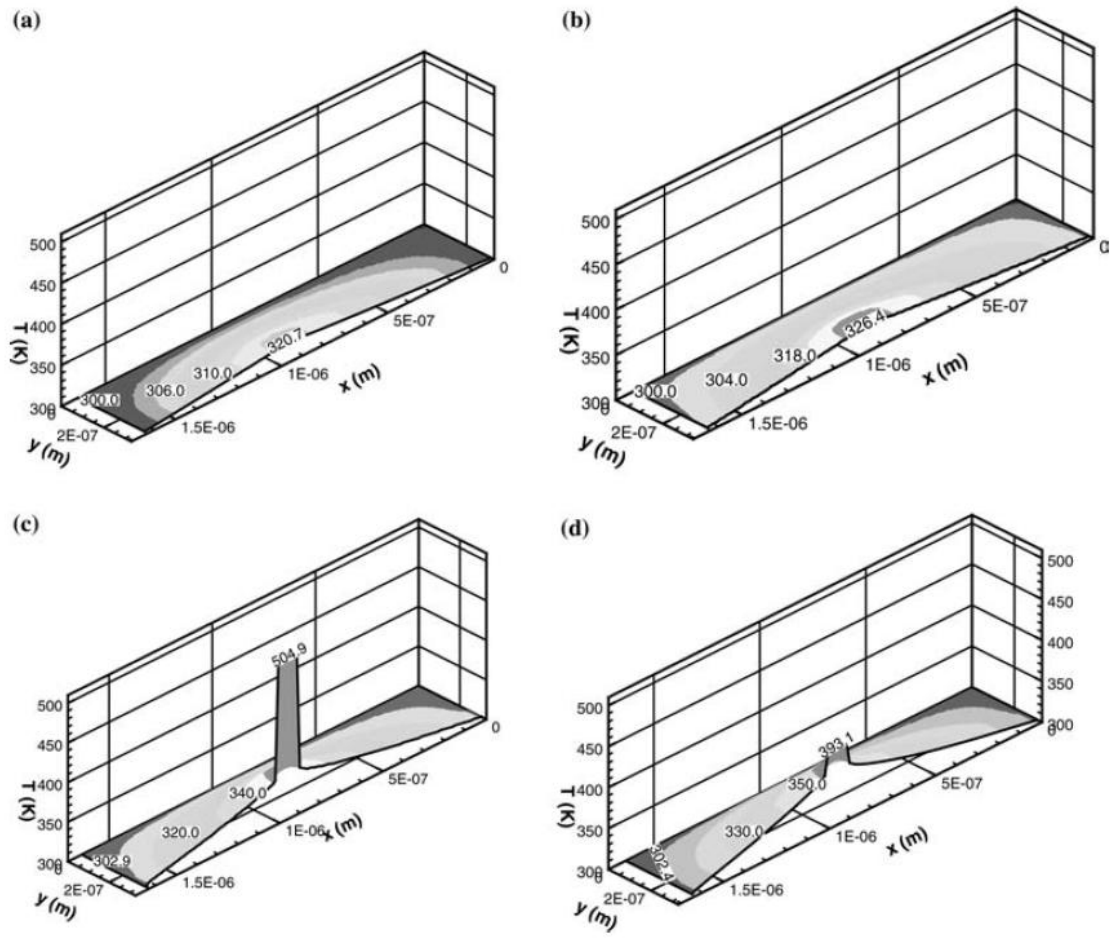

Figure 2.4. Temperature distributions in a Silicon Transistor for a) the Fourier Model, b) Grey BTE model, c) two-fluid BTE model, and d) full dispersion BTE model from [29].

2.4.2 Stress Modeling

Less effort seems to have been expended on stress modeling in AlGaN/GaN HFETs than thermal modeling. Stress modeling includes the effects of thermal expansion stresses as well as inverse piezoelectric stresses. Thermal expansion stresses include lattice mismatch and thermal gradient stresses and require knowledge of the temperature field within the device, which implies that thermal modeling is performed alongside the stress modeling. Inverse piezoelectric stress requires a knowledge of the electric field which is inside the device, which implies electrical modeling has been performed. Both of these modeling techniques rely on knowing the amount of residual stress in the devices in order to calculate absolute stress values.

Sarua et. al. study a coupled piezoelectric model during pinch-off. The shape of the results from their model match well to stress measurements using Raman, but the magnitudes are off by an order of magnitude[56]. Figure 2.5 shows the piezoelectric stress distribution within an AlGaN HFET operating at a source drain bias of 40 V as measured by Raman spectroscopy and as calculated from a coupled electro-mechanical model.

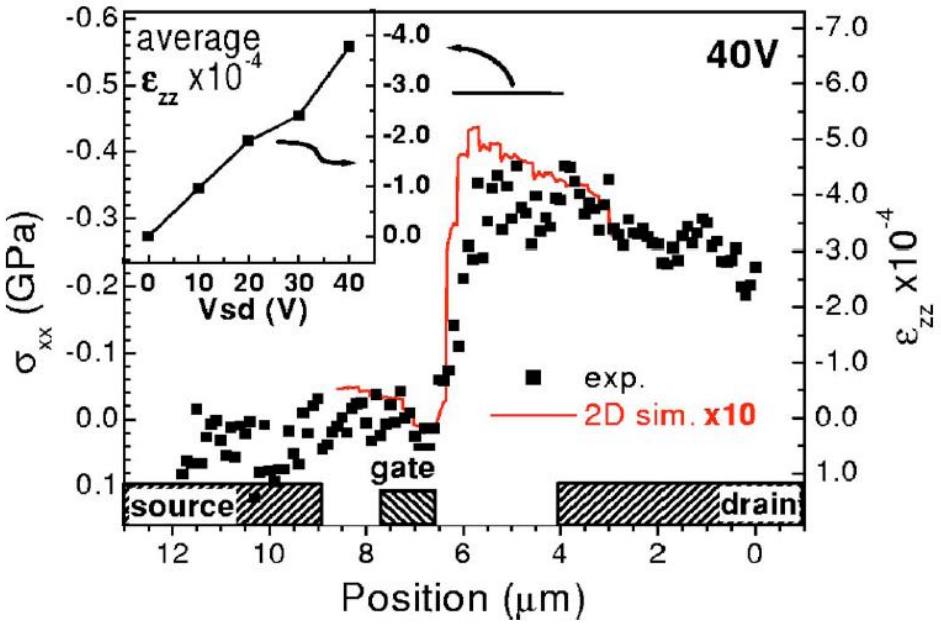

Figure 2.5. Piezoelectrically induced Strain ε_{zz} and stress σ_{xx} in the GaN layer in an AlGaN/GaN HFET as measured by shift in the E_2 phonon mode and by simulation. They disagree by an factor of 10 but the shape of the simulation appears to be correct. From [56].

A second study by Sarua et. al. looked at thermal gradient and inverse piezoelectric stresses with a finite difference model and Raman measurements and concluded that they are opposing in sign and thermal expansion can actually cause a lower net stress near the drain side of the gate during on operation[55]. Gao performs coupled electro-thermal-mechanical simulations of AlGaN/GaN HFETs to determine stresses in the AlGaN layer are on the order of 3 GPa for most bias conditions[65]. Jogai et. al. examined the strain in the AlGaN layer in an analytical electro-mechanical model and concluded that piezoelectric coupling effects are somewhat important in determining

strain and electric field within the AlGaN layer[22]. At the very least their study represents the complexity of the electro-mechanical system at hand. Mastro *et. al.* simulate the effect of passivation layers on AlGaN/GaN HEMT devices and conclude that SiN deposited with PECVD at high frequency induces tensile stresses in the SiN which, due to the discontinuity in the film at the gate, can cause large stress gradients at the gate edges in the AlGaN[46]. These stress concentrations lead to strain relaxation in the AlGaN and modification of the 2DEG concentration and therefore the electrical properties, which can lead to reliability and performance issues.

CHAPTER 3

CONTINUUM SCALE MODELING

3.1 Introduction

This chapter of the thesis focuses on studying continuum level effects such as the peak temperatures and stresses within the devices while they operate under different conditions with the software package COMSOL combined with an electrical device simulator Sentaurus, which solves the drift diffusion and hydrodynamic electron models. This chapter attempts to determine the effect of different bias conditions on the shape of the heat generation region, which has not been studied in detail before. Previous studies have assumed a constant heat generation region of varying size that is fitted to experiments on the devices and cannot predict some effects. In addition, limited study has been made of the thermal and mechanical operating conditions of AlGaN/GaN MOSHFET devices, which are examined here. Temperature, as well as the thermal expansion mismatch stress and thermal gradient stress contributions to device stress are studied, forming part of the picture of the device conditions during operation.

Continuum-scale implies the object of study is large enough to ignore atomic effects so that macro-scale physics models can be used, such as the Fourier law of heat conduction, linear elastic continuum mechanics, and linear piezo-electric effects. All of the analysis in this section is performed using a commercial finite element solver called COMSOL Multiphysics. The advantage of continuum-scale modeling is obviously in the simplicity and availability of commercial codes which are capable of performing the simulations required with little development time.

First, this chapter covers a brief discussion of COMSOL multiphysics, the main software package used in the analysis. Next the approach used in heat transport modeling in GaN HFETs is briefly introduced and the equations solved are introduced. Next, the

approach used in modeling thermally related stress is presented. Finally, the following case studies are performed with analysis of the results:

- The effect of using temperature dependent thermal conductivities within the GaN and SiC layers is analyzed.
- The effect of changing the drain to source bias on the peak temperature and stress is modeled.
- The effect of adding different oxide materials and using different oxide thicknesses under the gate on the peak temperature and stresses in AlGaN/GaN MOSHFETs is investigated. Insulated gate structures are being pursued as a way to reduce gate leakage currents and improve power output from the device.
- Finally, the effect of changing the bias conditions under constant power is analyzed. It would not be possible to observe this effect with most modeling procedures that have been used in the past to capture the size of the heat generation region. It is postulated that the changes in electric field gradient will change the region of heat dissipation and impact the temperature in devices. Such effects have yet to be considered in a full three dimensional device model in GaN.

3.2 COMSOL Multiphysics

COMSOL Multiphysics is a general purpose Finite Element Solver. It contains a graphical user interface which has full CAD, meshing, and post-processing capabilities. COMSOL contains a number of preset applications modes, which include Heat Transfer, Solid Stress/Strain, and Piezoelectric (and Inverse Piezoelectric) effects. Generally, when using COMSOL, one specifies the geometry in question, selects material parameters, meshes the domain and selects the type of elements, selects a solver or

solvers, and then solves and postprocesses the results. All of this can be done from within the COMSOL interface[66].

3.3 Heat Transport Modeling

Heat Transport in the continuum has been studied for hundreds of years since Newton first established his law of cooling in 1701[67] and Fourier wrote his famous *Théorie analytique de la chaleur* in 1822[68-69]. Fourier's law of heat transport is extremely widely used, although it is not accurate at small scales where boundary effects play a large role in heat transport properties. However, it is far simpler and easier to implement than the more detailed methods which are discussed in Chapter 4 of this thesis. It is used here to gain insight into how power input, power distribution, and device dimensions affect the temperature distribution, and to obtain a temperature distribution for coupling to a stress model.

One of the main differences between all of the models in literature is the treatment of the shape of the heat generation region; a few approaches have been employed. Some investigators use a square constant heat generation region embedded next to the gate contact on the drain side[70], while others use a constant heat flux or dual heat source on the top surface of the device[35, 57]. These approaches, while they may be tweaked to match experimental results, do not accurately capture the size of the heat generation region which is important in determining peak temperatures in these devices. These approaches are zero dimensional in that the total heat generation is calculated and then applied uniformly to an arbitrary or empirically fitted area.

Within this body of work an approach is used that captures the spatial distribution of the heat generation in one dimension as opposed to zero dimensions. The heat generation region from Sentaurus simulations is a two dimensional field which gives the spatial variation of heat generation in the x-z plane. Sentaurus is a two dimensional simulator that does not capture variations in the y direction, and it is assumed the heat

generation is constant along the width of the finger. In addition, since the spatial resolution of the Sentaurus model is significantly higher in two dimensions than is possible in the three dimensional continuum model within COMSOL, it is necessary to find a way to couple this high resolution field to a lower resolution finite element model. The method used here is a one dimensional integration in the z direction which flattens the two dimensional field into a one dimensional heat generation as a function of x position. Then this one dimensional function is applied uniformly along the width of the finger to the surface of the GaN layer. This is reasonable because the heat generation field from Sentaurus is concentrated exclusively along the AlGaN/GaN interface and is spatially at least an order of magnitude larger in the x direction than in the z direction, see Figure 3.1.

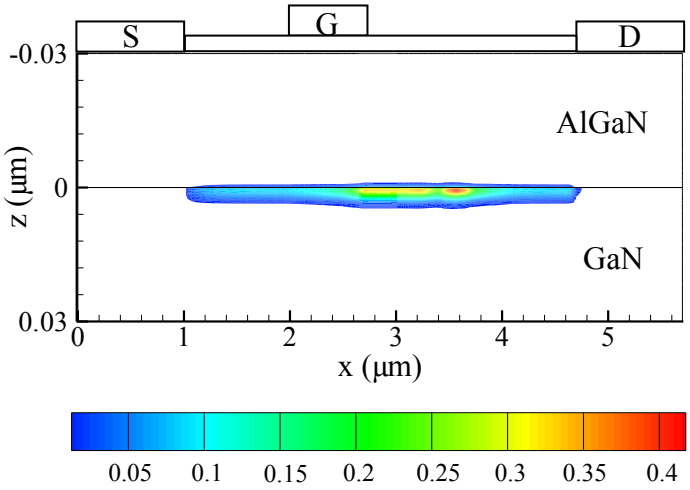

Figure 3.1. Shape of heat generation region as calculated by Sentaurus for 20nm of SiO_2 under the gate at VDS=14 V in W/μm^3.

There are some inaccuracies introduced by assuming that the heat generation is constant along the width of the finger or across different fingers. In an actual device, temperatures are lower at the edges of fingers or in fingers that are further from the middle of the device, since they are further from the center of heat generation, but there will be higher current flow and therefore more Joule heating[48]. However it has been shown that this effect is small enough to be ignored[61].

After flattening, the heat generation region is applied to the surface of the GaN in the thermal model. The applied heat generation profile for a MOSHFET with 20nm of SiO_2 is shown in Figure 3.2. This flattening is performed within COMSOL with the use of projection coupling.

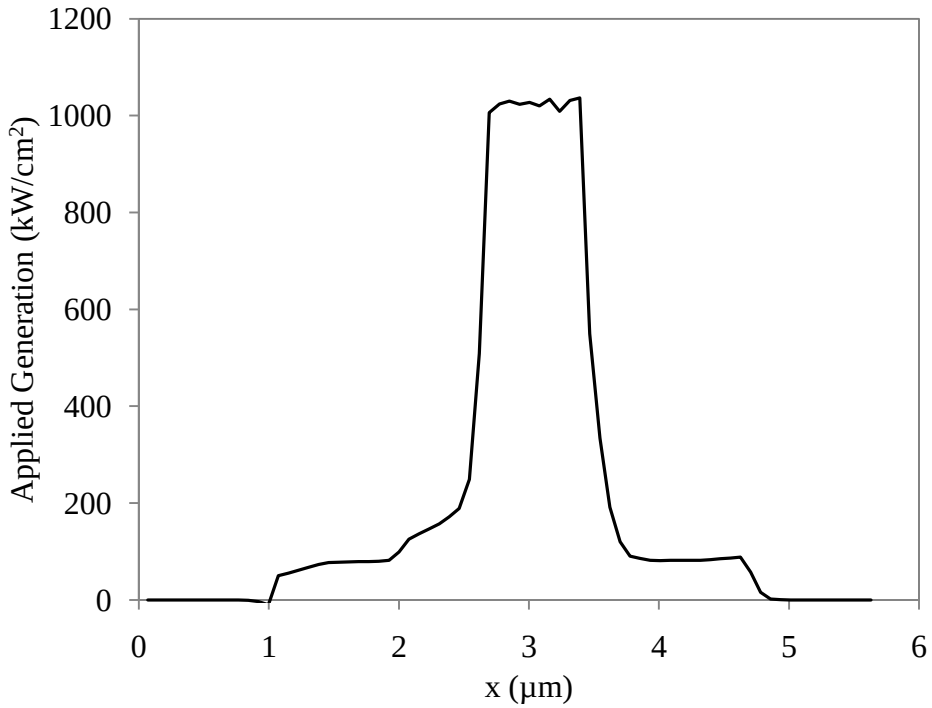

Figure 3.2. Flattened heat generation profile applied to the surface of the GaN layer in the thermal-mechanical model of a MOSHFET with 20nm of SiO_2.

COMSOL solves the Fourier diffusion equation

$$-\nabla(k\nabla T) = S \tag{3.1}$$

where T is the temperature and S is the generation at any point given the device boundary conditions, which are described in Section 3.5. In all of these studies the generation from Sentaurus is applied as a heat flux boundary condition on the top surface of the device, while the volumetric generation term S is set to zero. This is realistic because the generation region is extremely thin and almost directly under the top of the device.

3.4 Thermal Stress Modeling

There are several different kinds of stress effects that likely play a role in the reliability of GaN HEMT devices. All of these stresses act at once and form a complex system which is dependent strongly on growth conditions, operational temperatures, and electrical conditions. It is one of the goals of this work to further the understanding of thermal stresses in order to aid in failure prediction and improve device reliability.

Most of this work focuses on determining thermal gradient and CTE mismatch stresses for a variety of different bias conditions. Residual and inverse piezoelectric stress models are under development at this time. There are several major hurdles to overcome in the prediction of residual and piezoelectric stresses. Inverse piezoelectric stress is dependent on electric field which peaks in the AlGaN layer. The AlGaN layer is so thin that it is difficult to capture in a full 3D stress model. 2D models can be employed with a plane strain approximation, but will tend to over-predict stresses significantly due to constraint of expansion and contraction in the third direction. Residual stresses are dependent strongly on growth and processing conditions and the interactions of the materials at the interfaces. It is not difficult to predict the contribution that thermal expansion has to residual stress, but predicting the formation of dislocations and defects that relieve stresses as well as inter-metallic layer formation at the interfaces makes the residual stress problem one not suited to the current modeling methodology. However, preliminary results suggest that there could be significant stress concentrations at the edges of the metallic contacts due to larger thermal expansion coefficients in Ti/Al/Nu/Au stack (19×10^{-6}/K for the Ohmic versus 3.2×10^{-6}/K for GaN).

COMSOL solves the coupled linear elastic thermal-mechanical constitutive equation for the displacements given boundary conditions and temperatures within the device,

$$\sigma = D\varepsilon_{el} \tag{3.2}$$

where σ is the six dimensional stress vector, ε_{el} is the six dimensional elastic strain, and D is a 6-by-6 elasticity matrix which depends on the material. The elastic strain is related to the total strain by the following equation.

$$\varepsilon = \varepsilon_{el} + \varepsilon_{th} + \varepsilon_0 \tag{3.3}$$

Here the thermal strain is determined by the thermal expansion coefficients

$$\varepsilon_{th} = \alpha_{vec}\left(T - T_{ref}\right) \tag{3.4}$$

where α_{vec} is a vector form of the thermal expansion coefficients which can include anisotropic thermal expansion, which is included in this model. The initial strain, ε_0, is assumed to be zero. At free surfaces the stress is constrained to be zero, while at interfacing layers, the strain must be equal. The result from COMSOL is a stress field which represents the thermal gradient and thermal mismatch stresses inside GaN. The stresses presented here do not include inverse piezoelectric or residual stresses which are likely to considerably change the stress state within AlGaN/GaN HFETs.

3.5 Model Configuration

The basic device configuration for the model set up is described here, with changes described where they are applicable. The model was selected to have device dimensions matching typical AlGaN/GaN HFET devices on SiC. A typical device stack consists of a 30 nm thick AlGaN layer on top of a 3 μm GaN layer on top of a 200 μm SiC substrate attached to a Tungsten Copper package with a generic solder material of thickness 40μm. In all of the models the AlGaN layer was neglected because it is so thin it does not contribute significantly to stresses in the AlGaN, and it is at the top of the device so it does not contribute significantly to thermal resistance. Stresses in the AlGaN layer can be determined by assuming the AlGaN layer is strained exactly as the top of the GaN layer is strained.

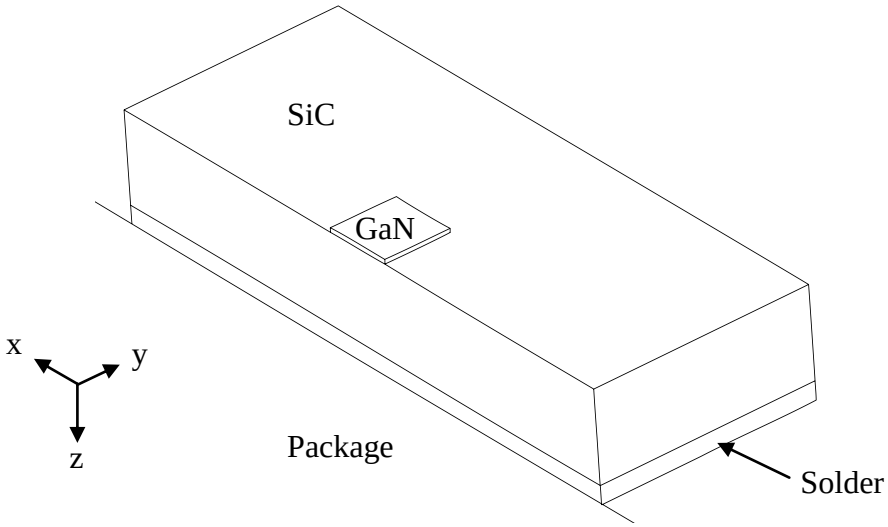

Figure 3.3. Model set-up of GaN, SiC, Solder and then some of the package is shown below. Note the size of the GaN layer which is reduced in size to save memory since the outer edges do not contribute to the thermal resistance of the package. The coordinate system is shown.

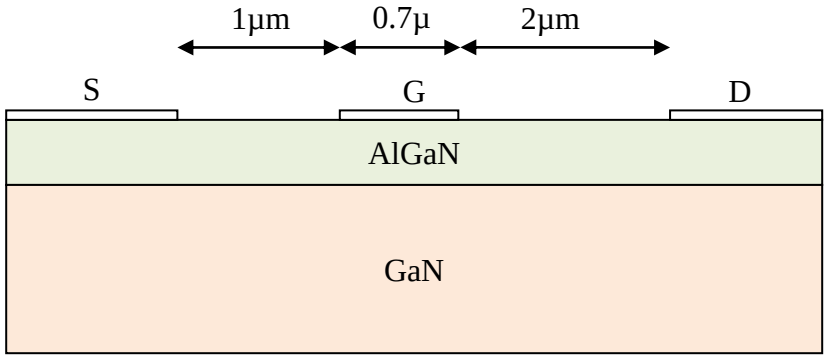

Figure 3.4. Basic device configuration for most of the cases, changes are described where applicable.

The base of the copper-tungsten package was set to 300 K and free to move physically (i.e. no thermal stress would occur as a result of package expansion being confined at the base). The substrate was assumed to be 1300µm by 980µm and in a real device the GaN and AlGaN layers would be deposited on top with the same dimensions. In the model however, the GaN layer was adjusted to a much smaller size in order to save memory since far away from the active region the GaN layer is not contributing significantly to the thermal resistance. This is visualized in Figure 3.3, along with the coordinate system used for the rest of this chapter. The GaN layer was typically modeled as a 200 µm by 200µm by 3µm layer on the surface of the SiC substrate. For the

purposes of these studies here single finger devices were modeled, since the effect of finger spacing or count was not being investigated. The symmetry of the device at the center of the finger was taken advantage of by splitting the model in half with a symmetry plane.

The device was meshed in COMSOL with tetrahedral elements. Automatic adaptive mesh refinement was used on all the models in order to resolve the peak temperature and stress areas with better resolution. A grid independence study was performed on the peak stress value to ensure that the solution was converged. Cubic lagrange elements were used for every study. A typical device mesh after refinement included approximately 73,000 elements with over 600,000 degrees of freedom for a coupled thermal and stress model. For good resolution of the temperature profile in plots such as Figures 3.9 and 3.13 a higher resolution model was used with manual refinement of the area of interest, and over 600,000 elements with over 1.2 million degrees of freedom. For most of the models the workstation used to perform the simulations had enough memory to perform a direct solution to avoid convergence issues, however with the larger simulations an iterative solver the Geometric Multigrid and GMRES was used. This did have significant convergence issues and the solver had to be tweaked to obtain convergence. A visualization of a typical device mesh is shown in Figure 3.5.

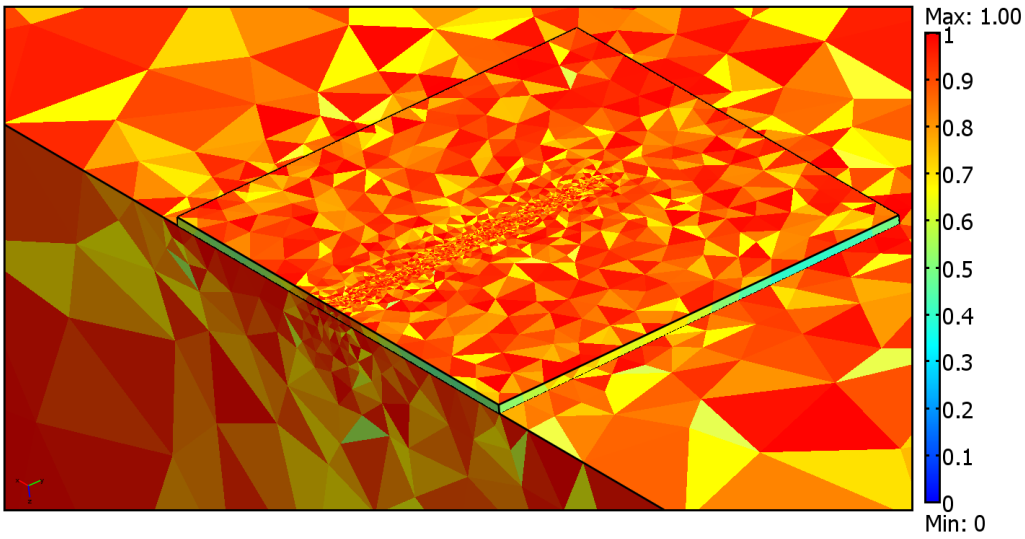

Figure 3.5. A visualization of the mesh. Note the refinement around the peak temperature region along the device finger. Coloration is showing mesh quality, with 1 being highest and 0 being lowest, as measured by aspect ratio.

For most of the studies here, the electrical simulations were performed in Sentaurus Device under steady state conditions yielding a heat generation region which was then applied to the device in COMSOL to find peak temperatures. This coupling was one-way, meaning that some of the effects of self-heating on electron transport were not captured. Sentaurus does include a limited thermal model with a small thermal domain, and was also calibrated with experiments on devices to ensure that the current and power were accurate for a given bias.

3.6 The Impact of Temperature Dependent Thermal Conductivities

The thermal conductivity of GaN decreases as a strong function of increasing temperature, which leads to nonlinear behavior in the peak temperatures in GaN devices with increasing device power. It was desirable to quantify this effect in order to know how important it is to use temperature dependent thermal conductivities in continuum scale modeling of high power GaN HEMTs.

A study was performed on a Transmission Line Measurement (TLM) device model to determine the effect of temperature dependent conductivities on peak

temperatures and thermal stresses. TLMs are used to test electrical properties of the GaN HEMT structure through evaluation of current flow of the two dimensional electron gas. They are essentially AlGaN/GaN HFETs with no gate metallization. The TLM device had a 20μm source-drain spacing. For the comparison, constant conductivities of GaN and SiC at room temperature of 2.54 W/cm-K and 3.3 W/cm-K respectively were used. Figure 3.6 shows the difference in using constant conductivities vs. more accurate temperature dependent thermal conductivities for a wide range of device powers for the gateless TLM. For this study a constant heat generation region was applied to the area between the contacts of the TLM to simulate joule heating in this region.

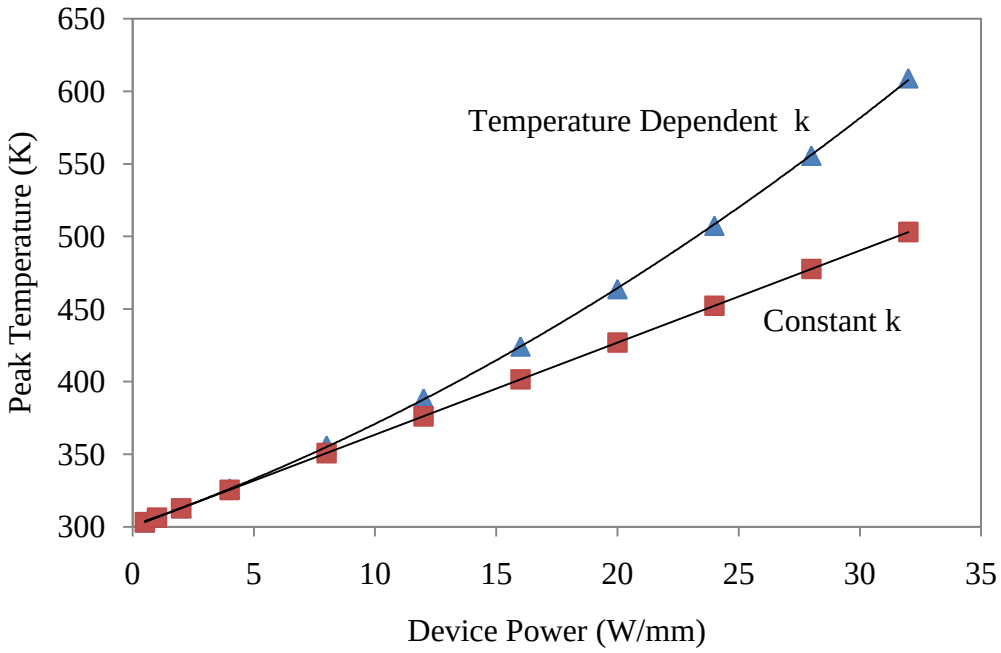

Figure 3.6. Peak temperature in a TLM as predicted by a constant conductivity and a temperature dependent conductivity are shown with squares and triangles respectively. Temperature dependent conductivities significantly affect peak temperature predictions at higher powers.

At higher powers the effect of using temperature dependent thermal conductivities is significant. The peak temperature error in using constant thermal conductivities will grow beyond 10% for powers above approximately 30 W/mm, and beyond 30% above about 50 W/mm. The peak temperature occurs at the center of the channel, approximately at the centroid of heat generation in the device. As expected, using

temperature dependent thermal conductivities also affects peak thermal stress in these devices.

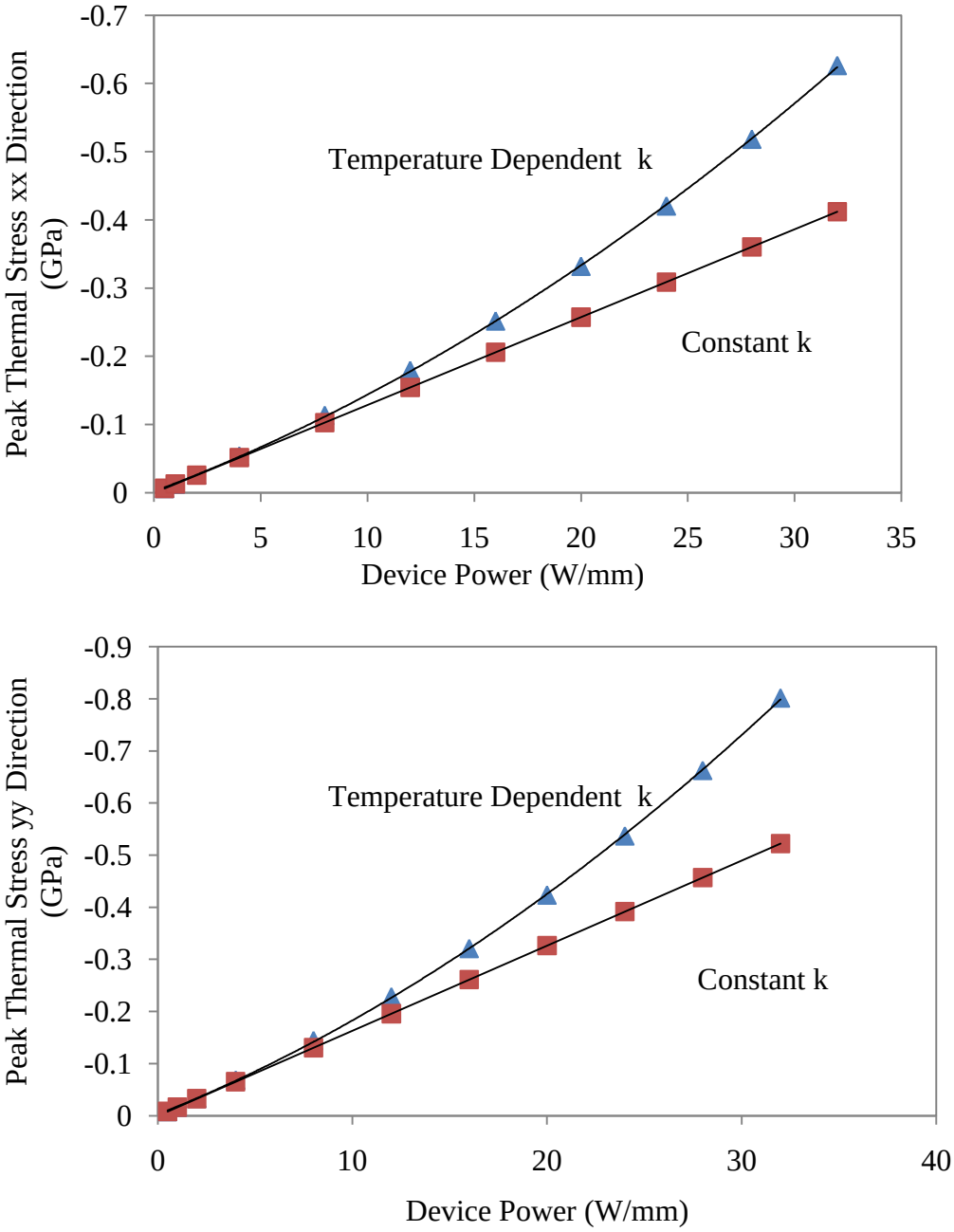

Figure 3.7. Peak thermal stress in a TLM as predicted by constant conductivities and temperature dependent conductivities are shown with squares and triangles respectively. Using temperature dependent conductivities significantly affects results at higher powers.

At the current level of development with AlGaN/GaN technology, power levels are usually low enough that constant temperature thermal conductivities can be used in

continuum scale modeling with minimal error, as has been done in literature for the most part. However, with some next generation devices showing power levels on the order of 40 W/mm, it will become important to use temperature dependent thermal conductivities in studying these devices. Furthermore, this model includes only a single finger device, which will be colder than a multiple finger device for the same total gate length and power density. It may be important to use temperature dependent thermal conductivities even with lower power devices if the peak temperatures are high.

3.7 Effect of Increased Device Bias

To begin to understand the effect of changing electrical conditions in the device, the first study to be done is to vary device drain to source bias while keeping everything else steady. The basic device outlined in Section 3.5 was run in COMSOL with drain to source voltages of between 2 and 14 V. The gate to source voltage was maintained at 0 V to keep the channel open. The heat generation information was taken from Sentaurus Device simulator and input into COMSOL and the resulting temperature profile and stress profiles were calculated. In a basic un-field-plated HFET, as expected, the peak temperature increases with increasing source drain bias. This is due to increased electron-phonon interactions as a result of higher electron velocities from the higher bias. When the bias on the device increases, the electric field across the channel increases which accelerates electrons. These faster moving electrons interact with the crystal lattice more frequently and impart more energy through Joule heating. This causes a temperature rise where the electrons are flowing the fastest, which in these devices is on the drain side of the gate. This temperature rise then causes thermal expansion which brings about thermal stresses. These stresses come from the thermal expansion coefficient mismatch between GaN and SiC and the thermal gradient. When material in the device is heated, it expands, unless it is constrained by surrounding material. In that case, it becomes compressively stressed. Figure 3.8 shows the increase in device

temperature and stress as a function of the increased device bias. Figure 3.9 shows the location of the peak x and y stresses in the device, near the drain side of the gate.

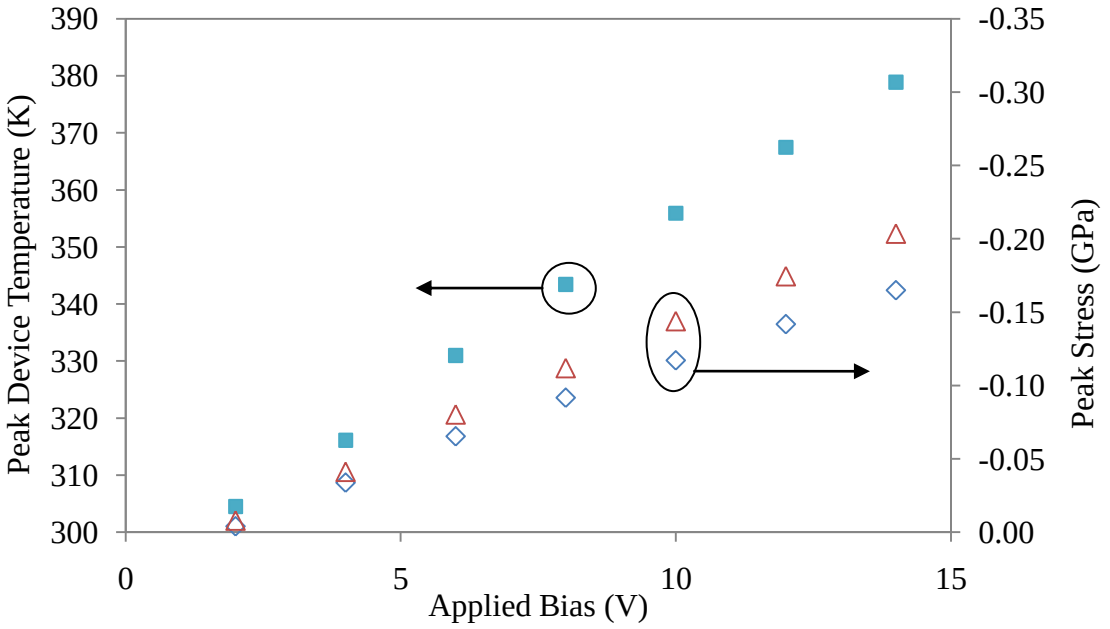

Figure 3.8. Peak device temperature and stresses in an HFET as a function of applied source-drain bias. Peak temperature is shown by solid squares, peak stress in the x and y direction are shown by hollow diamonds and triangles respectively. The peak temperatures and stresses are linear for low biases and device powers.

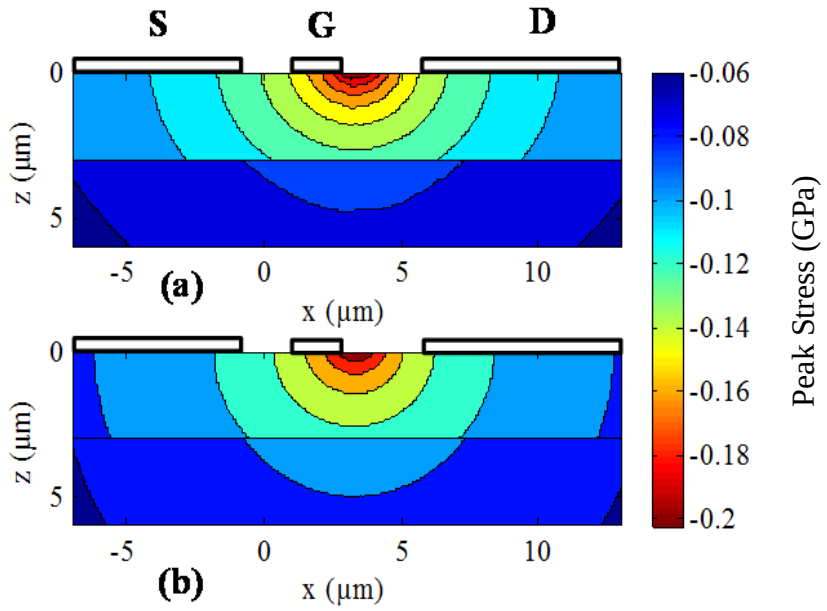

Figure 3.9. Spatial distribution of a) σ_x and b) σ_y, directional compressive stresses in GPa in HFET device at V_{DS}=14 V. Stresses are concentrated at the hottest section of the device, and are not biaxial.

3.8 Effect of Different Oxide Materials and Oxide Thicknesses

HFETs use Schottky contacts at the gate which utilize a difference in the work function of the metal contact and the AlGaN layer at the surface of the HEMT device to prevent current flow through the gate contact. However, high gate biases can induce a gate current which adversely affects device performance. Some groups have proposed using an oxide material between the gate metal and the AlGaN surface of the device forming a so called Metal Oxide Semiconductor Heterostructure Field Effect Transistor (MOSHFET) in order to further insulate the gate from the channel and prevent gate leakage currents[44, 71]. Several oxides and insulators such as, SiO_2, Al_2O_3, ZrO_2, Si_3N_4, and Sc_2O_3 have been used under the gate, giving rise to devices with lower gate leakage currents and hence increased output power[72-76]. This improvement comes at the cost of reduced gate transconductance, but it is assumed that the improvement in device performance is worth this cost. The goal of this study is to understand the effect that the oxide material and thickness has on the thermal performance of these devices. Al_2O_3, AlN, SiN, and SiO_2 with thicknesses between 10 and 20nm are modeled here to investigate device performance and specifically how changes in the electrical operation affect peak temperatures and stresses. Here the same procedure as in the previous section is followed again. The device is electrically simulated within Sentaurus and the heat generation region is applied to the thermal domain in COMSOL. The peak temperature for a number of different oxides and bias conditions are plotted in Figure 3.10. In addition, the peak stress for the x and y directions for a number of bias conditions and oxide materials are plotted in Figures 3.11 and 3.12 respectively. The stress in the z direction is not included because it is significantly lower than the x and y direction stresses. These peak stresses occur at the point of peak temperature, just on the gate side of the drain. The spatial distribution of stresses for an SiO_2 MOSHFET running at 14V drain source bias is shown in Figure 3.13.

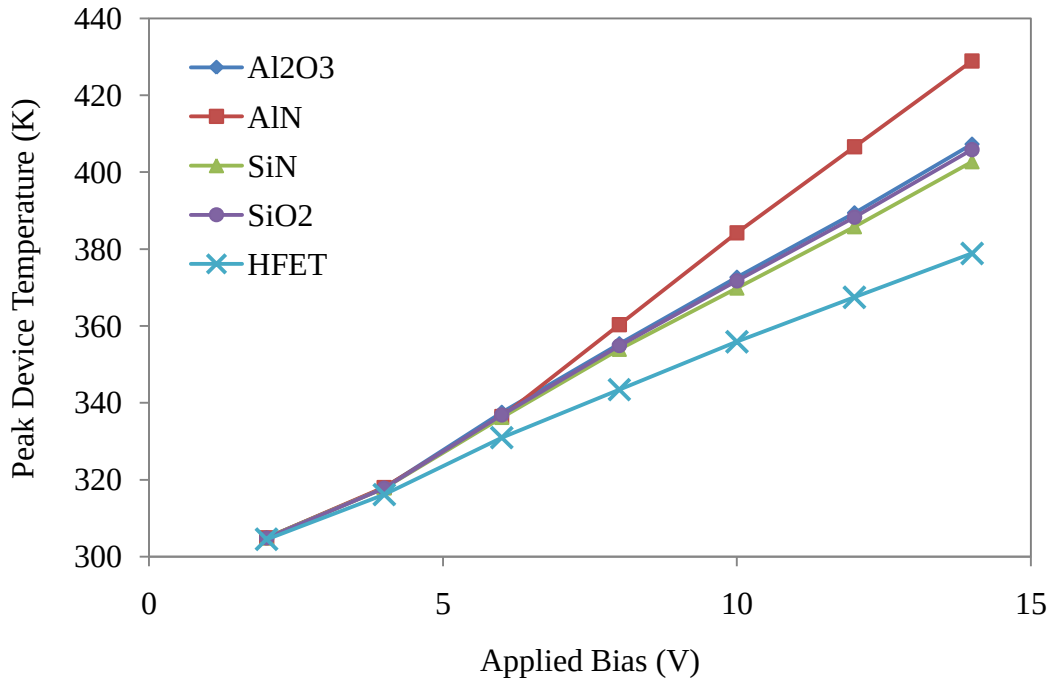

Figure 3.10. Peak temperatures of MOSHFETs with different oxide materials of thickness 10nm are shown for source-drain biases between 2 and 14 V. The HFET temperature is shown for reference.

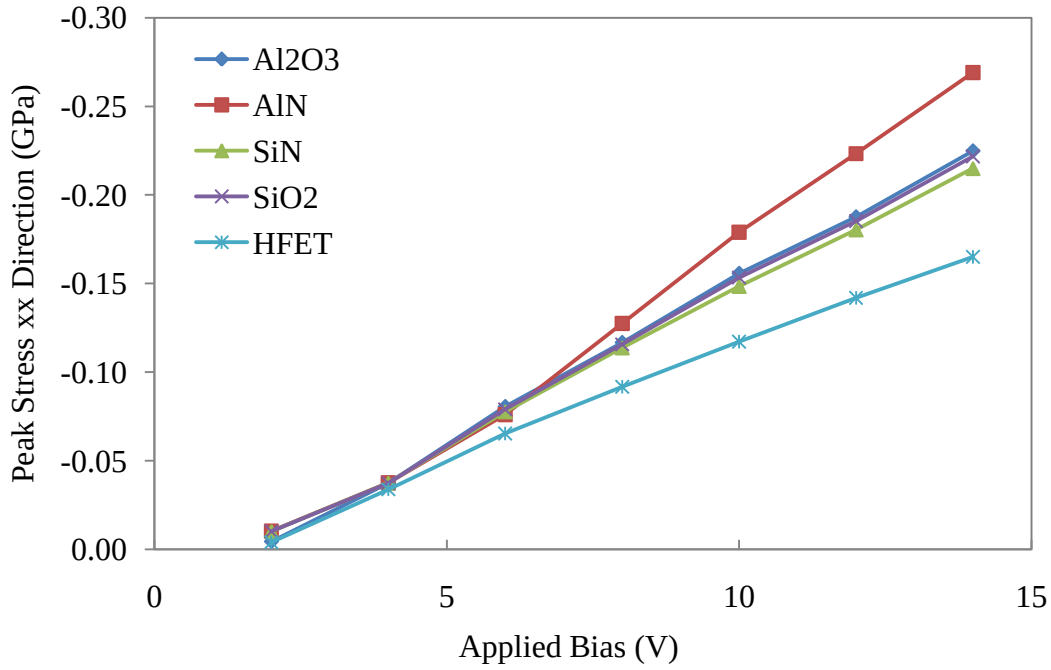

Figure 3.11. Peak stress in the x direction within AlGaN/GaN MOSHFETs. Again the HFET is included for reference.

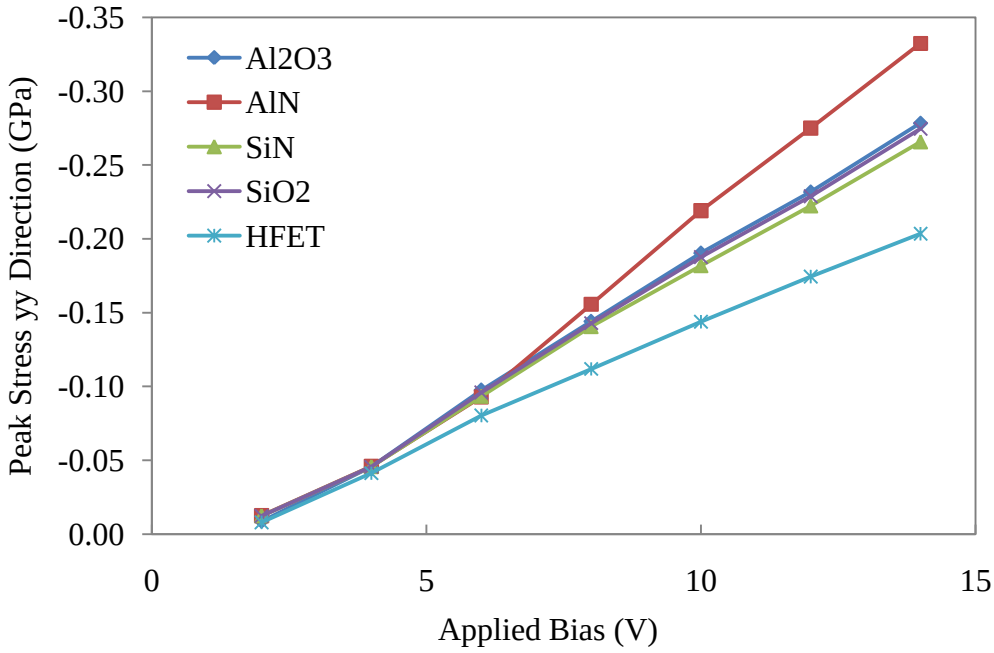

Figure 3.12. Peak stress in the y direction within AlGaN/GaN MOSHFETs. Again the HFET is included for reference. y direction stresses are significantly higher than x direction stresses.

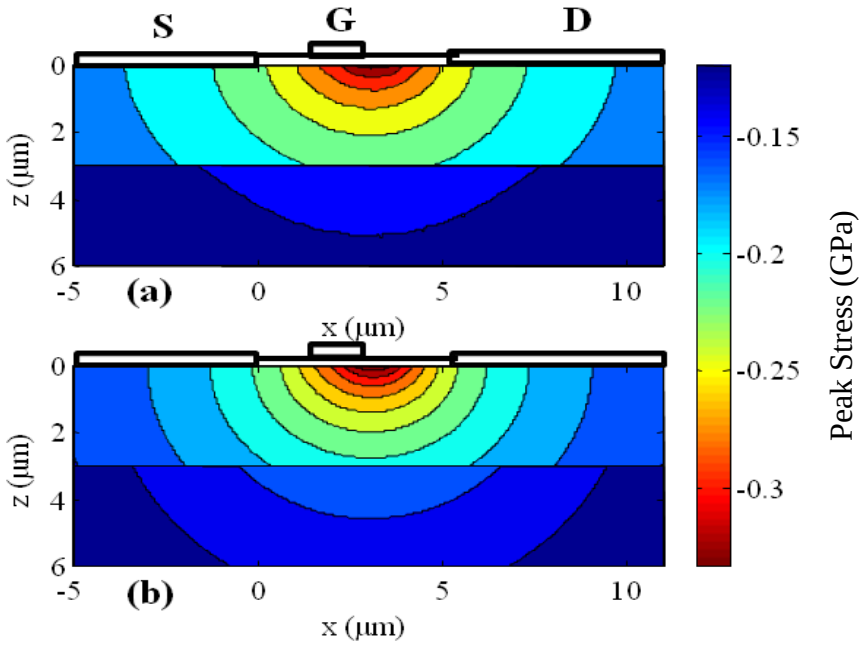

Figure 3.13. Spatial distribution of a) σ_x and b) σ_y, directional compressive stresses in GPa in an SiO_2 MOSHFET device at V_{DS}=14 V. Again, stresses are concentrated at the hottest point within the device.

In addition to investigating different Oxide Materials, the electrical, thermal, and mechanical effects of adjusting the thickness of these materials was studied. Figure 3.14

shows the peak temperatures in a MOSHFET as a function of increasing oxide thickness. Increasing oxide thickness appears to increase 2DEG electron density which in turn increases current flow in the devices leading to higher joule heating and higher temperatures as seen here. Changing the oxide material also changes the 2DEG density and mobility and leads to altered current flow under the same bias conditions[77]. Here the AlN is not plotted because there were convergence issues in the electrical model which prevented the results from being realistic, and so they were omitted. Peak stresses in the x and y directions are shown in Figures x and x. Again the shape of the curve is the same as for the temperature, indicating that stress is linearly related to temperature as is expected with a model with linear thermal expansion coefficients.

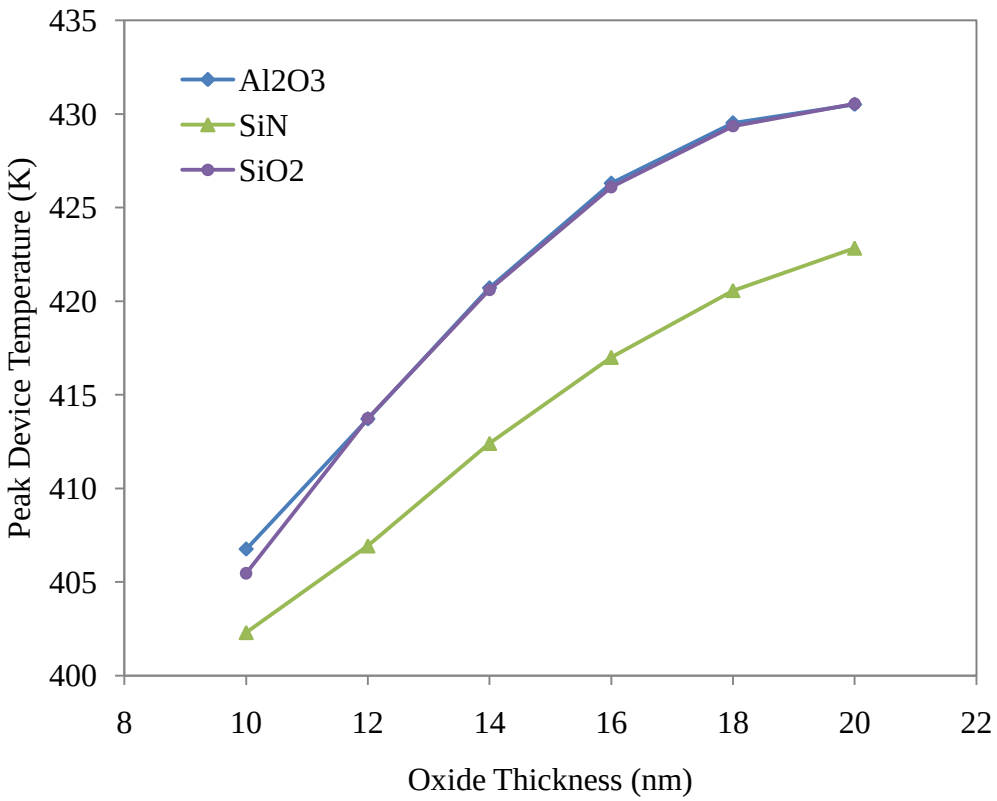

Figure 3.14. Peak temperatures of MOSHFETs with different oxide layer thicknesses for a source-drain bias 14V.

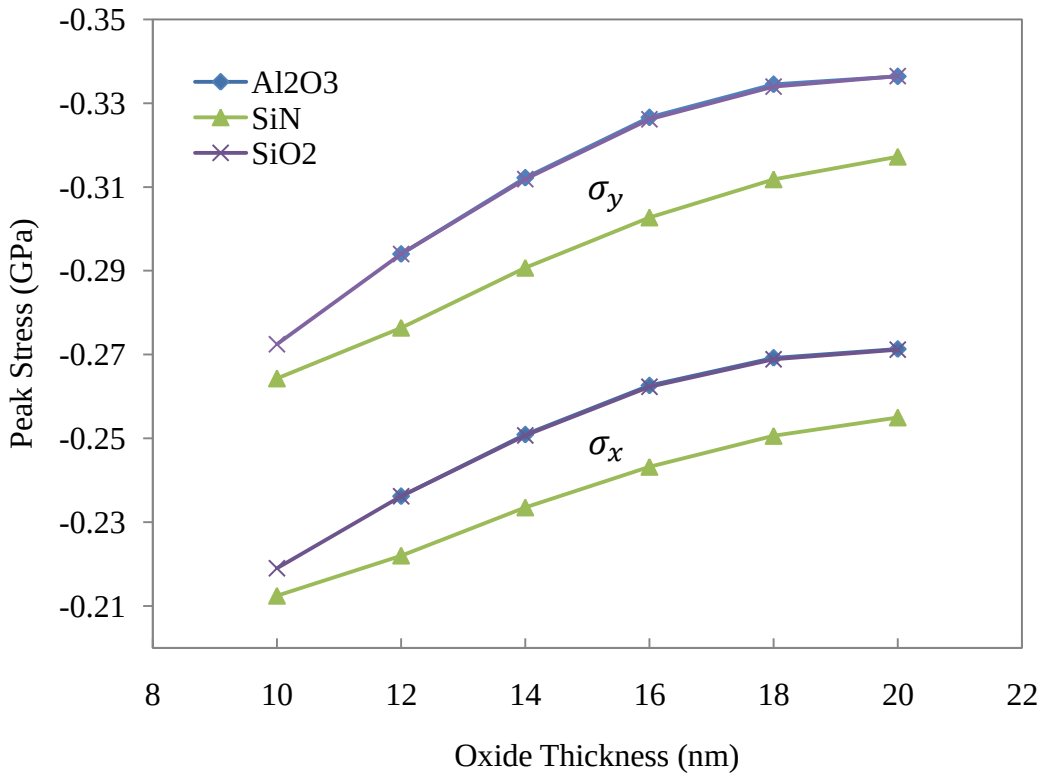

Figure 3.15. Peak stresses in AlGaN/GaN MOSHFETs with different oxide layer materials and thicknesses for source-drain bias 14V.

The thermal model did not actually include the oxide material, since its effect thermally is negligible as it is on the surface of the device, is extremely thin, and does not contribute significantly to heat transport away from the active region.

3.9 Effect of changing bias conditions under constant power

Since the electric field conditions in the device controls the Joule heating, it was postulated by Eric Heller of the Air Force Research Laboratory that changing the bias condition could impact device temperature (personal communication). For the same device power, the peak temperature of the device increased considerably for higher source-drain biases. To maintain the same device power with higher source-drain biases the gate must be pinched off considerably and this causes the heat generation region to shrink and become more concentrated near the drain edge of the gate. This effect has been modeled in this work. The electro-thermal coupling procedure used in this work

allows this effect to be captured. Without capturing the changing size of the heat generation region, the model would predict the same temperature for constant powers. Again the same procedure as in the previous sections was used: Sentaurus Device simulates the device electrically and determines the intensity of the heat generation within the device, which are then passed to COMSOL which calculates temperature and thermally induced stress. It is expected that the trend in temperature changes match closely with the temperature trend measured by Ramen spectroscopy on a device that was operated under the same conditions.

A slightly modified device geometry was used for this study to more closely match some experimental devices. The source-gate and gate-drain spacing were 9 and 8.5 µm respectively, while the gate was 3 µm in length. The AlGaN and GaN layers were 20nm and 2.05µm thick respectively. The device geometry is shown in Figure 3.16.

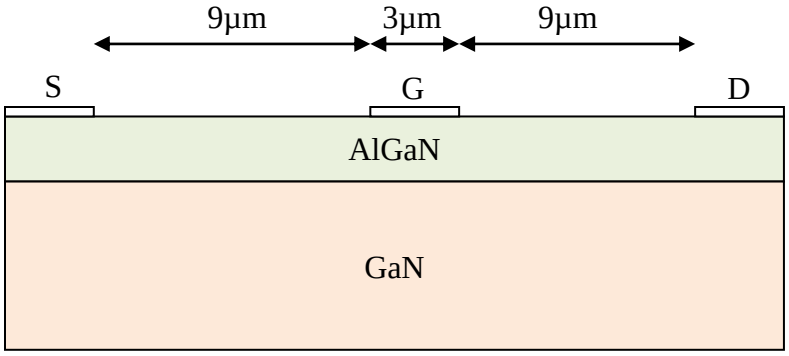

Figure 3.16. Active region device structure for the constant power device study. Contact separation is larger than in the device modeled in previous sections.

The effect of the changing bias conditions on the heat generation can be seen clearly in Figure 3.17, which shows the flattened heat generation throughout the AlGaN and GaN layers within the HEMT (here the 2 dimensional generation profile has been integrated in the y direction to yield a one dimensional generation profile that can then be applied uniformly on the top surface of the device).

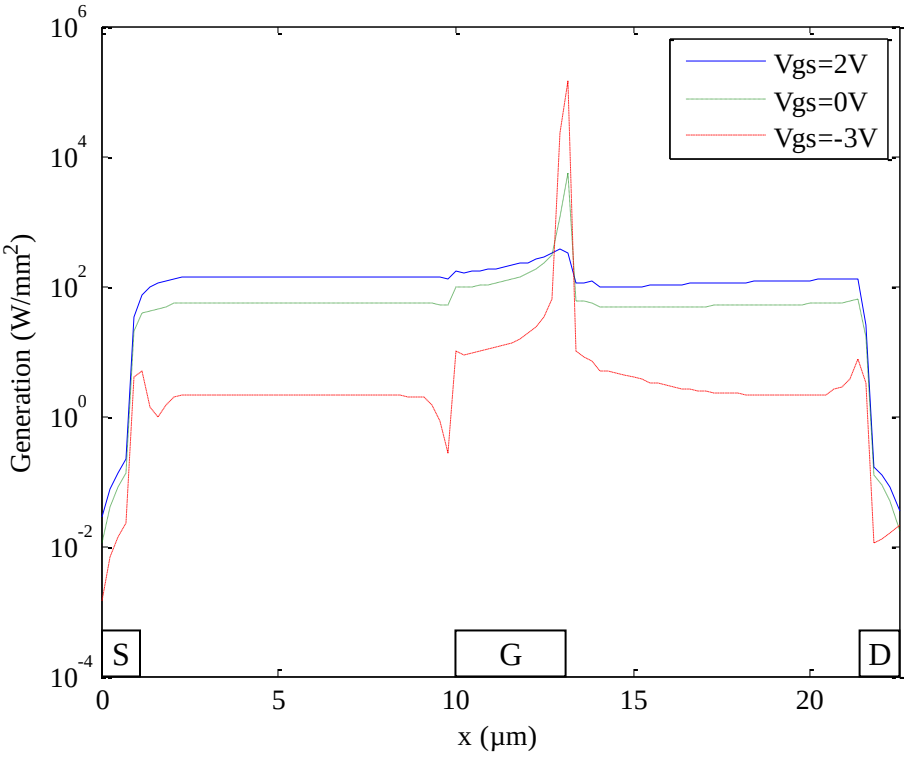

Figure 3.17. Heat generation shape as a function of gate bias. Note the increasing peak of heat generation near the gate edge.

Since the integrated power is not perfectly resolved when it is passed into COMSOL, it is necessary to normalize the generation so that all the different bias conditions are running at exactly the same power, in order to isolate the effect of the changing heat generation region shape. The power generation at any point is modified by the following formula

$$\hat{q}(\boldsymbol{r}) = q(\boldsymbol{r}) \frac{I_d V_{ds}}{\int_A q(\boldsymbol{r})\, d\boldsymbol{r}} \tag{3.5}$$

where $\hat{q}(\boldsymbol{r})$ is the adjusted power generation at a point $\boldsymbol{r}$, $q(\boldsymbol{r})$ is the raw generation at that point passed from Sentaurus, I_d is the drain current and V_{ds} is the drain-source voltage. This ensures that any numerical errors or meshing errors do not affect the peak device temperature as calculated using COMSOL.

Next, the generation profiles for several different bias conditions were run for the given device geometry, including the effect of a WCu package on the device. The peak temperature and the bias conditions are shown in Figure 3.18.

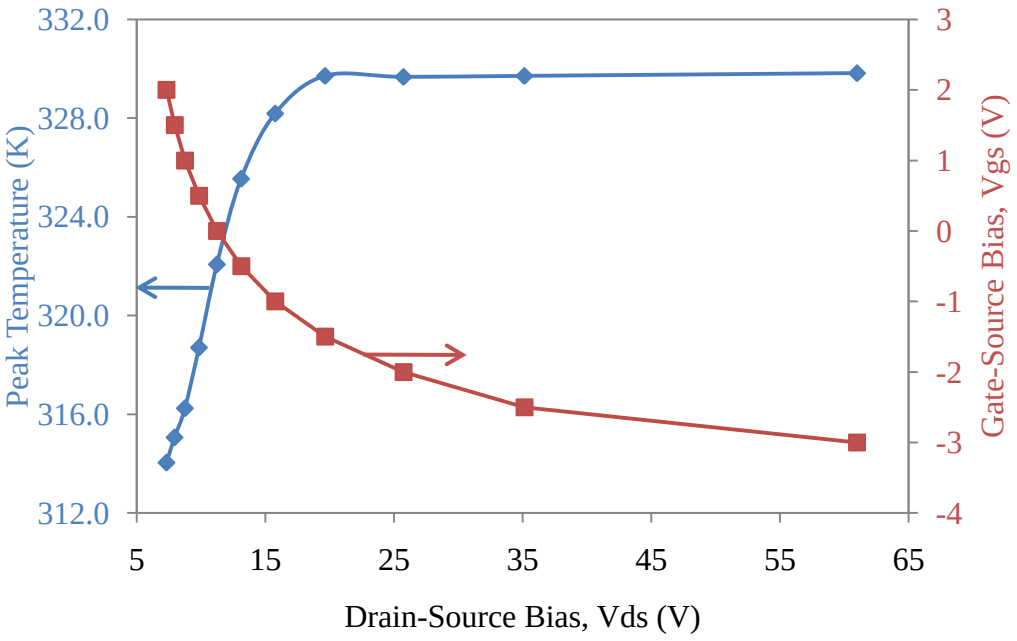

Figure 3.18. Peak temperature for different Vgs at a constant power of 3 W/mm.

It appears that after V_{ds} goes beyond 20 V or Vgs goes below -1.5 V the peak temperature flattens off. It should be noted that these induced temperature rises also lead to compressive thermal stresses due to the thermal gradient which increases linearly with peak temperature. Figure 3.19 shows the peak thermal stress within the device. The peak temperature increases by as much as 100% while the stress in the x and y directions increases by as much as 200% and 220% respectively.

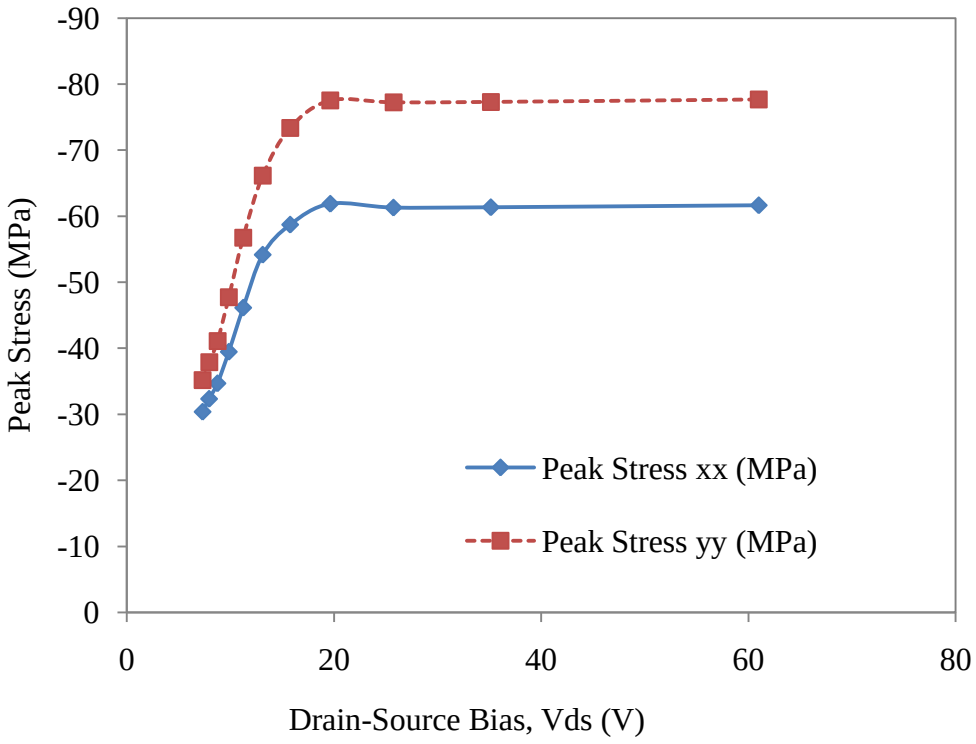

Figure 3.19. Peak Thermal Stress induced for different Vds at constant power of 3 W/mm

CHAPTER 4

MICRO-SCALE THERMAL MODELING

4.1 Introduction

While continuum-scale modeling has led to a fairly detailed understanding of device performance and a general grasp of device operating temperatures as well as their distributions, some aspects of device degradation are not understood well without accounting for micro-scale effects in thermal transport, specifically, out-of-equilibrium phonon generation, and ballistic effects in and around the hot spot. It is the goal of this chapter of the thesis is to determine as closely as possible the effect of inclusion of out-of-equilibrium phonons on the peak temperature in these devices.

This chapter is laid out as follows. First, micro-scale heat transport is introduced, along with methods of characterizing phonon information, and modeling heat transport at nano and micro-scales. Next, a solution method based on the Finite Volume Discrete Ordinates (FVDOM) method is developed using the scripting interface in COMSOL multiphysics. This method is verified against several analytical case studies. Finally, a detailed model that includes phonon interaction effects is developed, verified against several analytical solutions and then it is used to perform a preliminary simulation on an AlGaN/GaN HFET.

4.2 Micro-scale Heat Transport

Heat transport at the fundamental level occurs as the transmission of vibrations of atoms that make up the thermal medium. In a crystal (like GaN) where the atomic structure is well defined the atoms vibrate in predictable ways, and it is possible to develop a system of mathematics to rigorously calculate transmission of these vibrations. This science is called Lattice Dynamics and is concerned with the exploration of vibrational frequencies and amplitudes of atoms in crystal lattices. Due to quantum

mechanical effects these atomic crystal vibrations are quantized into particle like packets of energy called phonons.

It is important to understand how these phonon frequencies relate to their wavelengths in order to determine not only the speed of transmission of these phonons, but also the amount of energy they carry, and how they interact. The atoms within a crystal lattice are held in their positions with interatomic electrostatic forces and electron interactions. Atoms tend to reside near the equilibrium point where repulsive and attractive forces are equal to one another, but are able to oscillate around this equilibrium point. In order to quantify these vibrations, a harmonic model can be built where the atoms in a system move as if they were attached by springs to one another. By also assuming that the motion of these atoms is periodic according to some wavelength $\lambda = 2\pi/k$, where k is the wave vector of a certain vibrational mode, solving the harmonic system is a reasonable mathematical problem thanks to some tricks developed by Ewald for calculation of the lattice energy in a periodic configuration[78]. Given an accurate enough interatomic energy potential function, this simple model is accurate enough to determine a large set of properties.

In GaN there are two atoms in the unit cell, Gallium and Nitrogen, which are configured generally in the Wurtzite crystal structure when GaN is grown on Silicon Carbide (SiC) substrates.

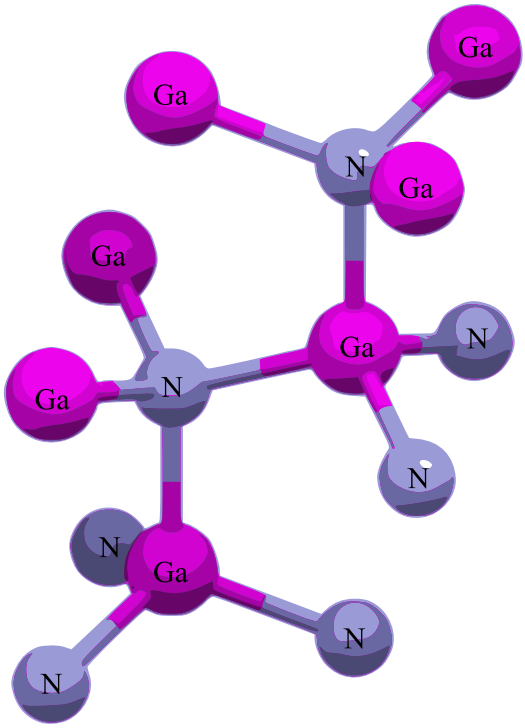

Figure 4.1. Gallium Nitride Wurtzite Atomic Structure

In the Wurtzite form, GaN has four atoms in its unit cell, which each have three degrees of freedom, leading to 12 individual polarizations of phonons, meaning that for any wave-vector, there are 12 different distinct energy levels that can be occupied by lattice vibrations in an ideal crystal. It is possible to calculate the energies or vibrational frequencies of all of these polarizations for every wavelength of phonon in the crystal using lattice dynamics. Figure 4.2 shows these energies as calculated by the *ab. initio* program ABINIT.

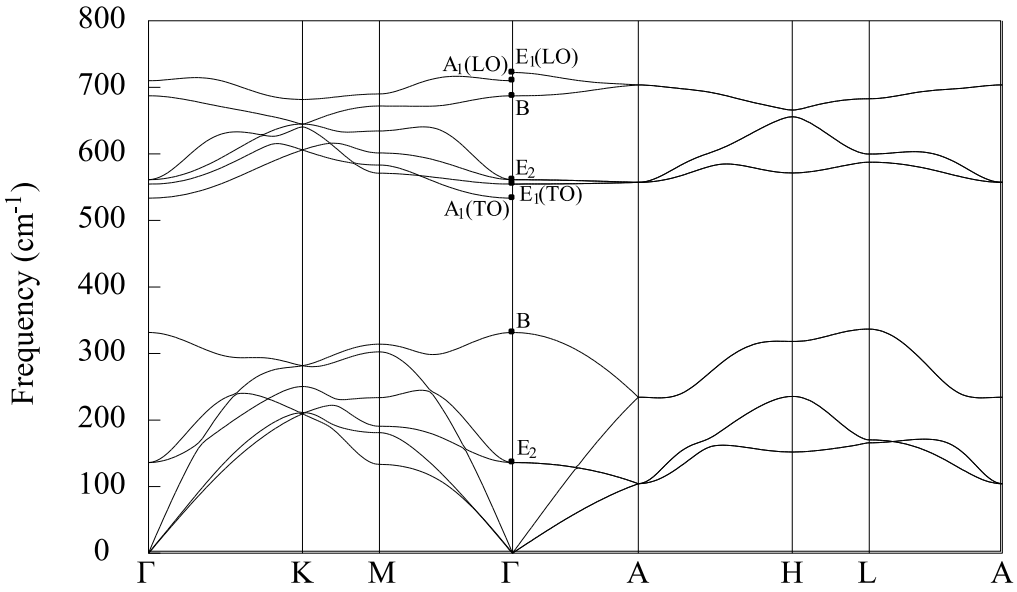

Figure 4.2. Phonon Dispersion curves for Gallium Nitride as calculated by ABINIT.

Because of the large discrepancy in mass of the two different types of atoms in GaN (Gallium and Nitrogen have a mass ratio of approximately 4.98) there develops a large phononic band-gap as can be seen in figure 4.2. This band-gap actually causes some problems for heat transport because it reduces the pathways for higher energy optical phonons to decay into acoustic phonons. Since higher energy modes are the phonons which predominantly interact with electrons, these are generated inside HFET devices[7]. However, they have low thermal conductivities and must decay into acoustic phonons before heat can be carried away from generation regions. This is referred to as a phonon bottleneck. This effect has not been quantified within the context of a GaN device, and one of the goals of this chapter is to determine how pronounced this effect is and whether it can contribute significantly to device reliability.

The Boltzmann Transport Equation (BTE) is known to effectively predict the motion of phonons within a crystal, and is the basis for much of the modeling of micro and nano-scale heat transport. The BTE for phonons is

$$\frac{\partial n_{qs}}{\partial t} + \boldsymbol{v}_{qs} \cdot \nabla n_{qs} = \left(\frac{\partial n_{qs}}{\partial t}\right)_{scattering} \tag{4.1}$$

where n_{qs} is a function of position, and time, and describes the distribution or density of phonons in the lattice and $\boldsymbol{v}$ is the phonon group velocity. The time dependent term on the right hand side represents the general scattering term which is difficult to determine accurately, while this is clearly the most important term. Without the scattering term the BTE cannot capture any of the important behavior of heat transport.

The BTE is a collection of Partial Differential Equations (PDEs) and is difficult to solve, mainly due to the as yet undefined scattering term. Some methods that are used are Monte Carlo Simulations, Lattice Boltzmann Method (LBM), and the Discrete Ordinates Method (DOM).

Monte Carlo Simulations work by statistically capturing only a small percentage of the actual interactions occurring within a material, but capturing enough of them to get a good idea of the device's behavior. This approach requires a large sample of phonons to be simulated, and is therefore computationally demanding.

The Lattice Boltzmann Method for phonons has been studied in detail in this group[2] and will not be discussed at length here. The basic premise is to discretize the domain using a finite difference technique and allow phonons to propagate only within this prescribed lattice. This method captures most basic phonon physics, but is not as rigorously derived as the DOM, and has issues with ray effects when dealing with complicated geometries or small domains. The LBM has its roots in rare gas dynamics and fluid dynamics, but is equally applicable to solution of heat transport in crystalline materials. The DOM, the method of choice in this work, originated in solution of radiation heat transfer problems, where ballistic effects are dominant. In the scale of devices being examined here, ballistic effects play an important role.

A commonly used approximation for the scattering term is the relaxation time approximation with elastic scattering where

$$\left(\frac{\partial n_{\boldsymbol{qs}}}{\partial t}\right)_{scattering} = \frac{n_{\boldsymbol{qs}}^0 - n_{\boldsymbol{qs}}}{\tau_{\boldsymbol{qs}}} \tag{4.2}$$

where $\tau_{\boldsymbol{qs}}$ is the relaxation time of a mode $\boldsymbol{qs}$ and $n_{\boldsymbol{qs}}^0$ is the equilibrium distribution of mode $\boldsymbol{qs}$, or the density of mode $\boldsymbol{qs}$ when all modes are in equilibrium. This simplification does not generally capture the inelastic effects within a material, and while it may preserve some aspects of the phonon dispersion, it will not capture true band-to-band transition rates in an out-of-equilibrium scenario.

A further simplification of this expression assumes that all modes in a particular direction can be represented by a single directional energy, so that the BTE becomes

$$\frac{\partial e}{\partial t} + \boldsymbol{v} \cdot \nabla e = \frac{e^0 - e}{\tau} \tag{4.3}$$

where now there is a single relaxation time τ. This is called the gray assumption and while this BTE preserves some ballistic effects it cannot capture the effect of dispersion within a material.

The COMSOL solution method presented below exclusively uses the DOM with a grey phonon approximation. It also uses the Equation of Phonon Radiative Transport (EPRT) introduced by Majumdar in [63] which is a form of the BTE with the occupation number $n_{\boldsymbol{qs}}$ replaced with a directional phonon intensity I which is a function of frequency, propagation direction, and position within they crystal. The EPRT is given by

$$\frac{1}{v}\frac{\partial I}{\partial t} + \mu\frac{\partial I}{\partial x} = \frac{I^0 - I}{v\tau} \tag{4.4}$$

where μ is the cosine of the angle between phonon propagation direction and the x direction.

4.3 Phonon Dispersion

In order to calculate how phonons interact with one another and how heat is transported within a material the properties of phononic vibration must be characterized. While some experimental methods exist for determining phonon frequencies, such as

neutron scattering, these methods are expensive, time consuming and limited in their resolution. It is possible however to predict phonon properties by simulating the motion of atoms in a crystal lattice via molecular dynamics or lattice dynamics calculations. The input for these calculations is an accurate representation of the potential energy of the atoms within a system with respect to their positions relative to one another. Some previous works in GaN have used a Tersoff potential with the General Utility Lattice Program (GULP)[2, 79]. This has the disadvantage of not being able to resolve high energy optical modes accurately because of the polarization of GaN. Outside of the GaN literature, works in Silicon typically use a simplified phonon dispersion along the [100] direction as all the information for an input into a phonon dispersion model[29], while a more recent work in Si has used the Environment-Dependent Interatomic Potential (EDIP) to accurately extract the thermal conductivity of Si with no fitting parameters[80]. This potential is available for Silicon only at this time, but it is capable of providing third order potential functions for calculating the anharmonic scattering strengths for three phonon interactions. This thesis uses the *ab. initio* program ABINIT for calculation of the phonon dispersion. At this time ABINIT does not seem capable of providing third order potential terms, although this feature is expected to become available shortly.

ABINIT is a first principles code for calculating total energy of molecules and periodic solids within Density Functional Theory (DFT), which gives rise to the ability to calculate a wide range of other parameters. It has been used effectively to determine structural, electrical, and vibrational properties of all kinds of materials, including III-V nitrides[81-83]. This work follows for all effective purposes the work of Pereira, *et al*[84] for parameters for GaN. Troullier-Martins LDA pseudopotentials were used following the lead of Stampfl and Van de Walle[85].

First the atomic structure must be optimized to ensure the atoms are at their equilibrium positions for calculation of the energy derivatives. This structural optimization is performed in two steps in ABINIT, firstly to determine the optimal atomic

coordinates given an initial lattice spacing, and then to optimize the lattice spacing as well as the atomic coordinates. This was performed with an initial guess of $a = 3.18$ Å and $c = 5.189$ Å with internal coordinate $u = 0.375$[86], using space group 186 (which corresponds to the Wrutzite GaN group C_{6v}^4-$P6_3mc$). A 4x4x4 Monkhorst-Pack grid[87] with a cutoff energy of 80 Hartree was used which resulted in optimized lattice parameters of $a = 3.184$ Å and $c = 5.183$ Å and internal coordinate $u = 0.377$.

After structural optimization, ABINIT can calculate the second derivatives of total energy for a given q-point grid. These q-points were generated by forming an evenly spaced grid over the Irreducible Brillouin Zone, including the Gamma point, and then using that grid as an input to the ABINIT derivative database analysis tool. ABINIT then gives the energies of the phonons with wave vectors at each q-point and polarization. Multiple different grid resolutions were investigated to ensure that properties of interest were grid independent. Figure 4.2 shows the phonon energy along the high symmetry directions in GaN, but phonon information is available for all points and polarizations within the Brillouin zone.

4.4 COMSOL Solution

COMSOL is a powerful general PDE solver which can be used to solve the BTE using the discrete ordinates method. This was first shown as viable in one dimension by Sihn and Roy[88]. COMSOL allows for quick deployment of a solution as the equation discretization and meshing are taken care of automatically. It is also reasonably powerful and can handle non-linearities in the equations in an easy and robust way. Unfortunately, this power comes at a cost in that imposition of boundary conditions and modification of solver settings is limited since the code is closed source. In addition, there is significant overhead in the way that COMSOL solves the equations and solving in sequential way is not possible. This section describes implementation of the grey BTE in COMSOL and shows some verification.

4.4.1 EPRT in COMSOL

The Equation of Phonon Radiative Transport, introduced in equation 4.4, must be discretized in the space and directional domains before it can be solved. This is done by introducing m independent variables which represent the phonon intensities in m discrete directions of phonon propagation as in [88] and [89]. Then the equilibrium phonon intensity equation above becomes a weighted sum, which can be expressed in COMSOL as an expression:

$$I_0 = \sum_{k=1}^{m} w_k I_k \tag{4.5}$$

where w_k are weights or solid angles for integration. These weights can be determined according to [90] such that if the angular domain is split uniformly in the azimuthal and polar directions the weights are

$$w_k = \frac{1}{4\pi}\int_{\varphi_k-\frac{\Delta\varphi}{2}}^{\varphi_k+\frac{\Delta\varphi}{2}}\int_{\theta_k-\frac{\Delta\theta}{2}}^{\theta_k+\frac{\Delta\theta}{2}} \sin\theta'\, d\theta'\, d\varphi' = \frac{1}{2\pi}\sin\theta_k \sin\frac{\Delta\theta}{2}\Delta\varphi \tag{4.6}$$

where φ and θ are shown in figure 4.3. This method of discretization is referred to as the Control Angle Discrete Ordinates Method (CADOM). More methods of ordinates selection and control angle discretization will be discussed in Section 4.5.2.

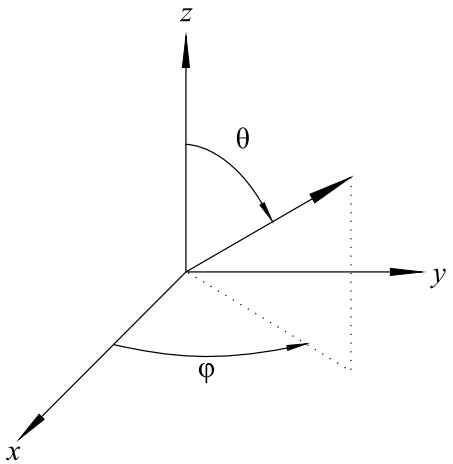

Figure 4.3. Polar and Azimuthal angles of discretization.

COMSOL requires the equations to be expressed in the 'weak' form[91] given by

$$0 = \int_{\Omega} W^{(2)} dA + \int_{B} W^{(1)} ds + \sum_{P} W^{(0)} - \int_{\Omega} v \cdot h^{(2)^T} \mu^{(2)} dA - \int_{B} v \cdot h^{(1)^T} \mu^{(1)} dA - \sum_{P} v \cdot h^{(0)^T} \mu^{(0)} \tag{4.7}$$

The sub domain contribution to the 'weak' form of the gray EPRT is

$$\int_{\Omega} \sum_{k=1}^{m} y_k (v\tau \overrightarrow{\boldsymbol{s}_k} \cdot \nabla I_k + I_k - I_0)\, dA = 0 \tag{4.8}$$

where y_k is the weight or test function for this variable (direction), and is the same as the shape function for the element, v is the group velocity, τ is the relaxation time, and I_k is the phonon intensity in the discrete k^{th} direction. Again, this simply states that the PDE must be orthogonal to the weight function, or shape function, over the domain of interest. The rest of the terms in the weak form depend on the boundary conditions.

4.4.2 Temperature Boundary Conditions

For a temperature boundary condition, one must relate the phonon intensity that is leaving the boundary to the temperature. The temperature can be related to the equilibrium phonon intensity as in [63, 89] by

$$I_0 = v \int_{T_{ref}}^{T} C dT \tag{4.9}$$

So the inward facing phonon intensities, where $\overrightarrow{\boldsymbol{s}_k} \cdot \vec{\boldsymbol{n}} < 0$, ($\vec{\boldsymbol{n}}$ is the outward facing normal) are given by

$$I_k = v \int_{T_{ref}}^{T} C dT \tag{4.10}$$

It is possible to write this as a boundary weak term as

$$\int_{B} \sum_{k=1}^{m} y_k (I_k - v \int_{T_{ref}}^{T} C dT)\, \alpha_k dA = 0 \tag{4.11}$$

where α_k, a function of position, is given by

$$\alpha_k = \alpha_k(s) = \begin{cases} 0, \text{ if } \overrightarrow{\boldsymbol{s_k}} \cdot \overrightarrow{\boldsymbol{n}(s)} \geq 0 \\ 1, \text{ if } \overrightarrow{\boldsymbol{s_k}} \cdot \overrightarrow{\boldsymbol{n}(s)} < 0 \end{cases} \tag{4.12}$$

and again y_k is the weight or test function equal to the shape function for each element. This constraint method requires that the inward phonon intensities reflect the exact temperature at the boundary, however, it doesn't guarantee that the temperature of the domain approaches the temperature of the boundary as the distance from the boundary approaches zero, since outward facing phonon intensities are not constrained. If the region near the boundary is in equilibrium, the temperature in that region should be the same as the boundary temperature, however if the region near the boundary is out of equilibrium, the boundary temperature might not exactly match the constraint temperature due to temperature slip at the boundaries.

4.4.3 Heat Flux Boundaries

Heat flux boundaries are similar to temperature boundaries. Again, using expressions from [88] and [63] to write the heat flux one obtains

$$\overrightarrow{\boldsymbol{Q}} = 4\pi \sum_{k=1}^{m} w_k I_k \overrightarrow{\boldsymbol{s_k}} \tag{4.13}$$

where $\overrightarrow{\boldsymbol{Q}}$ is the vector heat flux.

To implement this in COMSOL one must relate each inward facing phonon intensity to the boundary heat flux. First the Boltzmann equation is substituted into the heat flux expression.

$$\overrightarrow{\boldsymbol{Q}} = 4\pi \sum_{k=1}^{m} w_k (I_0 - v\tau \overrightarrow{\boldsymbol{s_k}} \cdot \nabla I_k) \overrightarrow{\boldsymbol{s_k}} \tag{4.14}$$

Expanding this expression and splitting the summation one gets

$$\overrightarrow{\boldsymbol{Q}} = 4\pi \sum_{k=1}^{m} w_k I_0 \overrightarrow{\boldsymbol{s_k}} - 4\pi \sum_{k=1}^{m} w_k v\tau \overrightarrow{\boldsymbol{s_k}} \cdot \nabla I_k \overrightarrow{\boldsymbol{s_k}} \tag{4.15}$$

Here the first term cancels since the net flux of the equilibrium intensity must be zero since it is the same in all directions. Now using Einstein notation for convenience, the flux is

$$Q_i = -4\pi \sum_{k=1}^{m} w_k v\tau s_{k\gamma} \partial_\gamma I_k s_{ki} \tag{4.16}$$

where indices that appear twice in a term are summed, and s_{ki} represents the kth direction cosine in the ith dimension. Since the s_k terms are orthogonal over the solid angle (i.e. $\sum_{k=1}^{m} s_{ki} s_{k\gamma} = 0$ when $i \neq j$), all the non-diagonal terms are zero, and one can write

$$Q_i = -4\pi \sum_{k=1}^{m} w_k v\tau \partial_i I_k s_{ki}^2 \tag{4.17}$$

Finally, assuming that the directional intensity gradients are all approximately equal to the equilibrium intensity gradient, it is possible to solve for $v\tau\partial_i I_0$ to get

$$v\tau \partial_i I_0 = -\frac{Q_i}{4\pi \sum_{k=1}^{m} w_k s_{ki}^2} \tag{4.18}$$

Now, using the Boltzmann equation and assuming the solution is diffuse, the gradient of the directional intensities can be adequately represented by the gradient of the equilibrium intensity,

$$I_k - I_0 - \sum_i \frac{Q_i s_{ki}}{4\pi \sum_{l=1}^{m} w_l s_{li}^2} = 0 \tag{4.19}$$

Now considering that at the boundary an inward facing heat flux Q_{flux} is specified,

$$\vec{\boldsymbol{Q}} = -Q_{flux} \vec{\boldsymbol{n}} \tag{4.20}$$

the final boundary constraint expression is

$$I_k - I_0 + \sum_i \frac{Q_{flux} s_{ki} n_i}{4\pi \sum_{l=1}^{m} w_l s_{li}^2} = 0 \tag{4.21}$$

where again the dimension index i is summed in the last term. This can be written in weak form as

$$\int_B \sum_{k=1}^{m} y_k \left(I_k - I_0 + \sum_{i=x,y,z} \frac{Q_{flux} s_{ki} n_i}{4\pi \sum_{l=1}^{m} w_l s_{li}^2} \right) \alpha_k \, dA = 0 \tag{4.22}$$

where α_k is defined as above.

4.4.4 Reflection Boundaries

Both specularly reflecting and diffusely reflecting boundaries may be desired. It is possible to constrain these using a weak term, although a point-wise constraint implementation is also possible.

For diffuse reflection this reduces to the simple constraint that

$$I_k = I_0 \tag{4.23}$$

for inward facing phonon directions, such that $\vec{s_k} \cdot \vec{n} < 0$. This constraint can be written as one weak term

$$\int_B \sum_{k=1}^{m} y_k (I_k - I_0) \, \alpha_k dA = 0 \tag{4.24}$$

where α_k is defined as above.

Specular reflection is slightly more complicated in that

$$I_k = I_r \tag{4.25}$$

is enforced, where I_r is the intensity of the source of the reflection.

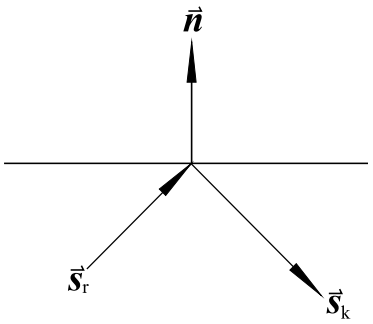

Figure 4.4. Reflection boundary condition geometry

In general the reflection direction $\overrightarrow{\boldsymbol{s_r}}$ of I_r is related to the direction $\overrightarrow{\boldsymbol{s_k}}$ of I_k by [89].

$$\overrightarrow{\boldsymbol{s_r}} = \overrightarrow{\boldsymbol{s_k}} - 2(\overrightarrow{\boldsymbol{s_k}} \cdot \vec{\boldsymbol{n}})\vec{\boldsymbol{n}} \tag{4.26}$$

When the directions are discretized there is no guarantee that there will be a phonon intensity direction that exactly matches the reflection condition, so it may be necessary to interpolate between multiple directions, or select the closest direction.

The current implementation includes specular reflection that finds the closest reflection vector. The weak term for specular reflection is

$$\int_{\mathrm{B}} \sum_{k=1}^{m} y_k (I_k - I_r)\, \alpha_k dA = 0 \tag{4.27}$$

4.4.5 Coupling to a Fourier Domain

Modeling an entire device with this COMSOL DOM method would be extremely costly, considering the dimensions of a AlGaN/GaN HFET. In order to reduce the computational requirements, it is possible to simulate the non-equilibrium or non-ballistic portion of the device with the Fourier method while maintaining the accuracy introduced with this method near the region of the hot spot. This requires coupling two regions with different governing physics. Coupling the DOM solution domain to a Fourier domain is not difficult using this COMSOL method, and this is the first implementation of a multi-scale DOM to Fourier coupling published, to the knowledge of the author. The Fourier weak form for implementation in COMSOL is

$$\int_{\Omega} y_f \nabla(k\nabla T) d\mathrm{A} = -\int_{\Omega} \nabla y_f \cdot (k\nabla T) d\mathrm{A} + \int_{B} y_f (k\nabla T \cdot \vec{\boldsymbol{n}}) dA \tag{4.28}$$

where k is the thermal conductivity, T is the temperature, y_f is the shape function for the Fourier domain, Ω is the Fourier domain, $d\Omega$ is the boundary of the Fourier domain, and $\vec{\boldsymbol{n}}$ is the outward facing normal vector at the boundary of the Fourier domain. Where the

boundary of the Fourier domain coincides with the boundary of the discrete ordinates domain, one must couple the two with a weak boundary expression. This expression ensures that the flux from the discrete ordinates domain is matched by a boundary flux into the Fourier domain, and also that the distribution of discrete ordinates intensities are appropriate. This distribution is derived by using the Boltzmann equation and assuming local equilibrium. Take the discrete Boltzmann equation in a particular direction k

$$I_k - I_0 + v\tau\overrightarrow{\boldsymbol{s}_k} \cdot \nabla I_k = 0 \tag{4.29}$$

Now assume that $\nabla I_k = \nabla I_0$, and substitute $I_0 = v\int_{T_{ref}}^{T} CdT$, which yields

$$I_k - v\int_{T_{ref}}^{T} CdT + v^2\tau[C]_{T_{ref}}^{T}\overrightarrow{\boldsymbol{s}_k} \cdot \nabla T = 0 \tag{4.30}$$

One must write the boundary expression so that it ensures that this expression is true for all inward facing phonon intensities; the outward facing phonon intensities are not prescribed, but do play a role in coupling the total flux to the Fourier flux. The flux on both sides of the boundary must be equal, so that

$$4\pi\sum_{k=1}^{m} w_k I_k \overrightarrow{\boldsymbol{s}_k} \cdot \vec{\boldsymbol{n}} = k\nabla T \cdot \vec{\boldsymbol{n}} \tag{4.31}$$

Using these expressions one can write the weak form for a Fourier-DOM boundary as

$$\int_B y_f\left(4\pi\sum_{k=1}^{m} w_k I_k \overrightarrow{\boldsymbol{s}_k} \cdot \vec{\boldsymbol{n}}\right) + \sum_{k=1}^{m} y_k\left(I_k - v\int_{T_{ref}}^{T} CdT + v^2\tau[C]_{T_{ref}}^{T}\overrightarrow{\boldsymbol{s}_k} \cdot \nabla T\right)\alpha_k\, dA \tag{4.32}$$

The first term in the integral is the contribution of the flux from the DOM going into the Fourier domain, and the last terms are the contribution of the flux going from the Fourier domain into the DOM domain.

4.4.6 Verification

Several different test cases are presented in order to demonstrate the viability of this method for solving the BTE in two dimensions. The test cases are two cases that were considered by Christensen in his thesis[2], a square domain with one non-homogenous boundary condition (a heat flux and a prescribed temperature boundary condition are considered) and a long rectangular domain with specular or diffusely reflecting edges to test the effect of reflection boundaries. The square domain is compared to analytical solutions in the diffuse case, and the reflection boundary condition results are compared with analytical solutions from [92]. The test cases were all run for GaN with one group velocity of $3427\ \mathrm{m/s}$, one relaxation time of $24.836\ \mathrm{ps}$ and a volumetric specific heat of $2.5707\times10^{6}\ \mathrm{J/K\,m^3}$. The thermal conductivity for use in the analytical calculation was calculated from kinetic theory to be

$$k = \frac{1}{3}\tau v^2 C \tag{4.33}$$

where C is the volumetric specific heat. The linear system solver SPOOLES and 3rd order Lagrange elements were used with all other settings left default.

4.4.6.1 Diffuse Analytical Comparison

The square domain tested is large enough to be fully diffuse so one can verify the results against an analytical solution of Fourier's heat transfer equation. Solve

$$\Delta T = 0 \tag{4.34}$$

$$T = T_0 \text{ at } x = 0, \text{ and } x = a \tag{4.35}$$

$$T = T_0 \text{ at } y = 0 \tag{4.36}$$

and for a prescribed temperature condition one obtains

$$T = T_1 \text{ at } y = a \tag{4.37}$$

while for a prescribed heat flux condition one obtains

$$\frac{\partial T}{\partial y} = -\frac{Q_b}{k} \text{ at } y = a \tag{4.38}$$

The domain was assumed to have a Knudsen number of 0.01 making it fully diffuse ($a = 8.5113$ μm). The domain was split into 924 freely meshed triangular elements as shown in figure 4.5.

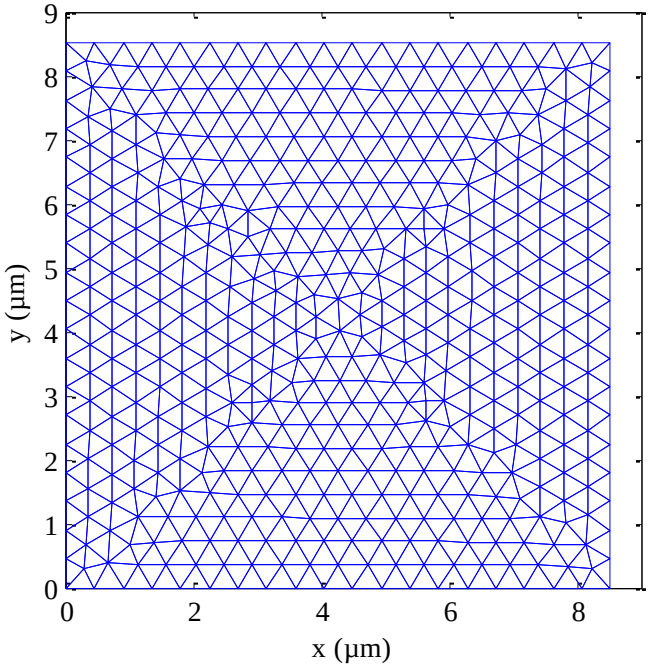

Figure 4.5. Meshing for square domain case

4.4.6.2 Temperature Boundary Condition

The analytical Fourier solution temperature distribution is given by

$$T = 2\frac{(T_1 - T_0)}{\pi}\sum_{m=1}^{\infty}\frac{1}{m}(1 - \cos m\pi)\,\text{csch}(m\pi)\sin\left(\frac{m\pi x}{a}\right)\sinh\left(\frac{m\pi y}{a}\right) + T_0 \tag{4.39}$$

The boundary temperatures T_0 and T_1 were set to 300 and 301 K respectively. The COMSOL solution matches the analytical solution as was visually verified by plotting the temperature distribution along a vertical cross section with $x = {}^{a}/_{2}$ and along the top of the domain at $y = a$ as shown in figures 4.6 and 4.7. The 100 term

analytical solution at the top boundary oscillates considerably due to the discontinuities at the corners.

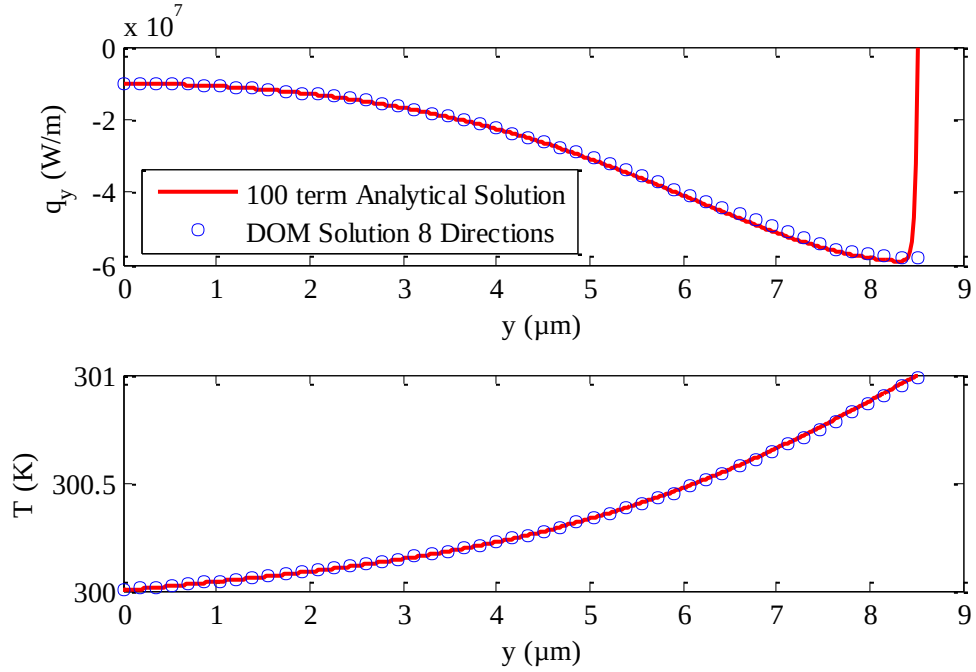

Figure 4.6. Heat flux and Temperature distributions for a vertical cross section of the domain at $x = {}^{a}/_{2}$ for the COMSOL solution and the analytical solution.

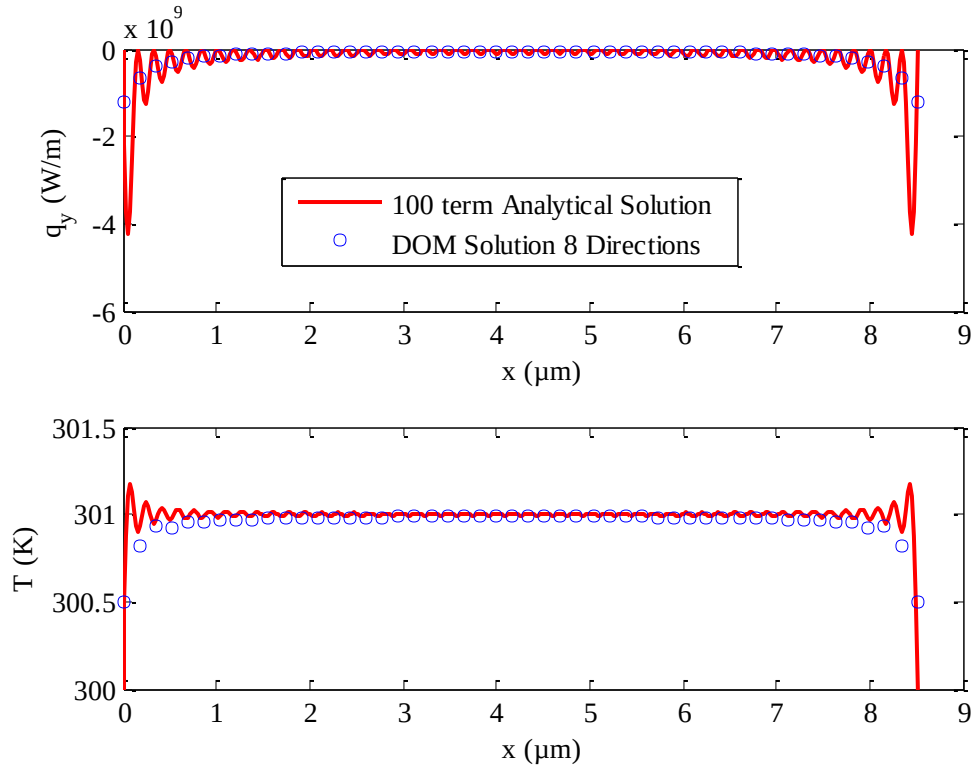

Figure 4.7. Heat flux and Temperature at the boundary of the domain where $y = a$ for the COMSOL solution and the analytical solution. The analytical solution is oscillatory at the boundary due to discontinuities at the corners.

4.4.6.3 Heat Flux Boundary Condition

The analytical Fourier solution gives a temperature of

$$T = 2\frac{Q_b}{k}\frac{a}{\pi^2}\sum_{m=1}^{\infty}\frac{1}{m^2}(1-\cos m\pi)\,\mathrm{sech}(m\pi)\sin\left(\frac{m\pi x}{a}\right)\sinh\left(\frac{m\pi y}{a}\right)+T_0 \qquad (4.40)$$

The comparison between the COMSOL solution and the analytical solution is shown for a vertical cross section of the domain in figure 4.8, and for the top of the domain in figure 4.9. The heat flux prescribed at the top of the domain was 1000 W, which comes to 2.439×10^8 W/m divided along the length of the boundary.

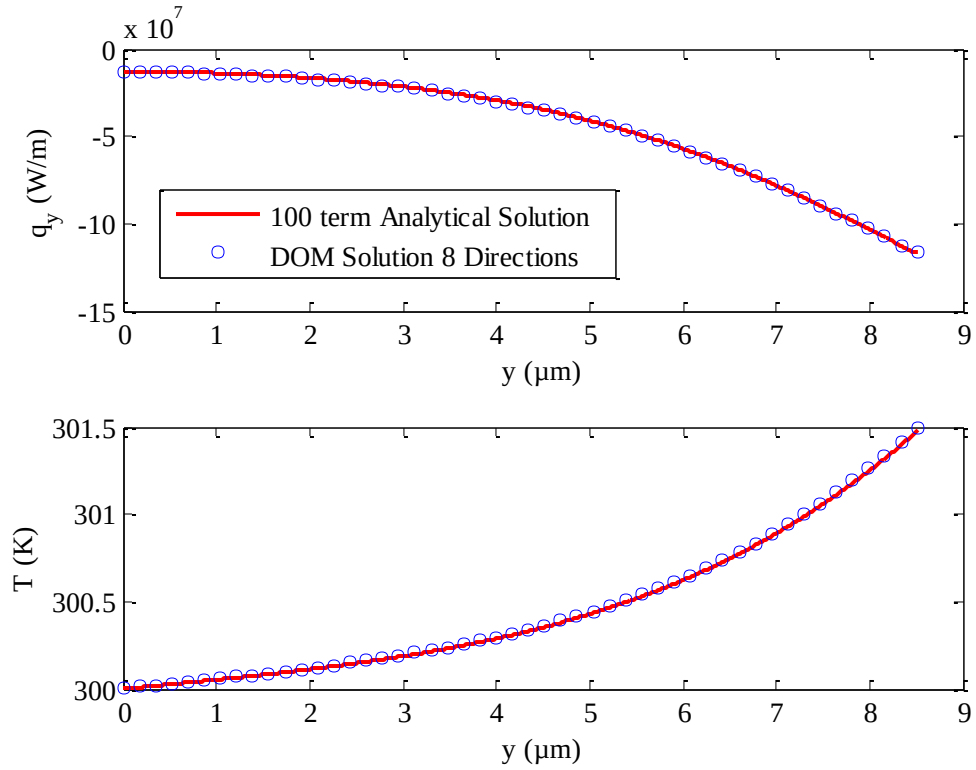

Figure 4.8. Heat flux and Temperature distributions for a vertical cross section of the domain at $x = {}^{a}/_{2}$ for the COMSOL solution and the analytical solution.

In both a heat flux and a temperature non-homogenous boundary condition the COMSOL DOM solution matches the analytical solution aside from oscillations due to discontinuities in the boundary conditions.

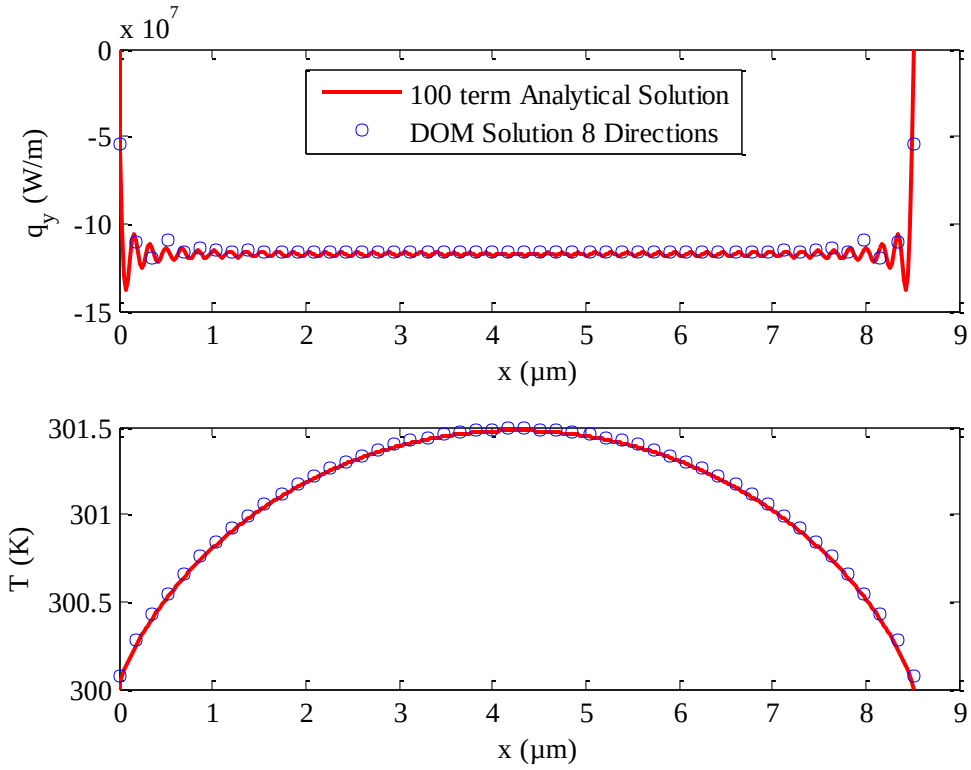

Figure 4.9. Heat flux and Temperature distributions for $y = a$ at the top of the domain for COMSOL and the analytical solution. Again the analytical heat flux solution is oscillatory due to a heat flux discontinuity at the corners.

4.4.6.4 Reflective Boundary Condition Comparison

In order to test the accuracy of the implementation of reflective boundary conditions, the COMSOL Discrete Ordinates solution is compared with Flik and Tien's solution. The domain is a long rectangle with the top and bottom boundaries set to diffuse or specular reflection. The effective thermal conductivity is compared with the theoretical thermal conductivity size effect as given by [92]. The specular boundary condition should not reduce the conductivity at all, while the diffuse case should reduce the conductivity according to

$$\frac{k_x}{k_\infty} = \begin{cases} 1 - \frac{4\mathrm{Kn}}{3\pi}(1 - S^3) - \frac{2}{\pi}\mathrm{ArcCos}\left(\frac{1}{2\mathrm{Kn}}\right) + \frac{1}{\mathrm{Kn}\pi}\left(1 + e^{-\frac{m}{2\mathrm{Kn}}}\right)\ln\left(\frac{2\mathrm{Kn} + 1 + 2\mathrm{Kn}S^2}{2\mathrm{Kn} + 1 - 2\mathrm{Kn}S^2}\right) & \text{for } \mathrm{Kn} \geq \frac{1}{2} \\ 1 - \frac{4\mathrm{Kn}}{3\pi} & \text{for } \mathrm{Kn} < \frac{1}{2} \end{cases} \quad (4.41)$$

where $S = \sqrt{1-(2\mathrm{Kn})^{-2}}$ and as always $\mathrm{Kn} = \Lambda/\mathrm{d}$ where Λ is the phonon mean free path, d is a characteristic length, in this case, the film thickness, and m is a fitting parameter (Flik and Tien choose $m = 6$).

The effective conductivity for the COMSOL calculations was calculated by selecting a sub-region of the rectangular domain well away from the side walls to remove any effect of the edges and using the following equation.

$$k_x = -q_x \frac{\Delta x}{\Delta T} \tag{4.42}$$

Figure 4.10 shows the reduction in thermal conductivity as the domain is made thinner.

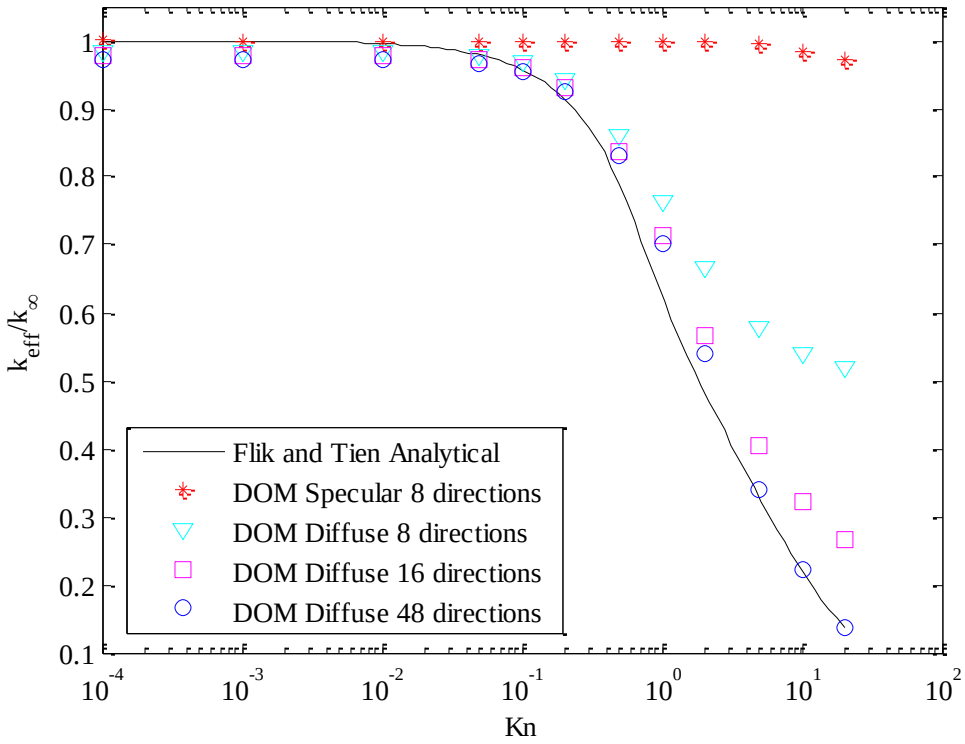

Figure 4.10. Comparison of COMSOL Discrete Ordinates solution for thin film reduction of conductivity vs. Flik and Tien's analytical result [92]. Two diffuse cases are shown, D8 representing 8 directions and D32 representing 32 directions.

As expected the thermal conductivity for the specularly reflecting case is not at all reduced, while the diffuse case roughly follows Flik and Tien. Increasing the number of directions correlates to a closer fit to the analytical result. The thermal conductivity

decreases for the diffuse case because the diffusion reflection condition reflects some of the phonons backwards. The thinner the domain, the more phonons make contact with the wall and are reflected back.

4.5 Finite Volume Solution

Although COMSOL has its advantages, the overhead in running such a large number of degrees of freedom is too large to allow further study to be done, specifically on band to band transition processes in Gallium Nitride. Therefore, it is necessary to develop an in-house solution and write the code from scratch. The method selected for solution of the BTE is the Finite Volume Method with the Discrete Ordinates Method (DOM) for discretizing in the directional domain. Within this model, an attempt has been made to maintain as much accuracy as possible about phonon transition probabilities in order to capture the pathway of energy from electrons to slower moving optical phonon modes to fast moving acoustic phonon modes.

4.5.1 Model Derivation

4.5.1.1 Cannonical Boltzmann Equation

The development of the finite volume model which includes polarization and dispersion begins with one of the most fundamental forms of the Boltzmann Equation, which describes the evolution of individual phonon modes within a crystal lattice. This equation can be linearized and in steady state is referred to as the canonical form of the Boltzmann equation. Srivastava[93] gives the canonical form of the linearized Boltzmann equation as

$$-\boldsymbol{v}_{\boldsymbol{qs}} \cdot \nabla n_{\boldsymbol{qs}} = \sum_{\boldsymbol{q's'}} P^{ss'}_{\boldsymbol{qq'}} \psi_{\boldsymbol{q's'}} \tag{4.43}$$

where $\boldsymbol{v_{qs}}$ is the mode group velocity, n_{qs} is the occupation number of the mode, and the scattering matrix on the right hand side given by

$$P_{\boldsymbol{qq'}}^{ss'} = \sum_{q''s''} \sum_{q'''s'''} \left[\frac{1}{2} \delta_{ss'} \delta_{\boldsymbol{qq'}} \left(\bar{P}_{\boldsymbol{qs},\boldsymbol{q'''s'''}}^{\boldsymbol{q''s''}} + \bar{P}_{\boldsymbol{qs}}^{\boldsymbol{q'''s'''},q''s''} + \bar{P}_{\boldsymbol{qs},\boldsymbol{q''s''}}^{\boldsymbol{q'''s'''}} \right) - \delta_{s's'''} \delta_{\boldsymbol{q'q'''}} \left(\bar{P}_{\boldsymbol{qs},\boldsymbol{q''s''}}^{\boldsymbol{q'''s'''}} + \bar{P}_{\boldsymbol{q''s''},\boldsymbol{q'''s'''}}^{\boldsymbol{qs}} - \bar{P}_{\boldsymbol{q'''s'''},\boldsymbol{qs}}^{\boldsymbol{q''s''}} \right) \right] \tag{4.44}$$

or, splitting the diagonal and off diagonal portions,

$$P_{\boldsymbol{qq'}}^{ss'} = \delta_{ss'} \delta_{qq'} \Gamma_{qs} + \Lambda_{qq'}^{ss'} \tag{4.45}$$

where $\Gamma_{\boldsymbol{qs}}$ is the diagonal portion, and $\Lambda_{\boldsymbol{qq'}}^{\mathbf{ss'}}$ is the off-diagonal portion of the interaction matrix. $\psi_{\boldsymbol{q's'}}$ is a measure of the disequilibrium of phonon mode $\boldsymbol{q's'}$. This equation describes the distribution of phonons of a particular mode $\boldsymbol{q}$ and polarization s within a domain. $P_{\boldsymbol{qq'}}^{ss'}$ is a scattering matrix, which describes the rate of phonon transition between a mode $\boldsymbol{q}$ and polarization s and a mode $\boldsymbol{q'}$ and polarization s'. The matrix is calculated by considering all three-phonon interactions that are possible, using Fermi's golden rule to calculate the transition probabilities. Here the $\bar{P}$ terms are the intrinsic transition probabilities which will be discussed in detail later. The canonical Boltzmann equation relates the gradient of the particle distributions and the particle velocities to their scattering rate with one another.

4.5.1.2 Interaction Probabilities Calculation

It is necessary to calculate the interaction probabilities that represent how the phonon modes scatter with one another. Srivastava[93] gives expressions for individual terms $\bar{P}_{\boldsymbol{qs}}^{\boldsymbol{q's'},\boldsymbol{q''s''}}$ and $\bar{P}_{\boldsymbol{qs},\boldsymbol{q's'}}^{\boldsymbol{q''s''}}$ by using the isotropic continuum model

$$\bar{P}_{\boldsymbol{qs},\boldsymbol{q's'}}^{\boldsymbol{q''s''}} = \frac{\pi\hbar}{4\rho^3 N_0 \mathrm{V}} \frac{qq'q''}{c_s c_{s'} c_{s''}} \left| A_{\boldsymbol{qq'q''}}^{ss's''} \right|^2 \bar{n}_{\boldsymbol{qs}} \bar{n}_{\boldsymbol{q's'}} \left(\bar{n}_{\boldsymbol{q''s''}} + 1 \right) \delta_{\boldsymbol{q+q'+q''},\boldsymbol{G}} \delta\left(\omega_{\boldsymbol{q''s''}} - \omega_{\boldsymbol{qs}} - \omega_{\boldsymbol{q's'}} \right) \tag{4.46}$$

$$\bar{P}_{\boldsymbol{qs}}^{\boldsymbol{q's'},\boldsymbol{q''s''}} = \frac{\pi\hbar}{4\rho^3 N_0 \mathrm{V}} \frac{qq'q''}{c_s c_{s'} c_{s''}} \left| A_{\boldsymbol{qq'q''}}^{ss's''} \right|^2 \left(\bar{n}_{\boldsymbol{qs}} + 1 \right) \bar{n}_{\boldsymbol{q's'}} \bar{n}_{\boldsymbol{q''s''}} \delta_{\boldsymbol{q+q'+q''},\boldsymbol{G}} \delta\left(\omega_{\boldsymbol{qs}} - \omega_{\boldsymbol{q's'}} - \omega_{\boldsymbol{q''s''}} \right) \tag{4.47}$$

In these expressions $\bar{P}^{\boldsymbol{q''s''}}_{\boldsymbol{qs},\boldsymbol{q's'}}$ is the interaction probability of a fusion process where $\boldsymbol{qs} + \boldsymbol{q's'} \rightarrow \boldsymbol{q''s''}$, $\bar{P}^{\boldsymbol{q's',q''s''}}_{\boldsymbol{qs}}$ represents the interaction probability of a fission process where $\boldsymbol{qs} \rightarrow \boldsymbol{q's'} + \boldsymbol{q''s''}$, $\bar{n}_{\boldsymbol{qs}}$ represents the equilibrium occupation number of a given mode $\boldsymbol{qs}$, $A^{ss's''}_{\boldsymbol{qq'q''}}$ is the scattering strength for this phonon process, $\delta_{\boldsymbol{q+q'+q''},G}$ represents the delta function that enforces conservation of momentum, $\delta(\omega_{\boldsymbol{q''s''}} - \omega_{\boldsymbol{qs}} - \omega_{\boldsymbol{q's'}})$ represents the delta function that enforces conservation of energy, ρ is the density, V is the total crystal volume, and c_s is the phonon speed.

Following Christensen[2] and Srivastava, among others, the scattering strength is represented by the Gruneisen parameter, γ, which represents a measure of the change of the frequencies of the phonons with changing temperature, the solid density and the modal speeds of the phonons,

$$\left|A^{ss's''}_{\boldsymbol{qq'q''}}\right|^2 = \frac{4\rho^2}{\bar{c}^2}\gamma^2 c_s^2 c_{s'}^2 c_{s''}^2 \tag{4.48}$$

Keeping in mind that (under the isotropic continuum assumption) $\omega_{\boldsymbol{qs}} = c_s \cdot q$, and using $F^2 = \frac{\gamma^2}{\bar{c}^2}$, it is then possible to write the two interaction probabilities for fusion and fission processes as

$$\bar{P}^{\boldsymbol{q''s''}}_{\boldsymbol{qs},\boldsymbol{q's'}} = \frac{\pi\hbar}{\rho N_0 \mathrm{V}} F^2 \omega_{\boldsymbol{qs}}\omega_{\boldsymbol{q's'}}\omega_{\boldsymbol{q''s''}}\bar{n}_{\boldsymbol{qs}}\bar{n}_{\boldsymbol{q's'}}(\bar{n}_{\boldsymbol{q''s''}} + 1)\delta_{\boldsymbol{q+q'+q''},\boldsymbol{G}}\delta(\omega_{\boldsymbol{q''s''}} - \omega_{\boldsymbol{qs}} - \omega_{\boldsymbol{q's'}}) \tag{4.49}$$

$$\bar{P}^{\boldsymbol{q's',q''s''}}_{\boldsymbol{qs}} = \frac{\pi\hbar}{\rho N_0 \mathrm{V}} F^2 \omega_{\boldsymbol{qs}}\omega_{\boldsymbol{q's'}}\omega_{\boldsymbol{q''s''}}(\bar{n}_{\boldsymbol{qs}} + 1)\bar{n}_{\boldsymbol{q's'}}\bar{n}_{\boldsymbol{q''s''}}\delta_{\boldsymbol{q+q'+q''},\boldsymbol{G}}\delta(\omega_{\boldsymbol{qs}} - \omega_{\boldsymbol{q's'}} - \omega_{\boldsymbol{q''s''}}) \tag{4.50}$$

Upon insertion of a Gaussian peak instead of the delta functions the following expressions are obtained.

$$\bar{P}^{\boldsymbol{q''s''}}_{\boldsymbol{qs},\boldsymbol{q's'}} = \frac{\sqrt{\pi}\hbar}{\rho N_0 \mathrm{V}} \frac{F^2 \omega_{\boldsymbol{qs}}\omega_{\boldsymbol{q's'}}}{\sigma} \bar{n}_{\boldsymbol{qs}}\bar{n}_{\boldsymbol{q's'}}(\bar{n}_{\boldsymbol{q''s''}} + 1)\delta_{\boldsymbol{q+q'+q''},\boldsymbol{G}} \exp\left[-\left(\frac{\omega_{\boldsymbol{q''s''}} - \omega_{\boldsymbol{qs}} - \omega_{\boldsymbol{q's'}}}{\sigma\omega_{\boldsymbol{q''s''}}}\right)^2\right] \tag{4.51}$$

$$\bar{P}^{\boldsymbol{q's',q''s''}}_{\boldsymbol{qs}} = \frac{\sqrt{\pi}\hbar}{\rho N_0 \mathrm{V}} \frac{F^2 \omega_{\boldsymbol{q's'}}\omega_{\boldsymbol{q''s''}}}{\sigma} (\bar{n}_{\boldsymbol{qs}} + 1)\bar{n}_{\boldsymbol{q's'}}\bar{n}_{\boldsymbol{q''s''}}\delta_{\boldsymbol{q+q'+q''},\boldsymbol{G}} \exp\left[-\left(\frac{\omega_{\boldsymbol{qs}} - \omega_{\boldsymbol{q's'}} - \omega_{\boldsymbol{q''s''}}}{\sigma\omega_{\boldsymbol{qs}}}\right)^2\right] \tag{4.52}$$

As discussed before both of the three phonon processes can be represented by a single scattering matrix given in eq. 4.44 that represents nominal transition rates between phonons of different modes and polarization given their deviation from equilibrium.

To eliminate the sum over $\boldsymbol{q'''}$ in eq. 4.44 one should be able to write $\boldsymbol{q'''} = -\boldsymbol{q} - \boldsymbol{q''} + \boldsymbol{G}$ where $\boldsymbol{G}$ is a reciprocal lattice vector. Then the scattering matrix becomes

$$P_{\boldsymbol{qq'}}^{ss'} = \sum_{\boldsymbol{q''}s''} \sum_{s'''} \left[\frac{1}{2} \delta_{ss'} \delta_{\boldsymbol{qq'}} \left(\bar{P}_{\boldsymbol{qs},\boldsymbol{q'''}s'''}^{\boldsymbol{q''}s''} + \bar{P}_{\boldsymbol{qs}}^{\boldsymbol{q'''}s''',\boldsymbol{q''}s''} + \bar{P}_{\boldsymbol{qs},\boldsymbol{q''}s''}^{\boldsymbol{q'''}s'''} \right) - \delta_{s's'''} \delta_{\boldsymbol{q'q'''}} \left(\bar{P}_{\boldsymbol{qs},\boldsymbol{q''}s''}^{\boldsymbol{q'''}s'''} + \bar{P}_{\boldsymbol{q''}s'',\boldsymbol{q'''}s'''}^{\boldsymbol{qs}} - \bar{P}_{\boldsymbol{q'''}s''',\boldsymbol{qs}}^{\boldsymbol{q''}s''} \right) \right] \tag{4.53}$$

and now there is no longer a delta function for momentum conservation in the *P*s in eqs. 4.51 and 4.52 so the interaction rates become

$$\bar{P}_{\boldsymbol{qs},\boldsymbol{q's'}}^{\boldsymbol{q''}s''} = \frac{\sqrt{\pi}\hbar}{\rho N_0 \mathrm{V}} \frac{F^2 \omega_{\boldsymbol{qs}} \omega_{\boldsymbol{q's'}}}{\sigma} \bar{n}_{\boldsymbol{qs}} \bar{n}_{\boldsymbol{q's'}} \left(\bar{n}_{\boldsymbol{q''}s''} + 1 \right) \exp\left[- \left(\frac{\omega_{\boldsymbol{q''}s''} - \omega_{\boldsymbol{qs}} - \omega_{\boldsymbol{q's'}}}{\sigma \omega_{\boldsymbol{q''}s''}} \right)^2 \right] \tag{4.54}$$

$$\bar{P}_{\boldsymbol{qs}}^{\boldsymbol{q's'},\boldsymbol{q''}s''} = \frac{\sqrt{\pi}\hbar}{\rho N_0 \mathrm{V}} \frac{F^2 \omega_{\boldsymbol{q's'}} \omega_{\boldsymbol{q''}s''}}{\sigma} \left(\bar{n}_{qs} + 1 \right) \bar{n}_{\boldsymbol{q's'}} \bar{n}_{\boldsymbol{q''}s''} \exp\left[- \left(\frac{\omega_{\boldsymbol{qs}} - \omega_{\boldsymbol{q's'}} - \omega_{\boldsymbol{q''}s''}}{\sigma \omega_{\boldsymbol{qs}}} \right)^2 \right] \tag{4.55}$$

Expanding sums from $P_{\boldsymbol{qq'}}^{ss'}$ above

$$P_{\boldsymbol{qq'}}^{ss'} = \frac{1}{2} \delta_{ss'} \delta_{\boldsymbol{qq'}} \sum_{\boldsymbol{q''}s''} \sum_{s'''} \left(\bar{P}_{\boldsymbol{qs},\boldsymbol{q'''}s'''}^{\boldsymbol{q''}s''} + \bar{P}_{\boldsymbol{qs}}^{\boldsymbol{q'''}s''',\boldsymbol{q''}s''} + \bar{P}_{\boldsymbol{qs},\boldsymbol{q''}s''}^{\boldsymbol{q'''}s'''} \right) - \delta_{\boldsymbol{q'q'''}} \sum_{\boldsymbol{q''}s''} \sum_{s'''} \delta_{s's'''} \left(\bar{P}_{\boldsymbol{qs},\boldsymbol{q''}s''}^{\boldsymbol{q'''}s'''} + \bar{P}_{\boldsymbol{q''}s'',\boldsymbol{q'''}s'''}^{\boldsymbol{qs}} - \bar{P}_{\boldsymbol{q'''}s''',\boldsymbol{qs}}^{\boldsymbol{q''}s''} \right) \tag{4.56}$$

where the two terms are the diagonal and off diagonal portions of the matrix written as

$$P_{\boldsymbol{qq'}}^{ss'} = \delta_{ss'} \delta_{qq'} \Gamma_{qs} + \Lambda_{\boldsymbol{qq'}}^{\boldsymbol{ss'}} \tag{4.57}$$

The diagonals, $\Gamma_{\boldsymbol{qs}}$, are given by

$$\Gamma_{\boldsymbol{qs}} = \frac{1}{2} \sum_{\boldsymbol{q's'}} \sum_{s''} \left(\bar{P}_{\boldsymbol{qs},\boldsymbol{q''}s''}^{\boldsymbol{q's'}} + \bar{P}_{\boldsymbol{qs}}^{\boldsymbol{q''}s'',\boldsymbol{q's'}} + \bar{P}_{\boldsymbol{qs},\boldsymbol{q's'}}^{\boldsymbol{q''}s''} \right) \tag{4.58}$$

where now $\boldsymbol{q''} = -\boldsymbol{q} - \boldsymbol{q'}$. The off diagonals, $\Lambda_{\boldsymbol{qq'}}^{\boldsymbol{ss'}}$, are given by

$$\Lambda_{\boldsymbol{qq'}}^{ss'} = -\sum_{\boldsymbol{q''s''}} \left(\bar{P}_{\boldsymbol{qs,q''s''}}^{\boldsymbol{q's'}} + \bar{P}_{\boldsymbol{q''s'',q's'}}^{\boldsymbol{qs}} - \bar{P}_{\boldsymbol{q's',qs}}^{\boldsymbol{q''s''}} \right) \tag{4.59}$$

and now $\boldsymbol{q'} = -\boldsymbol{q} - \boldsymbol{q''}$, or $\boldsymbol{q''} = -\boldsymbol{q} - \boldsymbol{q'}$. This should make it possible to eliminate the final sum, giving

$$\Lambda_{\boldsymbol{qq'}}^{ss'} = -\sum_{\boldsymbol{s''}} \left(\bar{P}_{\boldsymbol{qs,q''s''}}^{\boldsymbol{q's'}} + \bar{P}_{\boldsymbol{q''s'',q's'}}^{\boldsymbol{qs}} - \bar{P}_{\boldsymbol{q's',qs}}^{\boldsymbol{q''s''}} \right) \tag{4.60}$$

4.5.1.3 Grouped Phonon Model

In this work the goal is to capture how different phonon energies interact so that the effect of out of equilibrium phonons can be captured. It is impossible to consider every single phonon mode within a finite domain since there will be as many phonon modes as atoms in the crystal. Therefore this work considers a grouped phonon transport model where the phonon modes are represented as groups of phonons with associated energy. This grouped phonon model considers that the phonon dispersion is split into energy bands, as shown in Figure 4.11, where phonons within each band are considered to have the same equivalent temperature.

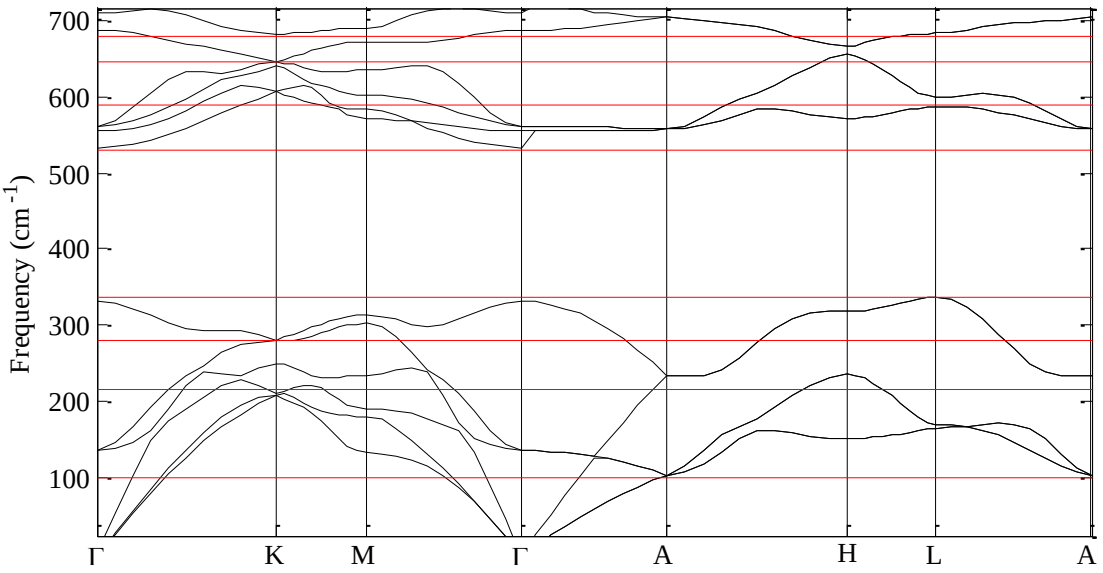

Figure 4.11. The phonon dispersion for GaN as calculated by ABINIT split into 9 bands where all the phonon modes within each band are considered for the purposes of this model to have the same equivalent temperautre.

The canonical BTE is converted to energy form by multiplying by $\hbar\omega_{\boldsymbol{qs}}/V$ where $\omega_{\boldsymbol{qs}}$ is the phonon frequency and V is the volume of the unit cell. It is necessary to divide by V in order to obtain energy density, in units of $\mathrm{J/m^3}$, not Brillouin zone energy. Then the BTE becomes

$$-\frac{\partial e_{\boldsymbol{qs}}}{\partial t} - \boldsymbol{v}_{\boldsymbol{qs}} \cdot \nabla e_{\boldsymbol{qs}} = \frac{\hbar\omega_{\boldsymbol{qs}}}{V} \sum_{\boldsymbol{q's'}} P_{\boldsymbol{qq'}}^{ss'} \psi_{\boldsymbol{q's'}} \tag{4.61}$$

with $e_{\boldsymbol{qs}}$ being the phonon energy.

Now, the isotropic grouped BTE should take a similar form, where now modes are represented by their respective groups, and the directionality of modes is lost, instead being captured isotropically by a continuous distribution of phonon energies around the solid angle. The scattering term on the right hand side now must include scattering from all directions of all groups. The grouped BTE is can be given by

$$-\frac{\partial e_{\boldsymbol{qs}}}{\partial t} - v_g \boldsymbol{s}(\boldsymbol{\Omega}) \cdot \nabla e_g(\boldsymbol{\Omega}) = \frac{\hbar\omega_g}{4V\pi} \sum_{g'} \int_{4\pi} P_{gg'} \delta(\boldsymbol{\Omega} - \boldsymbol{\Omega}') \psi_{g'}(\boldsymbol{\Omega})\, d\boldsymbol{\Omega}' \tag{4.62}$$

This equation includes the directional delta function which indicates directionally isotropic scattering. In this equation any disequilibrium in one direction will not affect other directions except through changing the lattice temperature, in contrast to how different modes interact; each mode can interact with other modes in specific ways. In this way the equation is directionally isotropic. Without this delta function the contributions from the individual modes would cancel and no band-to-band interactions would be captured. In this expression $\delta(\boldsymbol{\Omega} - \boldsymbol{\Omega}')$ indicates the directional delta function, which is defined by

$$\frac{1}{4\pi} \int_{4\pi} f(\boldsymbol{\Omega}') \delta(\boldsymbol{\Omega} - \boldsymbol{\Omega}') d\boldsymbol{\Omega}' = f(\boldsymbol{\Omega}) \tag{4.63}$$

To solve the grouped BTE, the group-wise phonon properties must be related to the individual phonon mode properties. These group-wise properties are the group

frequency ω_g, the interaction matrix $P_{gg\prime}$, how the disequilibrium term ψ_g relates to the energies of the groups e_g, and the group velocity v_g. Now, to find the different grouped variables, several assumptions are made. First, it is assumed the sum of the modal energies equals the group energy

$$e_g = \sum_{\boldsymbol{q}s\in g} e_{\boldsymbol{q}s} \tag{4.64}$$

and that the number of total phonons is conserved

$$n_g = \sum_{\boldsymbol{q}s\in g} n_{\boldsymbol{q}s} \tag{4.65}$$

This gives the requirement that the frequency of a group is a phonon occupation number weighted average of the frequencies of the constituent modes.

$$\omega_g = \frac{\sum_{\boldsymbol{q}s\in g} n_{\boldsymbol{q}s}\omega_{\boldsymbol{q}s}}{\sum_{\boldsymbol{q}s\in g} n_{\boldsymbol{q}s}} \tag{4.66}$$

This leads to the natural decision that the grouped phonons should all be of similar frequency. As shall be seen later, in this model the phonons are grouped into bands by frequency.

4.5.1.4 Group-wise Interaction Matrix, Group-wise Expression Derivation

Next, to calculate the group wise interaction matrix, it is assumed the contribution to the group-wise scattering term from each individual mode is the conserved, so that

$$P_{gg\prime}\psi_{g\prime} = \sum_{\boldsymbol{q}s\in g}\sum_{\boldsymbol{q}\prime s\prime\in g\prime} P^{ss\prime}_{\boldsymbol{q}\boldsymbol{q}\prime}\psi_{\boldsymbol{q}\prime s\prime} \tag{4.67}$$

In this way, all interactions within a group should be captured. The above simply states that all phonon interactions between phonons in group g and group g' are captured. Somehow the group-wise disequilibrium ψ_g must be related to the modal disequilibrium $\psi_{\boldsymbol{q}s}$. It must be assumed that all modes in a group are 'equally out of equilibrium' meaning they have the same equivalent lattice temperature T_g. In a sense, this means that

if these modes were the only modes within the material, that material would be in equilibrium at temperature T_g. $\psi_{\boldsymbol{qs}}$ is defined by Srivastava by the following expression

$$n_{\boldsymbol{qs}} = \frac{1}{\exp\left[\frac{\hbar\omega_{\boldsymbol{qs}}}{k_B T} - \psi_{\boldsymbol{qs}}\right] - 1} \tag{4.68}$$

where T is the lattice temperature. If the occupation of a mode can also be described by the commonly accepted Bose-Einstein distribution for phonons[78]

$$n_{\boldsymbol{qs}} = \frac{1}{\exp\left[\frac{\hbar\omega_{\boldsymbol{qs}}}{k_B T_{\boldsymbol{qs}}}\right] - 1} \tag{4.69}$$

where $T_{\boldsymbol{qs}}$ is the equivalent lattice temperature of mode $\boldsymbol{qs}$, then it is possible to equate the two and define the modal disequilibrium $\psi_{\boldsymbol{qs}}$ in terms of modal temperature and lattice temperature

$$\psi_{\boldsymbol{qs}} = \frac{\hbar\omega_{\boldsymbol{qs}}}{k_B T} - \frac{\hbar\omega_{\boldsymbol{qs}}}{k_B T_{\boldsymbol{qs}}} \tag{4.70}$$

Now, linearizing this expression in $T_{\boldsymbol{qs}}$, it is found that the modal disequilibrium can be roughly given by

$$\psi_{\boldsymbol{qs}} \approx \bar{\psi}_{\boldsymbol{qs}} = \frac{\hbar\omega_{\boldsymbol{qs}}(T_{\boldsymbol{qs}} - T)}{k_B T^2} \tag{4.71}$$

This equation will only be valid for small temperature differences, and the relative error in this approximation is given by

$$\epsilon = \frac{\bar{\psi}_{\boldsymbol{qs}} - \psi_{\boldsymbol{qs}}}{\psi_{\boldsymbol{qs}}} = \frac{T_{\boldsymbol{qs}}}{T} - 1 \tag{4.72}$$

This error grows above an acceptable level pretty quickly as the group temperature deviates from the lattice temperature, and so this model will not accurately capture large variations in temperature. Further study needs to be done on the effect of this linearization procedure. However, at the present time this model is simple and will

capture the effect of mode-to-mode transitions and offer insight into the behavior of out of equilibrium phonons.

Enforcing the group temperature equivalence ensures that the temperature of the individual phonons are equal to the temperature of the group those phonons are in, as given by

$$T_{\boldsymbol{q}s} = T_g \tag{4.73}$$

so that the disequilibrium term for a mode can also be related to the group temperature.

$$\psi_{\boldsymbol{q}s} \approx \frac{\hbar\omega_{\boldsymbol{q}s}(T_g - T)}{k_B T^2} \tag{4.74}$$

By equivalence, the group-wise disequilibrium term can also be related to the group temperature.

$$\psi_g \approx \frac{\hbar\omega_g(T_g - T)}{k_B T^2} \tag{4.75}$$

Substituting these expressions into the scattering equivalence expression eq. 4.67, it is found that

$$P_{gg\prime}\frac{\hbar\omega_{g\prime}(T_{g\prime} - T)}{k_B T^2} = \sum_{\boldsymbol{q}s\in g}\sum_{\boldsymbol{q}\prime s\prime\in g\prime} P^{ss\prime}_{\boldsymbol{q}\boldsymbol{q}\prime}\frac{\hbar\omega_{\boldsymbol{q}\prime s\prime}(T_{g\prime} - T)}{k_B T^2} \tag{4.76}$$

$$P_{gg\prime}\omega_{g\prime} = \sum_{\boldsymbol{q}s\in g}\sum_{\boldsymbol{q}\prime s\prime\in g\prime} P^{ss\prime}_{\boldsymbol{q}\boldsymbol{q}\prime}\omega_{\boldsymbol{q}\prime s\prime} \tag{4.77}$$

so that the group-to-group interaction matrix is simply a weighted average of the appropriate elements of the full interaction operator

$$P_{gg\prime} = \sum_{\boldsymbol{q}s\in g}\sum_{\boldsymbol{q}\prime s\prime\in g\prime} P^{ss\prime}_{\boldsymbol{q}\boldsymbol{q}\prime}\frac{\omega_{\boldsymbol{q}\prime s\prime}}{\omega_{g\prime}} \tag{4.78}$$

It is occasionally convenient to relate the diagonal portion of the group-wise matrix to the diagonal portion of the full scattering matrix by

$$\Gamma_g = \sum_{\boldsymbol{q}s \in g} \left(\Gamma_{\boldsymbol{q}s} \frac{\omega_{\boldsymbol{q}s}}{\omega_g} \right) \tag{4.79}$$

and the off-diagonal portions relate by

$$\Lambda_{gg'} = \sum_{\boldsymbol{q}s \in g} \sum_{\boldsymbol{q}'s' \in g'} \left(\Lambda^{ss'}_{\boldsymbol{q}\boldsymbol{q}'} \frac{\omega_{\boldsymbol{q}'s'}}{\omega_{g'}} \right) \tag{4.80}$$

so that the total group-wise scattering matrix is given by the combination of the diagonal and off diagonal portions of the matrix

$$P_{gg'} = \Gamma_g \delta_{gg'} + \Lambda_{gg'} \tag{4.81}$$

4.5.1.5 Linearization of Perturbation Terms

In order to solve the group BTE the group-wise disequilibrium terms $\psi_g(\boldsymbol{\Omega})$ must be related with $e_g(\boldsymbol{\Omega})$. It is known from above that the disequilibrium term approximately relates to the temperatures and frequencies of the group,

$$\psi_g(\boldsymbol{\Omega}) \approx \frac{\hbar \omega_g \left(T_g(\boldsymbol{\Omega}) - T \right)}{k_B T^2} \tag{4.82}$$

where $T_g(\boldsymbol{\Omega})$ is a directional temperature which is a representation of the temperature of band g in direction $\boldsymbol{\Omega}$. It is also known that the energy relates to the temperature via the specific heat with

$$e_g(\boldsymbol{\Omega}) = \int_{T_{ref}}^{T_g(\boldsymbol{\Omega})} C_g dT \approx C_g \left(T_g(\boldsymbol{\Omega}) - T_{ref} \right) \tag{4.83}$$

$$T_g(\boldsymbol{\Omega}) \approx \frac{e_g(\boldsymbol{\Omega})}{C_g} + T_{ref} \tag{4.84}$$

where C_g is the group specific heat, usually in units $\mathrm{J/m^3/K}$. This is obviously an approximation of constant specific heats which is another big assumption within this framework. The group specific heat is simply the sum of the modal specific heats,

$$C_g = \sum_{\boldsymbol{q}s \in g} C_{\boldsymbol{q}s} = \sum_{\boldsymbol{q}s \in g} k_B \left(\frac{\hbar\omega_{\boldsymbol{q}s}}{k_B T}\right)^2 \frac{\exp\left(\frac{\hbar\omega_{\boldsymbol{q}s}}{k_B T}\right)}{\left[\exp\left(\frac{\hbar\omega_{\boldsymbol{q}s}}{k_B T}\right) - 1\right]^2} \tag{4.85}$$

In the same way the lattice temperature relates to the total lattice energy and the total specific heat capacity by

$$T \approx \frac{e}{C} + T_{ref} \tag{4.86}$$

where C is the total specific heat capacity given by

$$C = \sum_g C_g \tag{4.87}$$

From these expressions it is possible to write the disequilibrium term as a function of group energy and lattice energy

$$\psi_g(\boldsymbol{\Omega}) \approx \frac{\hbar\omega_g}{k_B {T_0}^2}\left(\frac{e_g(\boldsymbol{\Omega})}{C_g} - \frac{e}{C}\right) \tag{4.88}$$

T_0 here represents the lattice temperature but for the linearity of the model it is pre-selected to be the temperature at which the scattering matrix is calculated, and is not dependent on the group-wise energies $e_g(\boldsymbol{\Omega})$.

This model has been developed in a strictly linear way, with constant specific heats and the assumption that the disequilibrium term can be linearly related to temperature. This model at this time also includes no impurity or boundary scattering. Boundary scattering will be implemented directly with geometry using the discrete ordinates method in the following sections, but isotope/defect scattering could play an important role in the reduction of the thermal conductivity especially at lower temperatures. While these are big assumptions, lifting them is likely to be much more complicated, and these expressions will be useful to discover behavior of the model and band-to-band behavior. In the future work section it is suggested that lifting these constant property assumptions is pursued.

4.5.1.6 Directional Discretization, Scattering Matrix Development

If the new linearized perturbation term (eq. 4.88) is substituted into the isotropic band-to-band BTE (eq. 4.62), and a generation term is added, the steady state BTE becomes

$$-v_g \boldsymbol{s}(\boldsymbol{\Omega}) \cdot \nabla e_g(\boldsymbol{\Omega}) = \frac{\hbar^2 \omega_g}{4\pi k_B {T_0}^2 V} \sum_{g'} \int_{4\pi} P_{gg'} \delta(\boldsymbol{\Omega} - \boldsymbol{\Omega}') \omega_{g'} \left(\frac{e_{g'}(\boldsymbol{\Omega}')}{C_{g'}} - \frac{e}{C} \right) d\boldsymbol{\Omega}' + G_g(\boldsymbol{\Omega}) \quad (4.89)$$

where $G_g(\boldsymbol{\Omega})$ is generation given in units of $\mathrm{W/m^3}$. The integral over the solid angle can be directly evaluated to give

$$-v_g \boldsymbol{s}(\boldsymbol{\Omega}) \cdot \nabla e_g(\boldsymbol{\Omega}) = \frac{\hbar^2 \omega_g}{4\pi k_B {T_0}^2 V} \sum_{g'} P_{gg'} \omega_{g'} \left(\frac{e_{g'}(\boldsymbol{\Omega})}{C_{g'}} - \frac{e}{C} \right) + G_g(\boldsymbol{\Omega}) \quad (4.90)$$

Now, following the discrete ordinates method (DOM) by integrating over selected control angles, and cancelling weight factors that appear on both sides, the discretized band-to-band BTE is

$$-v_g \boldsymbol{s}_i \cdot \nabla e_{gi} = \frac{\hbar^2 \omega_g}{4\pi k_B {T_0}^2 V} \sum_{g'} P_{gg'} \omega_{g'} \left(\frac{e_{g'i}}{C_{g'}} - \frac{e}{C} \right) + G_{gi} \quad (4.91)$$

where subscripts i indicate discrete directions, the total energy is simply the sum of all the weighted band-wise directional energies,

$$e = \sum_{g''} \sum_{i''} w_{i''} e_{g''i''} \quad (4.92)$$

and the generation term is the generation term evaluated at the direction cosine

$$G_{gi} = G_g(\boldsymbol{\Omega}_i) \quad (4.93)$$

If $G_g(\boldsymbol{\Omega})$ is not a function of $\boldsymbol{\Omega}$, as is often the case, then

$$G_{gi} = G_g \quad (4.94)$$

which is assumed to be the case for the rest of this document. If there is a directionally dependent generation field, these expressions may need to be modified.

It is desirable to put the BTE into the form

$$-v_g \boldsymbol{s}_i \cdot \nabla e_{gi} = \sum_{g'} \sum_{i'} K_{gg'ii'} e_{g'i'} + G_g \tag{4.95}$$

Where $K_{gg'ii'}$ is a scattering matrix which describes the evolution of scattering effects with respect to changing band-wise energy distributions. This form of the equation is fully linear and should be relatively easy to solve. To put the BTE into this form, the discretized band-to-band BTE (eq. 4.91) is expanded as follows.

$$-v_g \boldsymbol{s}_i \cdot \nabla e_{gi} = \frac{\hbar^2 \omega_g}{k_B {T_0}^2 \mathrm{V}} \sum_{g'} P_{gg'} \omega_{g'} \left(\frac{e_{g'i}}{C_{g'}} - \frac{1}{C} \sum_{g''} \sum_{i''} w_{i''} e_{g''i''} \right) + G_g \tag{4.96}$$

$$-v_g \boldsymbol{s}_i \cdot \nabla e_{gi} = \frac{\hbar^2 \omega_g}{k_B {T_0}^2 \mathrm{V}} \left[\sum_{g'} \frac{P_{gg'} \omega_{g'}}{C_{g'}} e_{g'i} - \frac{1}{C} \sum_{g''} P_{gg''} \omega_{g''} \sum_{g'} \sum_{i'} w_{i'} e_{g'i'} \right] + G_g \tag{4.97}$$

So, by comparison it is not difficult to see that the scattering matrix should be

$$K_{gg'ii'} = \frac{\hbar^2 \omega_g}{k_B {T_0}^2 \mathrm{V}} \left(\frac{P_{gg'} \omega_{g'}}{C_{g'}} \delta_{ii'} - \frac{w_{i'}}{C} \sum_{g''} P_{gg''} \omega_{g''} \right) \tag{4.98}$$

4.5.1.7 Scattering Matrix Correction

There are several conditions that the scattering matrix must meet in order to ensure physicality of the model. To ensure conservation of energy, the scattering terms on the right hand side should all sum to zero, so that there is no creation or destruction of energy due to scattering. This leads to

$$\sum_{g} \sum_{i} w_i \sum_{g'} \sum_{i'} K_{gg'ii'} e_{g'i'} = 0 \tag{4.99}$$

for any energy distribution e_{gi}. Substituting in the scattering matrix and performing some trivial algebra, one obtains

$$\sum_{g'}\left(\sum_{i'} e_{g'i'}w_{i'}\right)\sum_{g}\omega_g\left(\frac{P_{gg'}\omega_{g'}}{C_{g'}}-\frac{\sum_{g''}P_{gg''}\omega_{g''}}{C}\right)=0 \tag{4.100}$$

so to conserve energy in all cases,

$$\sum_{g}\sum_{g''}\omega_g\left(\frac{C\omega_{g''}}{C_{g''}}\delta_{g'g''}-\omega_{g''}\right)P_{gg''}=0 \tag{4.101}$$

In addition, the model should not allow unphysical movement of phonons, specifically, two conditions must be prevented. First, the total contribution of all of a bands forcing terms must never act to push the band beyond any other bands out of equilibrium. This can be stated by considering that if all bands are at a given deviation from equilibrium, the total sum of the forcing terms must be positive (pushing the band towards equilibrium). This is written as

$$\sum_{g'}P_{gg'}\omega_{g'}>0 \tag{4.102}$$

for all bands g. Furthermore, the band-to-band interaction matrix must have strictly positive diagonals and strictly negative off-diagonals; this is because it is not physical for a band that is below equilibrium to force another band in the other direction away from equilibrium. In addition, the net in-scattering term must contribute to a return to equilibrium.

As calculated $P_{gg'}$ will not necessarily meet these conditions, primarily because the interaction strengths of the phonons are being represented by a single Gruneisen parameter for all modes and not with the third derivative of energy term which represents the true strengths of interactions between phonon modes. However, it is possible to calculate a new $\hat{P}_{gg'}$ which is very close to the originally calculated $P_{gg'}$ but that also meets these conditions as accurately as possible. An iterative procedure is performed on the interaction matrix $P_{gg'}$ until it meets these requirements. Firstly, any off-diagonal

elements that are above zero are zeroed out. Then, the conservation condition is enforced by writing eq. 4.101 in matrix form as

$$\boldsymbol{W}\boldsymbol{x} = \boldsymbol{0} \tag{4.103}$$

where $\boldsymbol{x}$ is the stacked columns of $P_{gg'}$, and $\boldsymbol{W}$ is a $n \times n^2$ matrix which represents the conservation condition. To find $\hat{P}_{gg'}$ it is necessary to find an $\hat{\boldsymbol{x}}$ which meets the condition above which is as close to $\boldsymbol{x}$ as possible. If $\mathbb{W}$ is a linear basis for the null space of $\boldsymbol{W}$, then one finds $\boldsymbol{u}$ such that

$$\|\mathbb{W}\boldsymbol{u} - \boldsymbol{x}\| \tag{4.104}$$

is minimized, then find

$$\hat{\boldsymbol{x}} = \mathbb{W}\boldsymbol{u} \tag{4.105}$$

This can achieved by finding the $\hat{\boldsymbol{x}}$ in the null space of $\boldsymbol{W}$ which is closest to $\boldsymbol{x}$. In this implementation, the null space of $\boldsymbol{W}$ is found by MATLAB's `null` operator which uses Singular Value Decomposition, and then the over-determined system is solved using MATLAB's backslash division operator which uses QR decomposition to solve the least squares system. Then, the physicality of the forcing terms must be adjusted in a similar way so that eq. 4.102 is met. Finally, it is necessary to adjust the new $\hat{P}_{gg'}$ so that thermal conductivity is maintained, which can be done by multiplying $\hat{P}_{gg'}$ by a constant. This process is repeated until all three conditions are satisfactorily met. This results in a matrix that is somewhat modified from the original matrix, but that meets the required conditions for the model to conserve energy and behave physically.

To illustrate the need to modify the interaction matrix so that the model conserves energy and behaves physically, several band-to-band test cases were run with:

a) the raw matrix,

b) a partially corrected energy conserving matrix that behaves non-physically,

c) a partially corrected matrix that meets eq. 4.102 but has positive off-diagonals, and finally

d) a corrected matrix that meets all these conditions.

The test case run was a one dimensional slab with a thin generation region with a total of 10^{10} W/m^2 in the center where all the heat is input into the highest energy mode. The quadrature set used was Sn with $n = 3$ with 500 cells, and 5 bands. To check physicality the equivalent temperatures of each direction and mode are compared. No modes should extend below the boundary temperature or above the generation band temperature.

Figure 4.12 shows the temperature of each band and direction within the solution over the whole domain for case a, where the raw scattering matrix is used. It looks physical and performs well except that it is non-conservative as can been seen in Figure 4.16. Correcting the conservation issue leads the matrix to become non-physical, in that some band temperatures are pushed below the boundary temperatures, which is physically impossible, see Figure 4.13. Correcting the summation of the forcing terms according to 4.102 gives a slightly better result that still appears non-physical near the generation region, see Figure 4.14. Finally, after all three corrections are in place, the matrix is conservative and physically reasonable, giving the results seen in 4.15. It looks like some of the phonon modes have almost collapsed onto one another, which is possibly not a physical effect, and the resulting change in temperature distribution is not known.

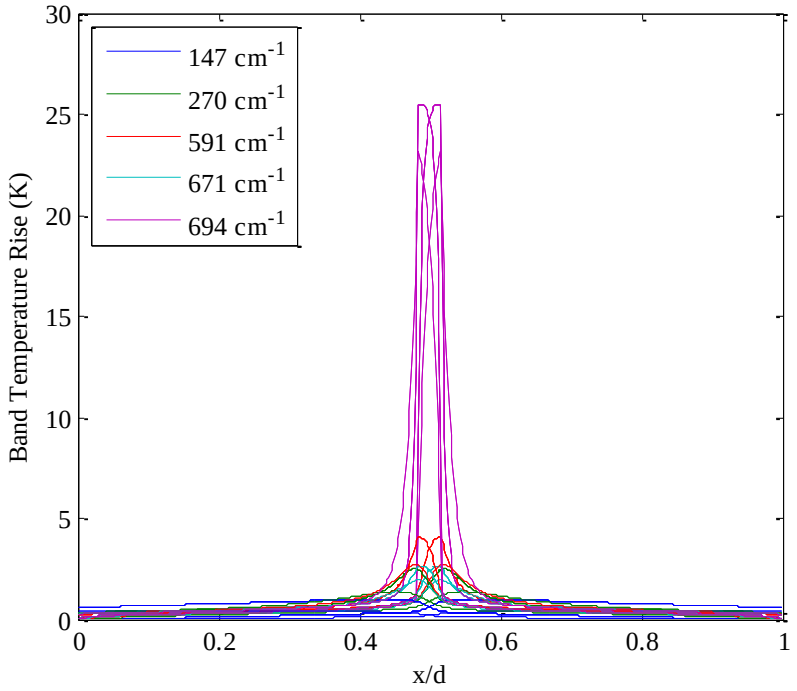

Figure 4.12. Band and direction-wise temperatures for case a. This looks physical, but note in Figure 4.16 it is highly non-conservative.

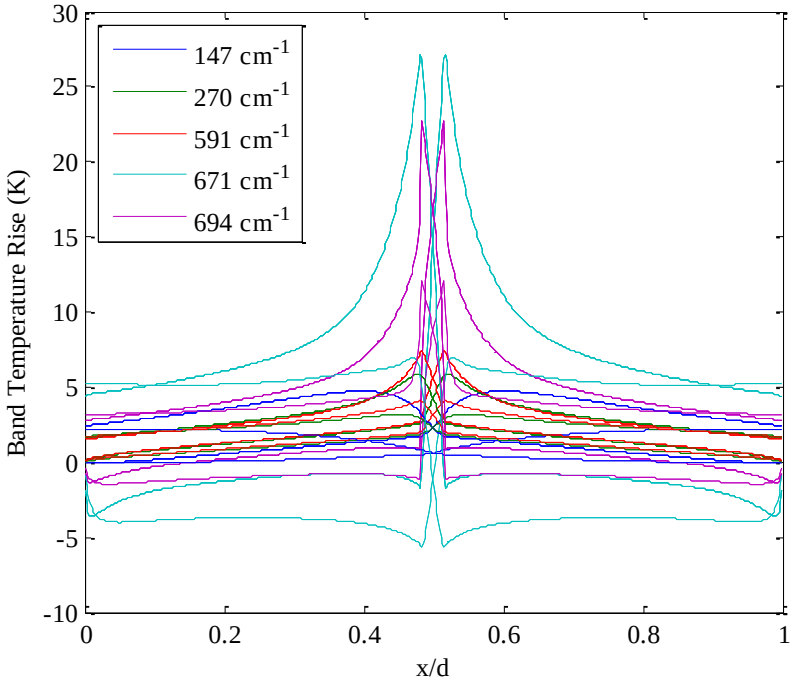

Figure 4.13. Band and direction-wise temperatures rises for case b. Note how some phonon temperatures are below the boundary temperature and above the generation band temperature.

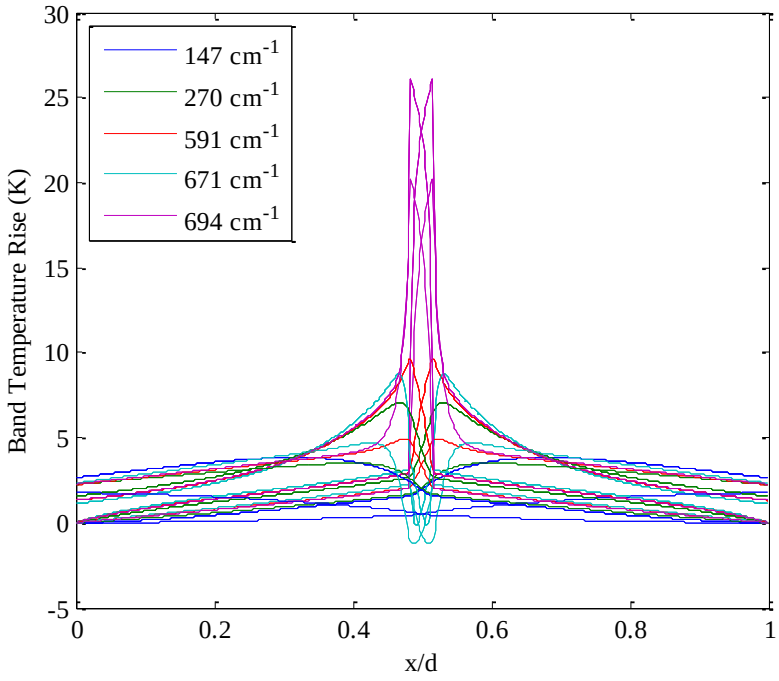

Figure 4.14. Band and direction-wise temperatures rises for case c. Note that there are strange dips in the 671 cm^{-1} mode near the generation region due to some negative interaction terms.

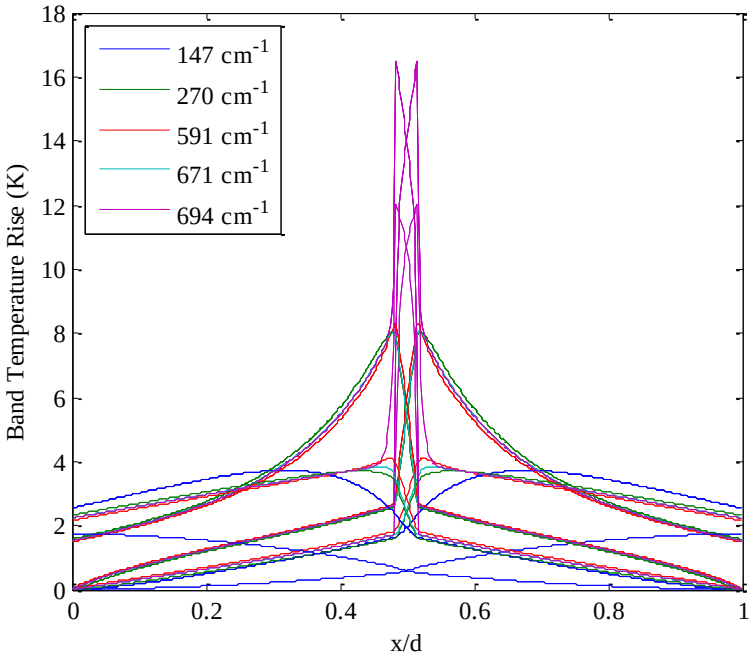

Figure 4.15. Band and direction-wise temperatures rises for case d, the final corrected matrix. This appears completely physical, and is conservative. There does appear to be some grouping and the curves look qualitatively different from those in case a. The peak phonon temperature is lower than in the other cases.

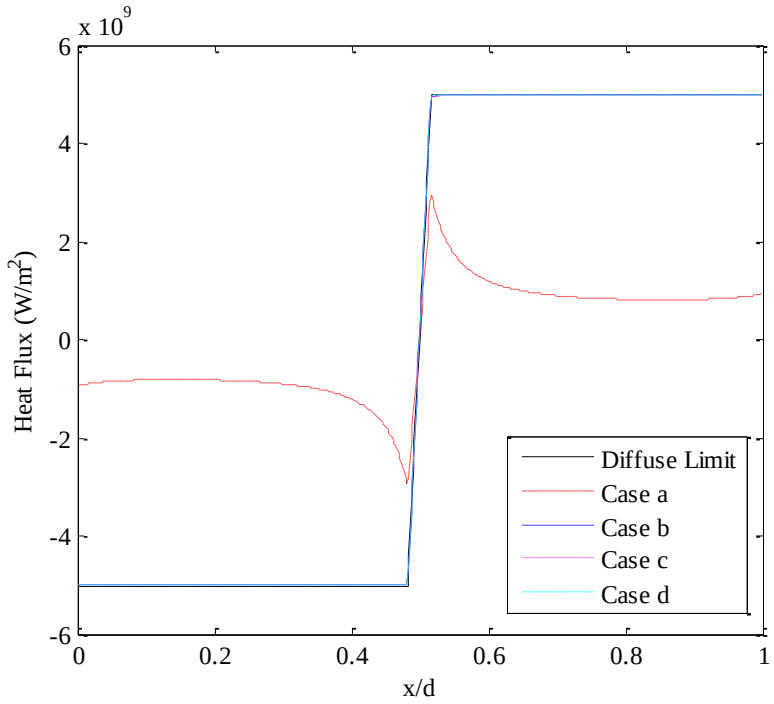

Figure 4.16. Heat fluxes for the different cases. All are conservative except case a.

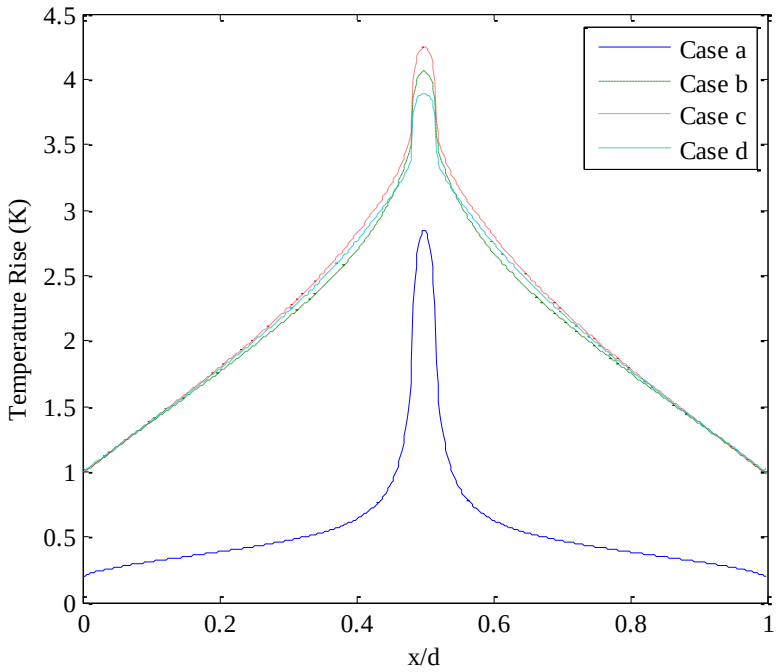

Figure 4.17. Total temperature rise for each case. There is a fairly large discrepancy between the raw case and the modified cases, most likely due to the non-conservative nature of the raw solution, since a large portion of the energy being input into the domain is being destroyed, and therefore the temperature is significantly lower.

The main question with all of these matrix corrections is: is the final scattering matrix even close to what was originally determined by the interaction expressions derived above, and does it physically represent what is occurring within a GaN crystal? The answer is hard to know. There is some limited information that can be used to determine how accurate a scattering matrix is. It is possible to compare the lifetime of a band to known phonon lifetimes from literature. Solving the transient zero-dimensional BTE (eq. 4.106) for an initial perturbation in the highest energy band (where the A_1(LO) resides), it is possible to see the decay of nearby phonon modes back to equilibrium neglecting any spatial effects.

$$-\frac{\partial e_{gi}}{\partial t} = \sum_{g'} K_{gg'} e_{g'} \tag{4.106}$$

$K_{gg'}$, the zero-dimensional scattering matrix, is given by

$$K_{gg'} = \frac{\hbar^2 \omega_g}{k_B {T_0}^2 \mathrm{V}} \left(\frac{P_{gg'} \omega_{g'}}{C_{g'}} - \frac{1}{C} \sum_{g''} P_{gg''} \omega_{g''} \right) \tag{4.107}$$

Discretizing in time the fully implicit form of the equation is

$$-e_g^{i+1} + e_g^i = \sum_{g'} K_{gg'} e_{g'}^{i+1} \left(t_{i+1} - t_i\right) \tag{4.108}$$

which can then be solved stepping through time given an initial phonon distribution.

The phonon spectrum was split into 9 separate frequencies and the A_1(LO) mode was excited and allowed to decay. The result of running case a and case d, and comparing to the upper and lower bounds from literature (τ=2.5 ps[94] and τ=0.75 ps[95])plotted in Figure 4.18. The matrix adjustment clearly affects the phonon decay by making it interact more strongly with the other phonons and relax more quickly, but the decay rate is still within the bounds from literature. It must be noted that the Raman measurements of these phonons are of the zone-center phonons while the model is capturing decay from phonons of all wave-vectors, and also capturing the effects of

some other phonon modes like the $E_1(LO)$. Comparing the band containing the $A_1(TO)$, $E_1(TO)$, and $E_2(high)$ again the decay is within the bounds of these phonon mode relaxation times. Finally, comparing the band containing the $E_2(low)$ mode shows that the un-corrected case is within the bounds of the literature but the corrected matrix has a relaxation time that is faster than the lowest bound in literature, around 5.5 ps which is about half the lowest literature value given by [50]. The decay of the band containing the $E_2(low)$ mode and those of the literature bounds are plotted in Figure 4.19.

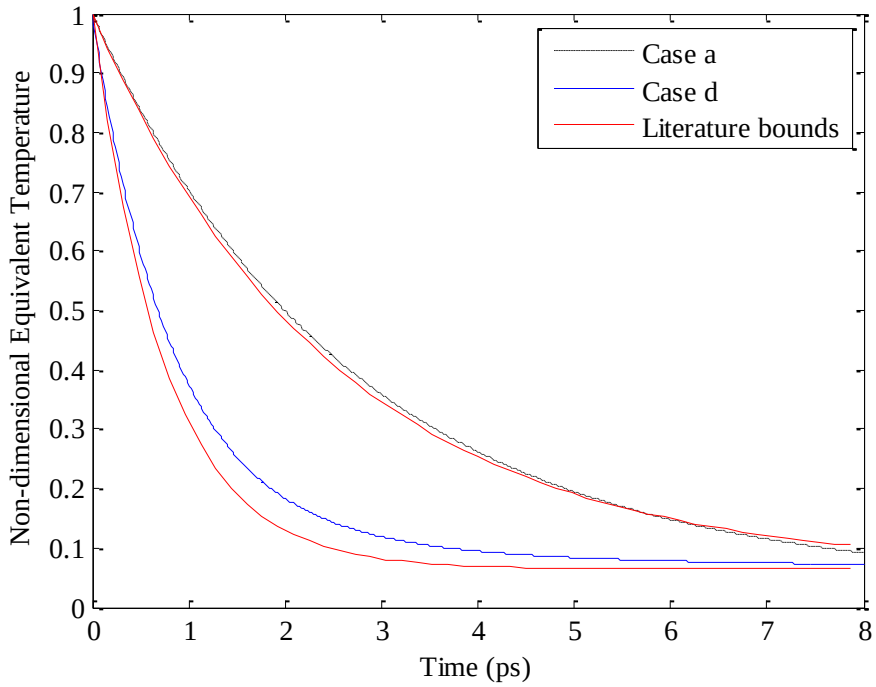

Figure 4.18. Transient response of two scattering matrices, a and d for an excited band that contains the $A_1(LO)$ mode plotted against the literature bounds for the relaxation of the phonon excitation.

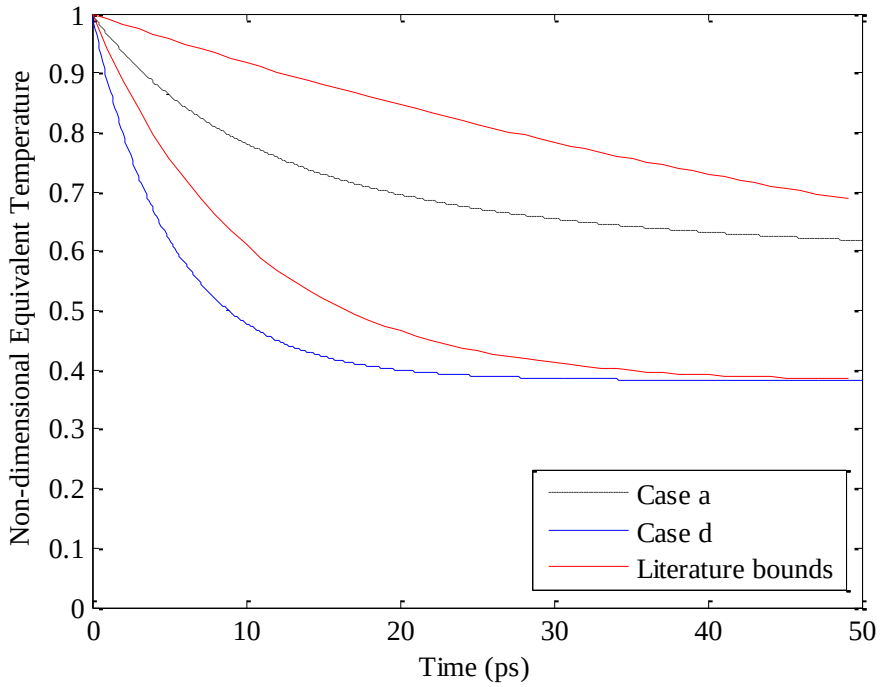

Figure 4.19. Transient response of the band containing the E_2(low) mode. Note that case a is not conservative and so is decaying to an incorrect final energy, but case d is outside the bounds of the E_2(low) phonon relaxation time.

4.5.1.8 Thermal Conductivity, Group velocities

To find the thermal conductivity related with this model, one must find a diffuse expression of energy expressed in terms of the gradient of the energy, and substitute this into the heat flux expression. First, the discretized BTE (eq. 4.91) is rearranged

$$\sum_{g'} \frac{P_{gg'}\omega_{g'}}{C_{g'}} e_{g'i} = \frac{1}{C}\sum_{g''} P_{gg''}\omega_{g''} \sum_{g'}\sum_{i'} w_{i'} e_{g'i'} - \frac{k_B {T_0}^2 \mathrm{V} v_g}{\hbar^2 \omega_g} \boldsymbol{s}_i \cdot \nabla e_{gi} \qquad (4.109)$$

This rearranged Boltzmann equation can be written in matrix form as

$$\boldsymbol{A}\boldsymbol{e}^i = \boldsymbol{b} - \boldsymbol{c}^i \qquad (4.110)$$

where $\boldsymbol{A}$ is a matrix given by

$$\{\boldsymbol{A}\}_{gg'} = \frac{P_{gg'}\omega_{g'}}{C_{g'}} \qquad (4.111)$$

(where the notation $\{\boldsymbol{M}\}_{ij}$ indicates the element in row i and column j), $\boldsymbol{e}^i$ is a column vector, dependent on direction i, given by

$$\{\boldsymbol{e}^i\}_g = e_{gi} \tag{4.112}$$

$\boldsymbol{b}$ is a column vector defined by

$$\{\boldsymbol{b}\}_g = \sum_{g''} \frac{P_{gg''}\omega_{g''}}{C} \sum_{g'} \sum_{i'} w_{i'} e_{g'i'} \tag{4.113}$$

and $\boldsymbol{c}^i$ is a column vector dependent on direction, given by

$$\{\boldsymbol{c}^i\}_g = \frac{k_B {T_0}^2 \mathrm{V}}{\hbar^2 \omega_g} v_g \boldsymbol{s}_i \cdot \nabla e_{gi} \tag{4.114}$$

In the diffusive regime, where conductivity is well defined, the energy gradient of all the modes is equivalent to the specific heat of that mode times the temperature gradient,

$$\nabla e_{gi} = \frac{\partial e_{gi}}{\partial T} \nabla T = C_g \nabla T \tag{4.115}$$

which allows $\boldsymbol{c}^i$ to also be given in terms of temperature gradient by the expression

$$\{\boldsymbol{c}^i\}_g = \frac{k_B {T_0}^2 \mathrm{V} C_g}{\hbar^2 \omega_g} v_g \boldsymbol{s}_i \cdot \nabla T \tag{4.116}$$

Heat flux in a material is related to directional band-wise phonon energy by the following integral expression

$$\boldsymbol{Q} = \sum_g \frac{1}{4\pi} \int_{4\pi} v_g \boldsymbol{s}(\boldsymbol{\Omega}) e_g(\boldsymbol{\Omega}) d\boldsymbol{\Omega} \tag{4.117}$$

which in discretized form is

$$\boldsymbol{Q} = \sum_g \sum_i v_g w_i \boldsymbol{s}_i e_{gi} \tag{4.118}$$

This equation can also be written in the matrix vector form adopted above, which is

$$\boldsymbol{Q} = \sum_i w_i \boldsymbol{s}_i \boldsymbol{v}^T \boldsymbol{e}^i \tag{4.119}$$

where $\boldsymbol{v}$ is a column vector of group velocities given by

$$\{\boldsymbol{v}\}_g = v_g \tag{4.120}$$

An expression for energy can be found by solving the matrix form of the BTE (eq. 4.110) for the phonon energy $\boldsymbol{e}^i$ in each direction i

$$\boldsymbol{e}^i = \boldsymbol{A}^{-1}\left(\boldsymbol{b} - \boldsymbol{c}^i\right) \tag{4.121}$$

Substituting this into the heat flux expression, one can find

$$\boldsymbol{Q} = \sum_i w_i \boldsymbol{s}_i \boldsymbol{v}^T \boldsymbol{A}^{-1}\left(\boldsymbol{b} - \boldsymbol{c}^i\right) \tag{4.122}$$

$$\boldsymbol{Q} = \boldsymbol{v}^T \boldsymbol{A}^{-1}\left[\sum_i w_i \boldsymbol{s}_i \boldsymbol{b} - \sum_i w_i \boldsymbol{s}_i \boldsymbol{c}^i\right] \tag{4.123}$$

Now, the first term drops out, because $\sum_i w_i \boldsymbol{s}_i = 0$ (see section 4.5.2), but the second term gives the thermal conductivity.

$$\boldsymbol{Q} = -\boldsymbol{v}^T \boldsymbol{A}^{-1} \sum_i \boldsymbol{c}^i w_i \boldsymbol{s}_i = -k\nabla T \tag{4.124}$$

Returning to tensor form to extract the temperature gradient, one finds

$$\boldsymbol{Q} = -\frac{k_B {T_0}^2 \mathrm{V}}{\hbar^2} \sum_g v_g \sum_{g\prime} \frac{\{\boldsymbol{A^{-1}}\}_{gg\prime} C_{g\prime} v_{g\prime}}{\omega_{g\prime}} \sum_i w_i \boldsymbol{s}_i \boldsymbol{s}_i \nabla T \tag{4.125}$$

which, since $\sum_i w_i \boldsymbol{s}_i \boldsymbol{s}_i = \delta/3$, gives the thermal conductivity

$$k = \frac{1}{3} \frac{k_B {T_0}^2 \mathrm{V}}{\hbar^2} \sum_g v_g \sum_{g\prime} \frac{\{\boldsymbol{A^{-1}}\}_{gg\prime} C_{g\prime} v_{g\prime}}{\omega_{g\prime}} \tag{4.126}$$

The conductivity for the band-wise case is not difficult to find, since $\boldsymbol{A}$ is relatively small. Calculation of the full conductivity using the same expressions but with the full scattering matrix $P^{SS\prime}_{\boldsymbol{qq}\prime}$ is difficult, since the matrix $\boldsymbol{A}$ is so large. Thankfully, the band-to-band conductivity is not much different from the Single Mode Relaxation Time (SMRT) conductivity[80], so that can be used to calibrate the final parameter, group-wise

group velocity. Under a SMRT assumption, where the off diagonal terms in the scattering matrix are dropped, the full conductivity is given by

$$k = \frac{1}{3}\frac{k_B {T_0}^2 \mathrm{V}}{\hbar^2}\sum_{\boldsymbol{qs}}\frac{v_{\boldsymbol{qs}}^2 C_{\boldsymbol{qs}}^2}{\Gamma_{\boldsymbol{qs}}\omega_{\boldsymbol{qs}}^2} \tag{4.127}$$

which gives relaxation times in the group-wise case to be

$$\tau_g = \frac{k_B {T_0}^2 \mathrm{V} C_g}{\hbar^2 \Gamma_g \omega_g^2} \tag{4.128}$$

and in the full case to be

$$\tau_{\boldsymbol{qs}} = \frac{k_B {T_0}^2 \mathrm{V} C_{\boldsymbol{qs}}}{\hbar^2 \Gamma_{\boldsymbol{qs}} \omega_{\boldsymbol{qs}}^2} \tag{4.129}$$

since the full thermal conductivity in the SMRT case is given by

$$k = \frac{1}{3}\sum_{\boldsymbol{qs}} v_{\boldsymbol{qs}}^2 C_{\boldsymbol{qs}} \tau_{\boldsymbol{qs}} \tag{4.130}$$

Then it is possible to determine group-wise group velocities so that they conserve bulk thermal conductivity.

$$k_g = \frac{1}{3} v_g^2 C_g \tau_g = \frac{1}{3}\sum_{\boldsymbol{qs}\in g} v_{\boldsymbol{qs}}^2 C_{\boldsymbol{qs}} \tau_{\boldsymbol{qs}} \tag{4.131}$$

$$v_g^2 = \frac{1}{C_g \tau_g}\sum_{\boldsymbol{qs}\in g} v_{\boldsymbol{qs}}^2 C_{\boldsymbol{qs}} \tau_{\boldsymbol{qs}} \tag{4.132}$$

This approach allows accurate calculation of grouped phonon parameters that will closely represent the fully dispersive model.

4.5.1.9 Reduction to Single Mode Relaxation Time (SMRT) Model

The group-wise BTE can be reduced to the standard SMRT model with a relaxation time approximation. The group-wise BTE again is given by

$$-v_g \boldsymbol{s}(\boldsymbol{\Omega}) \cdot \nabla e_g(\boldsymbol{\Omega}) = \frac{\hbar^2 \omega_g}{4\pi k_B {T_0}^2 V}\sum_{g'} P_{gg'}\omega_{g'}\left(\frac{e_{g'}(\boldsymbol{\Omega})}{C_{g'}} - \frac{e}{C}\right) \tag{4.133}$$

In the SMRT, the off-diagonal terms are neglected, so that the scattering matrix is given by

$$P_{gg'} = \Gamma_g \delta_{gg'} \tag{4.134}$$

so that the BTE becomes

$$-v_g \boldsymbol{s}(\boldsymbol{\Omega}) \cdot \nabla e_g(\boldsymbol{\Omega}) = \frac{\hbar^2 \omega_g}{4\pi k_B {T_0}^2 V} \sum_{g'} \Gamma_g \delta_{gg'} \omega_{g'} \left(\frac{e_{g'}(\boldsymbol{\Omega})}{C_{g'}} - \frac{e}{C} \right) \tag{4.135}$$

$$-v_g \boldsymbol{s}(\boldsymbol{\Omega}) \cdot \nabla e_g(\boldsymbol{\Omega}) = \frac{\hbar^2 \omega_g}{4\pi k_B {T_0}^2 V} \Gamma_g \omega_g \left(\frac{e_g(\boldsymbol{\Omega})}{C_g} - \frac{e}{C} \right) \tag{4.136}$$

Now, substituting the relaxation time given by eq. 4.128 one finds

$$-v_g \boldsymbol{s}(\boldsymbol{\Omega}) \cdot \nabla e_g(\boldsymbol{\Omega}) = \frac{C_g}{\tau_g} \left(\frac{e_g(\boldsymbol{\Omega})}{C_g} - \frac{e}{C} \right) \tag{4.137}$$

Now, if an equilibrium energy for each band is defined by the ratio of that band's specific heat to the total specific heat times the total lattice energy e, as in

$$e_g^0 = \frac{C_g}{C} e \tag{4.138}$$

then the SMRT BTE is

$$-v_g \boldsymbol{s}(\boldsymbol{\Omega}) \cdot \nabla e_g(\boldsymbol{\Omega}) = \frac{e_g(\boldsymbol{\Omega}) - e_g^0}{\tau_g} \tag{4.139}$$

This model obviously will not conserve energy, since the scattering terms on the right hand side will not sum to zero for all energy distributions and so energy will be created or destroyed by each band, unless the relaxation times (and specific heats used to calculate e_g^0) of all the bands are exactly the same or the model has only one group and so is gray.

4.5.2 Quadrature Selection

As mentioned before, the DOM was first developed for solution of radiation heat transfer problems. There are several different ways to discretize the directional space and perform the approximation to the integral over the solid angle in equation 4.90.

4.5.2.1 Control Angle Quadratures

The first method, the Control Angle or Finite Volume Discrete Oridinates method (CADOM or FVDOM) which was discussed in Section 4.4.1 and is discussed by Chai et. al.[96] is especially flexible in that the resolution can be extended to reduce ray effects. CADOM is the method of choice of some groups specifically those solving the BTE in Silicon devices[89].

The difficulty with CADOM is that it does not strictly conserve the moments,

$$\sum_i w_i \boldsymbol{s}_i \boldsymbol{s}_i = \delta/3 \tag{4.140}$$

in the same way that other quadratures can. In the limit of an infinite number of directions, CADOM does conserve this expression, but for any practical number of discrete directions, the flux is not conserved. This can be an issue when comparing the solution to the an analytical solution and expecting the fluxes to match exactly. There are ways to correct the direction cosines or group velocities of the phonons in order to enforce matching with bulk conductivity, but the ballistic effects of which have been investigated in Section 4.5.8.1. It has been common in Lattice Boltzmann Method solvers to correct the thermal conductivity for the dimensionality of the problem, because the moments in eq. 4.140 are not conserved, and in fact the thermal conductivity of a LBM style model will be exactly $3/2$ times greater than the actual thermal conductivity as calculated from kinetic theory using the phonon properties of the LBM model. Heino uses a dimensionality factor on the thermal conductivity in order to match the LBM model to the bulk conductivity[97].

In a CADOM model, one correction proposed here is to adjust the lengths of the direction cosines. Considering that eq. 4.140 must be conserved in order to match the bulk conductivity, this is enforced by finding a factor α_β that makes

$$\sum_i w_i {a_\beta}^2 s_{\beta i} s_{\beta i} = 1/3 \tag{4.141}$$

true for the particular ordinate set, where β represents a dimension x,y, or z. Solving for α one finds

$$\alpha_\beta = \sqrt{\frac{1}{3\sum_i w_i s_{\beta i}^2}} \tag{4.142}$$

which can be used to correct the direction cosines so that fluxes and conductivity exactly match in the diffuse case. In the results below an analysis of this quadrature correction is made in the ballistic case against an analytical solution.

4.5.2.2 Sn Quadratures

Better than correcting cosines in the CADOM is finding a set of ordinates that matches these second moments in all cases. Such a quadrature sets have been investigated in literature extensively. Sn quadratures, used extensively in radiation heat transfer match these requirements and perform well compared to CADOM for comparable numbers of discrete directions for reducing ray effects[98]. The development of Sn quadratures is based on conservation of flux, moment, and various other rotation invariance conditions which are discussed by Lemonnier[99].

4.5.2.3 DCT and other more exotic quadratures

Other quadrature sets have been investigated which conserve higher order moments and have higher symmetry constraints and lower rotational invariance. Koch and Becker[100] evaluate several different quadrature schemes to determine which has the lowest error, namely Equal Weight Quadrature schemes[101], Lebedev's quadrature schemes[102], various geometrically based quadrature schemes, and Koch's DCT quadrature sets[103].

4.5.3 Finite Volume Equation Discretization

Using the finite volume method it is possible to discretize and solve the BTE. First, the domain is discretized into control volumes.

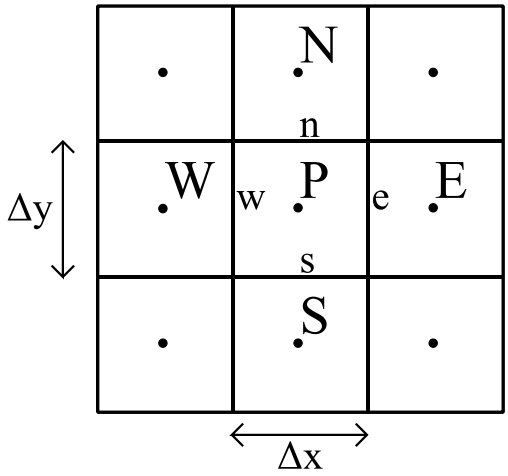

Figure 4.20. Discretization nomenclature for the FVDOM method here.

Then, the BTE is integrated over the control volume

$$-\int_{x_w}^{x_e}\int_{y_s}^{y_n} v_g \boldsymbol{s}_i \cdot \nabla e_{gi} dy\, dx = \int_{x_w}^{x_e}\int_{y_s}^{y_n} \sum_{g'}\sum_{i'} K_{gg'ii'} e_{g'i'}\, dy\, dx + \int_{x_w}^{x_e}\int_{y_s}^{y_n} G_g dy\, dx \tag{4.143}$$

$$-\int_{x_w}^{x_e}\int_{y_s}^{y_n} v_g \left(s_{ix}\frac{\partial e_{gi}}{\partial x} + s_{iy}\frac{\partial e_{gi}}{\partial y}\right) dy\, dx = \int_{x_w}^{x_e}\int_{y_s}^{y_n} \sum_{g'}\sum_{i'} K_{gg'ii'} e_{g'i'}\, dy\, dx + \int_{x_w}^{x_e}\int_{y_s}^{y_n} G_g dy\, dx \tag{4.144}$$

$$v_g s_{ix}\left(e_{gi}^{e} - e_{gi}^{w}\right)\Delta y + v_g s_{iy}\left(e_{gi}^{n} - e_{gi}^{s}\right)\Delta x + \sum_{g'}\sum_{i'} K_{gg'ii'} e_{g'i'}^{P}\,\Delta x\Delta y + G_g^P \Delta x\Delta y = 0 \tag{4.145}$$

Using upwinding, and dividing by $\Delta x \Delta y$ we can write the equation in the form

$$a_{gi}^{P} e_{gi}^{P} - a_{gi}^{E} e_{gi}^{E} - a_{gi}^{W} e_{gi}^{W} - a_{gi}^{N} e_{gi}^{N} - a_{gi}^{S} e_{gi}^{S} + \sum_{g'}\sum_{i'} K_{gg'ii'} e_{g'i'}^{P} + G_g^P = 0 \tag{4.146}$$

with

$$a_{gi}^{E} = \max\left(-v_g s_{ix}/\Delta x\,, 0\right) \tag{4.147}$$

$$a_{gi}^{W} = \max\left(v_g s_{ix}/\Delta x\,, 0\right) \tag{4.148}$$

$$a_{gi}^{N} = \max\left(-v_g s_{iy}/\Delta y\,, 0\right) \tag{4.149}$$

$$a_{gi}^{S} = \max\left(v_g s_{iy}/\Delta y\,, 0\right) \tag{4.150}$$

$$a_{gi}^{P} = \max\left(v_g s_{ix}/\Delta x\,, 0\right) + \max\left(-v_g s_{ix}/\Delta x\,, 0\right) + \max\left(v_g s_{iy}/\Delta y\,, 0\right) + \max\left(-v_g s_{iy}/\Delta y\,, 0\right) \tag{4.151}$$

In the QUICK scheme, a second order scheme developed by Leonard[104], the equation is derived in a similar way with the upwind face being defined by

$$e_{gi}^{e} = e_{gi}^{P} + \frac{e_{gi}^{E} - e_{gi}^{W}}{4} + \frac{e_{gi}^{E} + e_{gi}^{W} - 2e_{gi}^{P}}{8}$$

4.5.4 Boundary Conditions

4.5.4.1 Temperature Boundary Conditions

Temperature boundary conditions require that inward facing fluxes are equal to a given energy equivalent to the temperature. In general,

$$e_g(\boldsymbol{r_b}, \boldsymbol{\Omega})\big|_{\boldsymbol{\Omega}\cdot\mathbf{n}<0} = \int_{T_{ref}}^{T_b} C_g dT \tag{4.152}$$

In this model, the inward facing energy at a boundary face is given by

$$e_{gi}^{In} = C_g T_b \tag{4.153}$$

4.5.4.2 Diffuse Reflection Boundary Conditions

At a diffusively reflecting boundary the inward facing flux should be equal to the equilibrium energy at the boundary. This can be defined in several ways, but for the purposes of this model, the reflection occurs within each band, meaning the inward facing energy of a certain band is given by the integral over the solid angle of the energies within that band at the boundary. So at a boundary, energy is not scattered between energy bands. This reflection condition can be described by

$$e_g(\boldsymbol{r}_b, \boldsymbol{\Omega})\big|_{\boldsymbol{\Omega}\cdot\mathbf{n}<0} = \frac{1}{4\pi}\int_{4\pi} e_g(\boldsymbol{r}_b, \boldsymbol{\Omega}')\, d\boldsymbol{\Omega}' \tag{4.154}$$

So, in discrete terms, the inward facing boundary flux is given by

$$e_{gi}^{in} = \sum_{i\prime} w_{i\prime} e_{gi\prime} \tag{4.155}$$

Now, for UDS, at the boundary, the outward facing contribution to this sum can be given by upwinding, while the inward facing contribution is given by the reflection condition, so we have

$$e_{gi}^{in} = \sum_{i\prime} w_{i\prime} e_{gi\prime}^{P}\, a_{i\prime} + \sum_{i\prime} w_{i\prime} e_{gi}^{in}\,(1-\alpha_{i\prime}) \tag{4.156}$$

where $\alpha_i = 1$ if $\boldsymbol{s}_i \cdot \boldsymbol{n} < 0$, otherwise, $\alpha_i = 0$. It is now possible to solve for e_{gi}^{in},

$$e_{gi}^{in} = \frac{\sum_{i\prime} w_{i\prime} e_{gi\prime}^{P}\, \alpha_{i\prime}}{1 - \sum_{i\prime} w_{i\prime}\,(1-\alpha_{i\prime})} \tag{4.157}$$

For QUICK scheme reflection we have a similar sum,

$$e_{gi}^{in} = \sum_{i\prime} w_{i\prime} e_{gi\prime}^{out}\, a_{i\prime} + \sum_{i\prime} w_{i\prime} e_{gi}^{in}\,(1-\alpha_{i\prime}) \tag{4.158}$$

except that the outward facing contribution needs to be calculated from the quadratic upwinding method, so that $e_{gi\prime}^{out}$ is given by

$$e_{gi\prime}^{out} = e_{gi\prime}^{P} + \frac{e_{gi\prime}^{outside} - e_{gi\prime}^{inner}}{4} + \frac{e_{gi\prime\prime}^{P} + e_{gi\prime}^{inner} - 2e_{gi\prime}^{P}}{8} = \frac{3}{4}e_{gi\prime}^{P} + \frac{3}{8}e_{gi\prime\prime}^{P} - \frac{1}{8}e_{gi\prime}^{inner} \tag{4.159}$$

where $e_{gi\prime}^{inner}$ is the second cell from the wall and i'' is the direction that corresponds to the direction i' mirrored about the reflecting face. This is a partially specular reflection scheme because it includes some reflection from the same direction. To enforce full diffusivity would also require that the reflection portion $e_{gi\prime\prime}^{P}$ is also an average of all directions. However, this assumption improves the simplicity of the model for what is expected to be very little difference in the final result. This gives

$$e_{gi}^{in} = \sum_{i'} w_{i'} \left(\frac{3}{4} e_{gi'}^{P} + \frac{3}{8} e_{gi''}^{P} - \frac{1}{8} e_{gi'}^{inner} \right) a_{i'} + \sum_{i'} w_{i'} e_{gi}^{in} \left(1 - \alpha_{i'}\right) \tag{4.160}$$

and solving for e_{gi}^{in} we get

$$e_{gi}^{in} = \frac{\sum_{i'} w_{i'} \left(\frac{3}{4} e_{gi'}^{P} + \frac{3}{8} e_{gi''}^{P} - \frac{1}{8} e_{gi'}^{inner} \right) \alpha_{i'}}{1 - \sum_{i'} w_{i'} \left(1 - \alpha_{i'}\right)} \tag{4.161}$$

This term contributes in two places: the equation for the boundary cell, and the equation for the second cell from the wall.

In effect, these reflection equations calculate the inward facing energies by finding the average of the outward facing energies in all directions at the given point and band.

4.5.5 Solution Methods

4.5.5.1 Explicit Methods

The simplest solution method is to rearrange the discrete equation and sweep over the domain solving for energies in a sequential way, using a Gauss-Seidel method

$$e_{gi}^{P} = \frac{a_{gi}^{E} e_{gi}^{E} + a_{gi}^{W} e_{gi}^{W} + a_{gi}^{N} e_{gi}^{N} + a_{gi}^{S} e_{gi}^{S} - \sum_{g'} \sum_{i'} K_{gg'ii'} e_{g'i'}^{P} + b}{a_{gi}^{P}} \tag{4.162}$$

Some of the neighbor coefficients will be zero, which can be exploited when determining sweeping direction for a given equation. It should also be possible to rearrange the equations into a Tri-diagonal Matrix Algorithm. For details see [105].

4.5.5.2 Fully Implicit Methods

It is possible to form a fully implicit form of the equations. Form sparse diagonal matrices

$$\boldsymbol{A}_{gi} = \begin{bmatrix} a_{gi}^{P} + K_{ggii} & -a_{gi}^{E} & \dots & -a_{gi}^{S} & & & & \\ -a_{gi}^{W} & a_{gi}^{P} + K_{ggii} & & & -a_{gi}^{S} & & & \\ \vdots & & \ddots & & \vdots & & & \\ -a_{gi}^{N} & & & a_{gi}^{P} + K_{ggii} & -a_{gi}^{E} & & & \\ & -a_{gi}^{N} & \dots & -a_{gi}^{W} & a_{gi}^{P} + K_{ggii} & & & \\ & & & & & \ddots & & \\ & & & & & & a_{gi}^{P} + K_{ggii} & -a_{gi}^{E} \\ & & & & & & -a_{gi}^{W} & a_{gi}^{P} + K_{ggii} \end{bmatrix} \tag{4.163}$$

and form diagonal matrices

$$\boldsymbol{K}_{gg'ii'} = \begin{bmatrix} K_{gg'ii'} & & \\ & \ddots & \\ & & K_{gg'ii'} \end{bmatrix} \tag{4.164}$$

Next, form the full coefficient matrix

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{A}_{1,1} & \boldsymbol{K}_{1,1,1,2} & \cdots & \boldsymbol{K}_{1,1,N_{bands},N_{dir}} \\ \boldsymbol{K}_{1,2,1,1} & \boldsymbol{A}_{1,2} & & \boldsymbol{K}_{1,2,N_{bands},N_{dir}} \\ \vdots & & \ddots & \vdots \\ \boldsymbol{K}_{N_{bands},N_{dir},1,1} & \boldsymbol{K}_{N_{bands},N_{dir},1,2} & \cdots & \boldsymbol{A}_{N_{bands},N_{dir}} \end{bmatrix} \tag{4.165}$$

With the full energy vector

$$\boldsymbol{e} = \begin{bmatrix} e_{1,1}^{1,1} \\ e_{1,1}^{1,2} \\ \vdots \\ e_{1,1}^{N_y,N_x} \\ e_{1,2}^{1,1} \\ \vdots \\ e_{N_{bands},N_{dir}}^{N_y,N_x} \end{bmatrix} = \begin{bmatrix} \boldsymbol{e}_{1,1} \\ \boldsymbol{e}_{1,2} \\ \vdots \\ \boldsymbol{e}_{N_{bands},N_{dir}} \end{bmatrix} \tag{4.166}$$

Finally, the right hand side is formed by taking into account boundary conditions, so that the full matrix equation is given by

$$\boldsymbol{A}\boldsymbol{e} = \boldsymbol{b} \tag{4.167}$$

This equation quickly becomes intractable, even though $\boldsymbol{A}$ is very sparse. Consider using S_n quadratures with $n = 3$, yielding 12 directions, and using a 9 band frequency discretization, with 100x100 cells, the system has 12x9x100x100 or 1,080,000

degrees of freedom. The full $\boldsymbol{A}$ matrix will have approximately 120,960,000 nonzero elements, which equates to 1,845 MB of storage in double precision. On some smaller systems it is possible to form the entire $\boldsymbol{A}$ matrix and solve using direct or iterative methods. The GMRES iterative solver is one option for 1-D systems[106]. Preconditioning is achieved by performing incomplete LU decomposition on the spatial portion of the $\boldsymbol{A}_{gi}$ matrix in each direction and band,

$$\boldsymbol{L}_{gi}\boldsymbol{U}_{gi} = \boldsymbol{A}_{gi} \tag{4.168}$$

so that the preconditioning matrices are

$$\boldsymbol{M}_1 = \begin{bmatrix} \boldsymbol{L}_{1,1} & \cdots & \\ \vdots & \ddots & \vdots \\ & \cdots & \boldsymbol{L}_{N_{bands},N_{dir}} \end{bmatrix} \tag{4.169}$$

and

$$\boldsymbol{M}_2 = \begin{bmatrix} \boldsymbol{U}_{1,1} & \cdots & \\ \vdots & \ddots & \vdots \\ & \cdots & \boldsymbol{U}_{N_{bands},N_{dir}} \end{bmatrix} \tag{4.170}$$

Then the preconditioned system becomes

$$(\boldsymbol{M}_1\boldsymbol{M}_2)^{-1}\boldsymbol{A}\boldsymbol{e} = (\boldsymbol{M}_1\boldsymbol{M}_2)^{-1}\boldsymbol{b} \tag{4.171}$$

or

$$\boldsymbol{M}_2^{-1}\boldsymbol{M}_1^{-1}\boldsymbol{A}\boldsymbol{e} = \boldsymbol{M}_2^{-1}\boldsymbol{M}_1^{-1}\boldsymbol{b} \tag{4.172}$$

There are alternate ways to form the full matrix system, and convergence of them has not been studied closely, but this method works quite well for small systems.

There are likely other ways to use iterative solvers to solve these equations implicitly without forming the entire $\boldsymbol{A}$ matrix, which have also not been studied but are likely to be fairly quick if implemented in a low level code. For instance, implementing GMRES on a function $f(\boldsymbol{e}) = \boldsymbol{A}\boldsymbol{e}$.

4.5.5.3 Semi-Implicit Methods

Semi-implicit methods appear to be the fastest for larger systems, although convergence becomes poor for highly diffusive problems. Essentially, instead of solving the entire system in one step, each direction and band are solved sequentially until convergence. This involves moving the out-of-band scattering terms to the right hand side, which gives equations

$$\boldsymbol{A}_{gi}\boldsymbol{e}_{gi} = -\sum_{g'}\sum_{i'} K_{ggii'}\,\boldsymbol{e}_{g'i'}\left(1-\delta_{gg'}\delta_{ii'}\right) + \boldsymbol{b}_{gi} \tag{4.173}$$

Now again it is possible to use LU decomposition on the individual $\boldsymbol{A}_{gi}$ matrices before the main iteration in order to speed up the solution process. The solver finds

$$\boldsymbol{e}_{gi} = \boldsymbol{U}_{gi}^{-1}\boldsymbol{L}_{gi}^{-1}\left(-\sum_{g'}\sum_{i'} K_{ggii'}\,\boldsymbol{e}_{g'i'}\left(1-\delta_{gg'}\delta_{ii'}\right) + \boldsymbol{b}_{gi}\right) \tag{4.174}$$

repeatedly until the residual of the full system is reduced below a relative tolerance, i.e. we have for the full system

$$\|\boldsymbol{A}\boldsymbol{e}-\boldsymbol{b}\| < \tau_r\|\boldsymbol{b}\| \tag{4.175}$$

where τ_r is the relative tolerance of the solution.

If the complete LU decomposition is too memory intensive to perform, a partial LU decomposition can be performed and then a preconditioned GMRES or another iterative solver can be used to find the solution to eq. 4.173.

4.5.6 Implementing Boundary Conditions

4.5.6.1 Temperature B.C.s

Temperature boundary conditions are extremely straightforward to implement in any of these solver schemes. At a boundary cell, instead of a contribution to the $\boldsymbol{A}_{gi}$ matrix, the coefficient will contribute to $\boldsymbol{b}_{gi}$ by the amount $a_{gi}e_{gi}^{b}$, where a_{gi} is the

appropriate coefficient. For the QUICK scheme the contribution of second neighbors also has to be considered.

4.5.6.2 Reflection B.C.s

Reflection boundary conditions are slightly more difficult. The boundary contribution must be calculated but it is dependent on the outgoing fluxes since they are being reflected. For a single boundary with a reflection condition, it is possible to form a matrix expression where the contribution from reflection to the right hand side is given by

$$\boldsymbol{b}_{refl,gi} = B_{gi}^{(1)} \boldsymbol{M}_r \sum_{i\prime} B_{i\prime}^{(2)} \boldsymbol{e}_{gi} \tag{4.176}$$

where $B_{gi}^{(1)}$ is a scalar which is given by

$$B_{gi}^{(1)} = \frac{a_{gi}}{1 - \sum_{i\prime} w_{i\prime} (1 - \alpha_{i\prime})} \tag{4.177}$$

where a_{gi} is the appropriate boundary coefficient for this boundary cell. $\boldsymbol{M}_r$ is a diagonal matrix given by the appropriate sparsity pattern for the boundary, i.e., a 1 in the position for cells that contain the reflection condition, and a 0 for cells that do not. Finally, $B_{i\prime}^{(2)}$ is a scalar given by

$$B_{i\prime}^{(2)} = w_{i\prime} \alpha_{i\prime} \tag{4.178}$$

where again $\alpha_i = 1$ if $\boldsymbol{s}_i \cdot \boldsymbol{n} < 0$, otherwise, $\alpha_i = 0$.

The QUICK scheme reflection implementation is similar except that two summations are required to include the effect of the second neighbor.

For more general reflection boundary conditions with multiple boundaries that reflect, a full reflection matrix $\boldsymbol{R}_{gi}$ is constructed which gives the contribution from reflection on the right hand side in band gi as

$$\boldsymbol{b}_{refl,gi} = \boldsymbol{R}_{gi}\boldsymbol{e} \tag{4.179}$$

however forming a full matrix in this sense is expensive in terms of memory and computational power, so the first formulation is favorable in the case of one wall that is reflective.

4.5.7 Verification

After development of the model some verification is in order to ensure that it is producing reasonable results.

4.5.7.1 One dimensional Grey Transport

Firstly the one-dimensional transport is compared with the analytical thermal transport solution in a one dimensional plane wall case. The analytical solution is taken from radiation heat transport in the work of Heaslet and Warming[107]. Here the left and right boundaries are set to temperatures of 301 K and 300 K respectively. For the DOM solution the domain was split into 500 finite volume cells and Sn quadratures with $n = 8$ and the one band (grey) model was used. Good agreement is seen between the DOM model and the analytical solution.

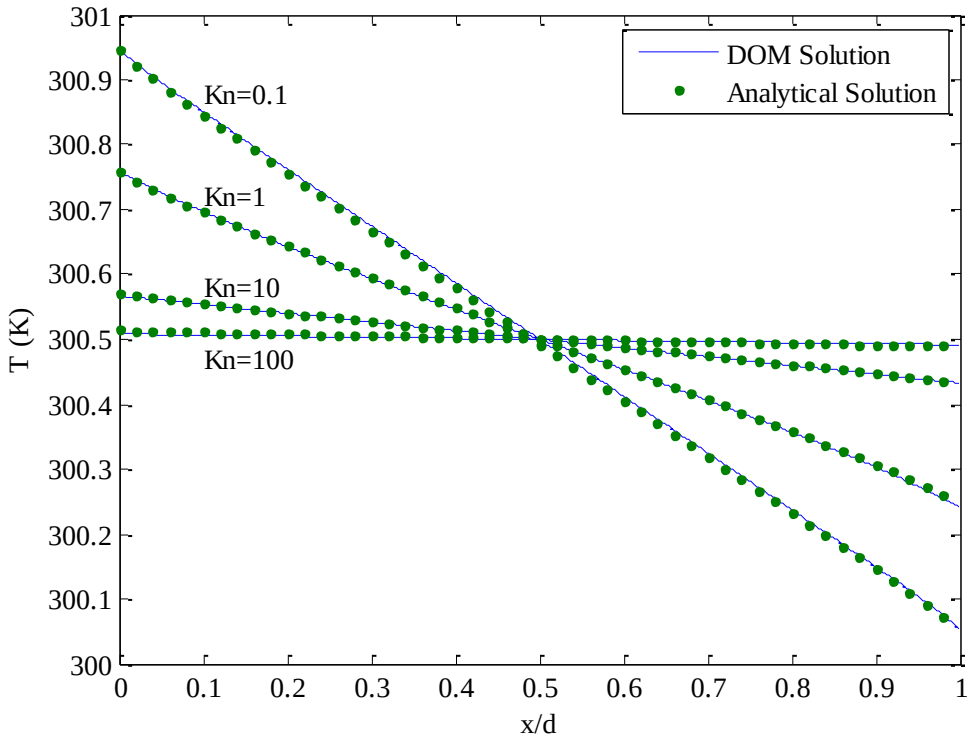

Figure 4.21. Comparison of DOM solution to Analytical solution. Kn between 0.1 and 100 are compared.

Next the transport is verified in the diffuse case against a one dimensional Fourier model. In the diffuse case the Knudsen number is well below 0.01. There are a number of variables which affect how closely the model can match the diffuse result, but it is expected as the Knudsen number approaches 0 the DOM solution approaches the diffuse heat flux. Other variables determining how closely the diffuse solution is matched are the solver tolerance and the number of finite volume cells used for the simulation. Here the solver tolerance was selected to be 10^{-10} and the solution used 1000 cells with Sn quadratures with $n = 3$. Increasing the tolerance and number of cells both have an effect on how closely the solution matches to the diffuse result. Below the difference between a the diffusive solution (using Fourier's law) and the DOM solution for decreasing (more diffusive) Knudsen numbers is shown in Figure 4.22. The error is given by the ratio of the norm of the difference between the heat flux of the DOM solution and the Fourier solution and the norm of the Fourier solution,

$$\epsilon = \frac{\sqrt{\sum_i \left(Q_{DOM}(x_i) - Q_{Fourier}(x_i)\right)^2}}{\sqrt{\sum_i Q_{Fourier}(x_i)^2}} \tag{4.180}$$

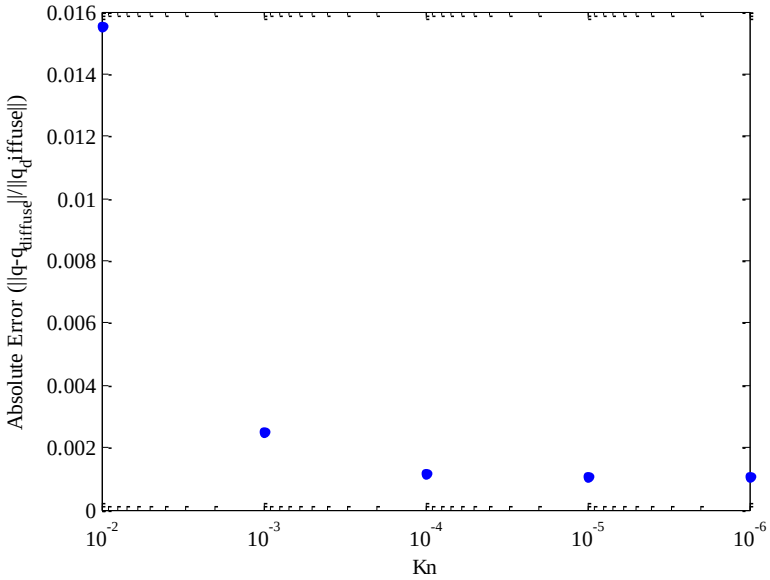

Figure 4.22. The absolute error between the DOM and the Fourier solution for decreasing Knudsen numbers (more diffusive regimes). As the Knudsen number approaches zero the error should decrease to zero.

There appears to be some flattening out of the error which is a result of the solver tolerance and the number of cells selected for the simulation which give somewhat of an upper bound to the accuracy of the solution.

For the selection of different ordinate sets which do not conserve the second moment of flux, the flux will never be exactly equal to the Fourier solution for any finite number of discrete ordinates.

4.5.7.2 Two dimensional Grey Transport

While the ballistic nature of two dimensional transport is hard to verify since there are no analytical solutions, it is possible to compare the diffusive regime to a Fourier solution and confirm that the model is matching transport at large scales. The two dimensional case analyzed in Section 4.4.6 is looked at again here with the same

conditions. The equivalent temperature for a Knudsen number of 10^{-3} and a 100×100 mesh is shown in Figure 4.23. In this case, Sn quadratures were again used with $n = 3$. The difference between the Fourier and DOM solution is not discernable from the contour plots.

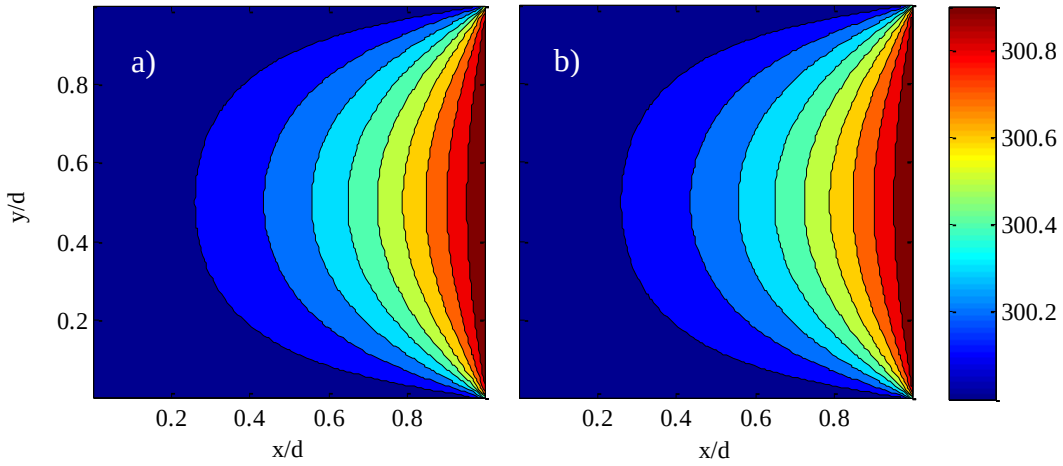

Figure 4.23. The DOM solution (a) is compared to a Fourier solution (b) in the case of Knudsen number 10^{-3} and a 100×100 mesh. The difference is not discernable here.

In order to take a closer look at how the DOM is matching the Fourier solution for low Knudsen numbers, Figure 4.24 shows the temperature and heat flux for the Fourier and DOM models directly compared and again shows excellent agreement. The solution time for these models was on the order of 15 seconds.

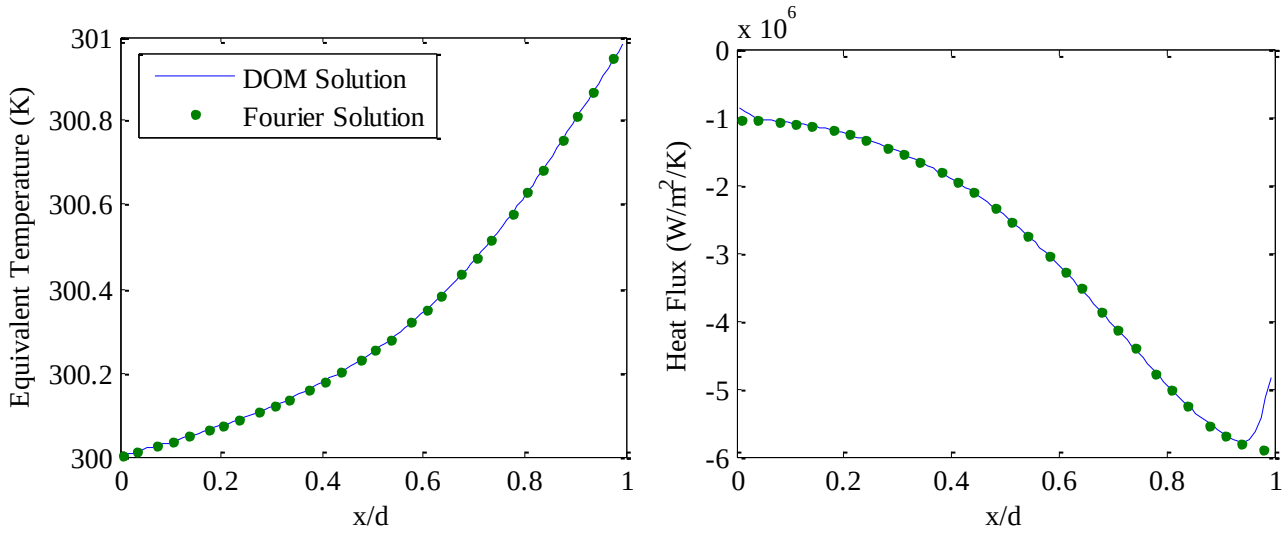

Figure 4.24. The Temperature (on left) and heat flux in the x-direction (on right) are shown as compared to the Fourier solution (green dots). Excellent agreement is seen except at the boundaries in the heat flux.

4.5.7.3 Boundary Scattering Reduction in Thermal Conductivity

The reduction in thermal conductivity is another good test for the performance of reflective boundary conditions, as seen in Section 4.4.6.4. The reduction in thermal conductivity for three different Sn quadrature sets $n = 3,6 \text{ and } 8$ are shown in Figure 4.25 and compared to an analytical solution by Flik et. al.[92]. The mesh was 20×400 cells and the solver tolerance was set to 10^{-5} which was determined to be mesh independent (increasing mesh to 30×600 and tolerance to 10^{-6} made no discernable difference.

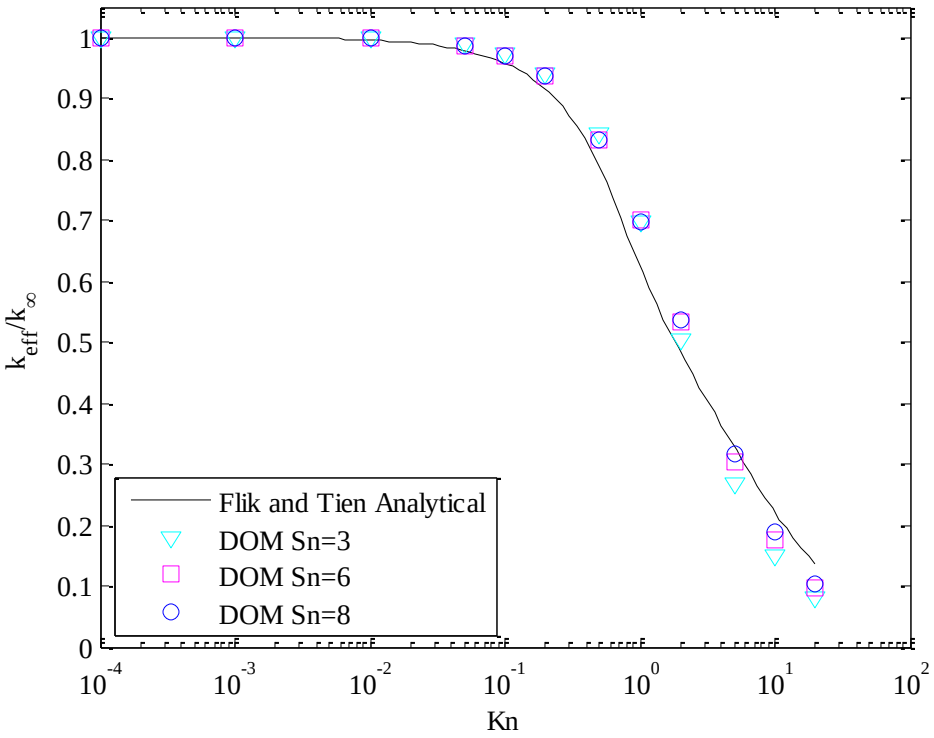

Figure 4.25. Reduction in thermal conductivity as a function of Knudsen number for a thin film. Sn quadratures with $n = 3,6$ and 8 are shown.

It seems that for low Knudsen numbers the conductivity reduction is slightly underestimated (conductivity is higher) while for higher Knudsen numbers the reduction is overestimated (conductivity is lower) by the DOM model presented here. It does appear that increasing the resolution of the quadrature set is slowly pushing the solution towards Flik and Tien's solution.

4.5.7.4 One Dimensional Grey Generation

In order to ensure that the model can capture heat generation accurately a constant heat generation region is inserted into a one dimensional plane wall model and the fluxes are compared for a single band case and for the multiple band case. Initially a UDS scheme was tested, but it was determined that for reasonable grid spacing UDS did not conserve flux, as can be seen in Figure 4.27, which prevents it from converging to the Fourier temperature distribution in the diffuse limit (see Figure 4.26). For more ballistic

solutions the UDS scheme becomes more conservative, but in a semi-diffuse regime the UDS solution is not accurate when generation is present in a domain.

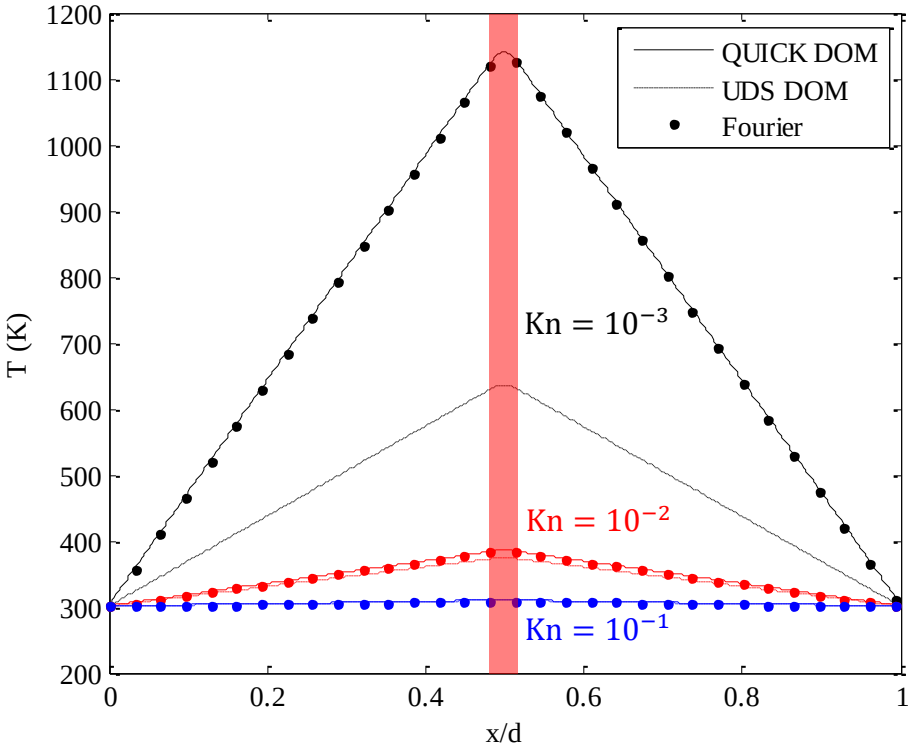

Figure 4.26. Temperature profiles for the one dimensional test case for different Knudsen numbers with a constant small generation region of 10^{10} W/m^2. Note that for smaller Knudsen numbers (more diffusive transport) the UDS scheme deviates significantly from the QUICK scheme DOM solution and the Fourier solution. It is expected that the ballistic DOM solution gives rise to higher temperatures than the Fourier solution due to temperature slip.

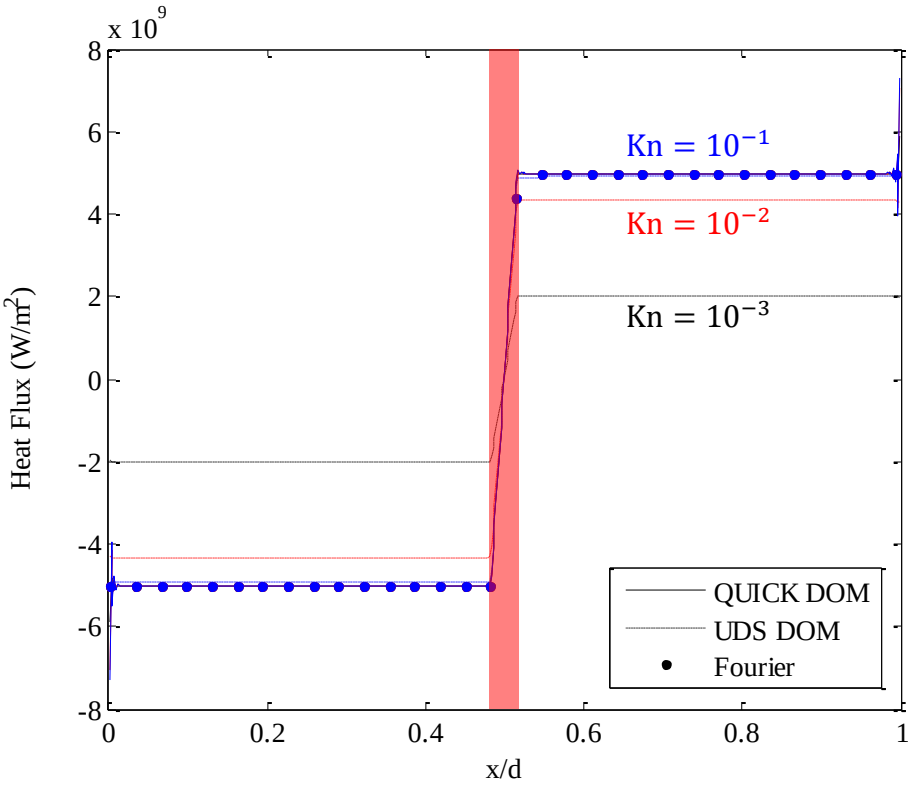

Figure 4.27. Comparison of heat flux profiles for the one dimensional test case for different Knudsen numbers with a small heat generation region of 10^{10} W/m^2. The QUICK solution is completely collapsed onto the Fourier solution for all Kn which is what one expects since the energy being generated in the domain must be carried out. The QUICK scheme heat flux does have oscillations at the boundaries and near the edges of the heat flux region.

As good as the QUICK scheme is at ensuring conservation for reasonable grid spacing, it does have oscillations at the boundaries and near discontinuities in the underlying heat generation. This kind of oscillation could be resolved with a third order SMART scheme or other flux limiting scheme that prevents local minima and maxima from being generated near discontinuities[28, 108]. However, implementation of that kind of scheme would require restructuring the solver significantly and was not attempted at this time.

4.5.7.5 Phonon Properties and Relaxation Times

The phonon relaxation times can be calculated from the diagonal portion of the scattering matrix by using equation 4.129. The relaxation times as shown in Table 4.1

are within the realistic range of parameters, except for the E_1(TO) and E_1(LO) modes, where the relaxation time is being overestimated significantly.

Table 4.1. Relaxation Times for Γ-point modes

Mode	ω (cm^{-1})	τ (ps), this work	τ (ps), literature
E_2(low)	137	36.1	10.1 [50], 70 [109]
	140	37.0	
B (low)	329	3.6	Raman inactive
A_1(TO)	540	1.9	0.46 [50], 0.63 [110], 5.1 [95]
E_1(TO)	556	3.9	0.94[95], 0.95 [50]
E_2(high)	560	4.2	1.4 [50], 2.5 [95], 2.56 [110], 5.8 [109]
B (high)	564	4.5	Raman inactive
A_1(LO)	686	2.5	0.75 [95], 0.8, 0.86 [110], 1.5 [109], 0.35-2.5 [94]
E_1(LO)	714	1.2	0.56 [95], 0.66 [110]

4.5.8 Results

4.5.8.1 One Dimensional Analysis of Ordinates Set Selection

Ordinate selection plays a role in the accuracy of ballistic and transient transport in the DOM. We focus here on the ballistic portion of transport, and compare various ordinate sets to an analytical solution provided by Heaslet and Warming[107]. The solution of three ballistic transport cases, Kn=1,10, and 100 were compared for a number of ordinate sets. The total squared error was calculated for each solution, and the results are plotted in Figure 4.28. As expected the error decreases with an increase in the number of directions. Probably the most interesting result here is that the corrected ordinate set performed as well as the uncorrected ordinate set, suggesting that there is not a penalty in correcting the CADOM quadratures with the method presented in Section 4.5.2.1.

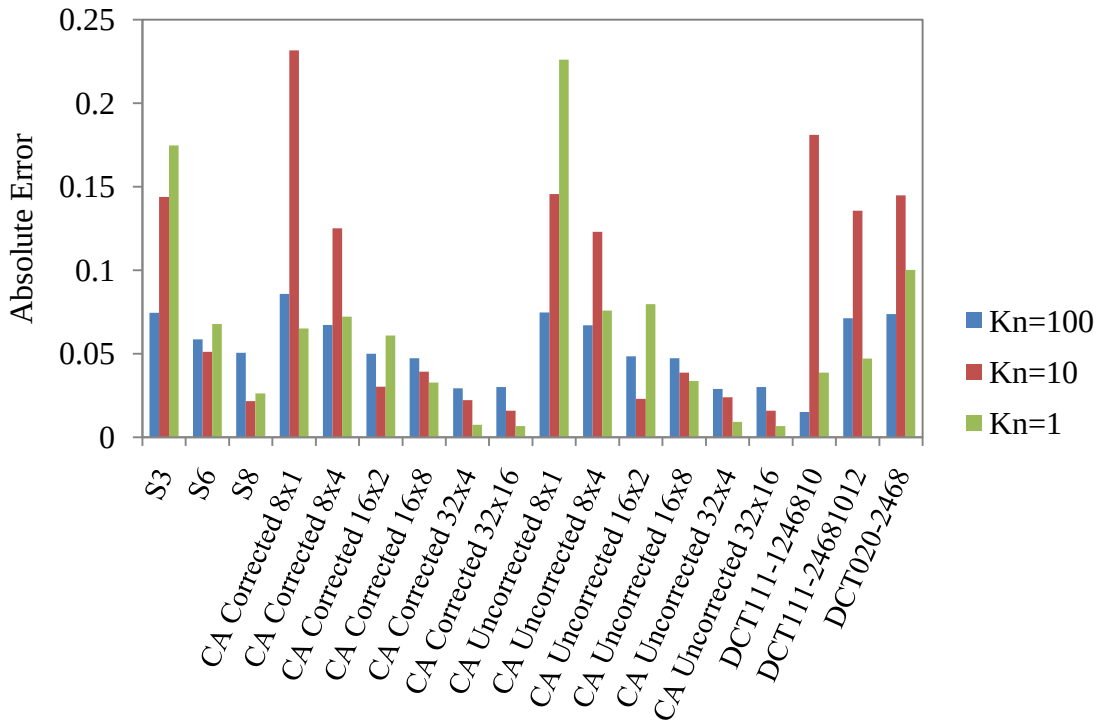

Figure 4.28. Ballistic error for Knudsen numbers 1, 10 and 100 for various quadrature sets. Three of the best performing DCT quadratures were selected.

4.5.8.2 Band-to-Band HFET Modeling

A limited test of a AlGaN/GaN HFET is executed. The domain modeled is much smaller and more simple than the domain in an actual device, and is shown in Figure 4.29. The domain was split into a mesh 100x143 cells maintaining a unity aspect ratio. The model used the 8 band discretization of the phonon spectrum illustrated in Figure 4.11. This 8 band discretization resulted in the phonon properties displayed in Table 4.2.

Table 4.2. Phonon properties for the 8 band model

$\omega_g(cm^{-1})$	v_g(m/s)	C_g(J/cm^3/K)
70	5420	0.059
160	4051	0.976
237	4265	0.264
305	2548	0.366
573	1972	0.331
612	4623	0.308
671	1036	0.102
694	1054	0.166

The generation region as calculated for an AlGaN/GaN HFET from Sentaurus for V_{ds}=14V and V_{gs}=0V was input into the model's highest energy band where all the longitudinal optical phonons lie. This is the same bias condition as modeled with the continuum scale model with more realistic boundary conditions in Section 3.7, except that in this case only the generation within the GaN region was captured. The AlGaN region was not modeled here since the phonon properties of AlGaN have not been accurately studied and the physics of thermal transport across the AlGaN/GaN interface are not yet captured by this model. S_n quadratures were used with n=3 resulting in a total of 1,372,800 degrees of freedom. The solution was performed with the semi-implicit solver and took approximately 6 hours to complete.

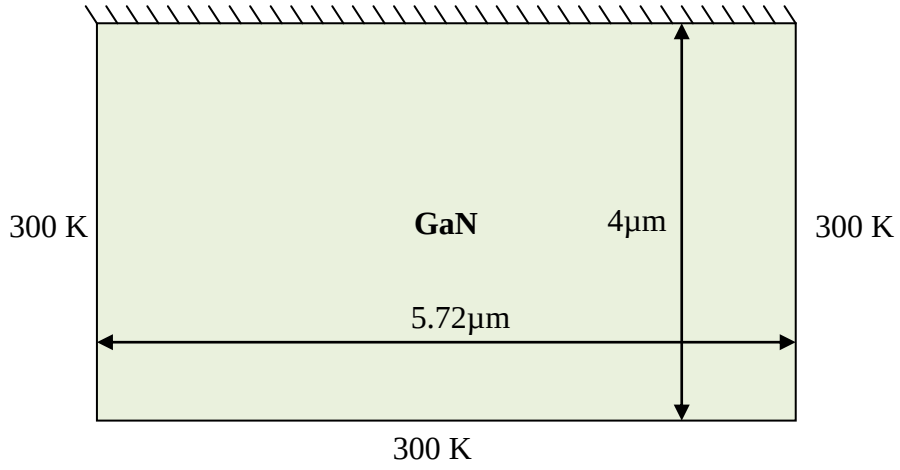

Figure 4.29. Simplified device domain run in the band to band phonon model. The top boundary is insulated while the other three boundaries are set to 300 K. This domain is much smaller than an actual device but includes the entire active area.

The case run is approximately 8.5 W/mm, putting it on the high end of the power spectrum for AlGaN/GaN HFET devices. The resulting equivalent temperature distribution is shown as a contour plot in Figure 4.30. The peak temperature occurs on the drain side of the gate next to the AlGaN/GaN boundary where the peak generation region is. This temperature profile is significantly different from a Fourier style solution in this domain, as shown in Figure 4.31. A horizontal cut along the top of the domain of the equivalent Band to Band temperature and the Fourier temperature is shown in Figure 4.32a, while a vertical cut is shown in Figure 4.32b. Finally, the individual band temperatures are plotted with the equivalent total temperature and Fourier temperature in Figure 4.33 for the horizontal cut line across the top of the domain.

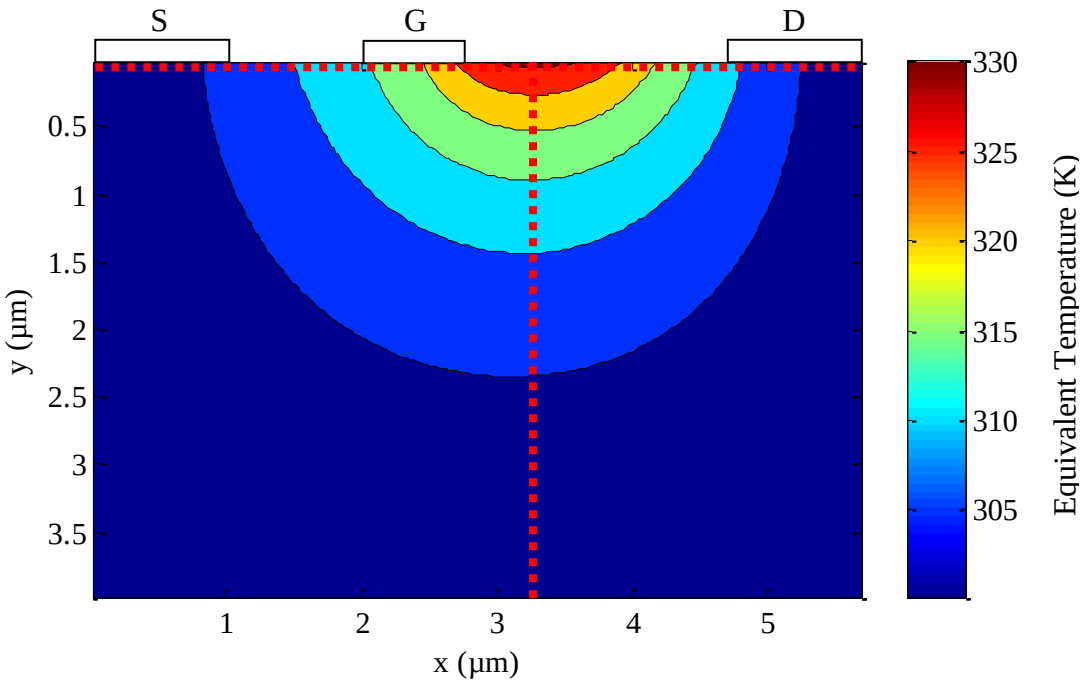

Figure 4.30. Contour plot of the equivalent temperature within the device. Source (S), Gate (G), and Drain (D) structures are shown for reference. The peak temperature occurs in the channel on the drain side of the gate. The horizontal and vertical red dashed lines indicate the locations of the cuts for Figure 4.32.

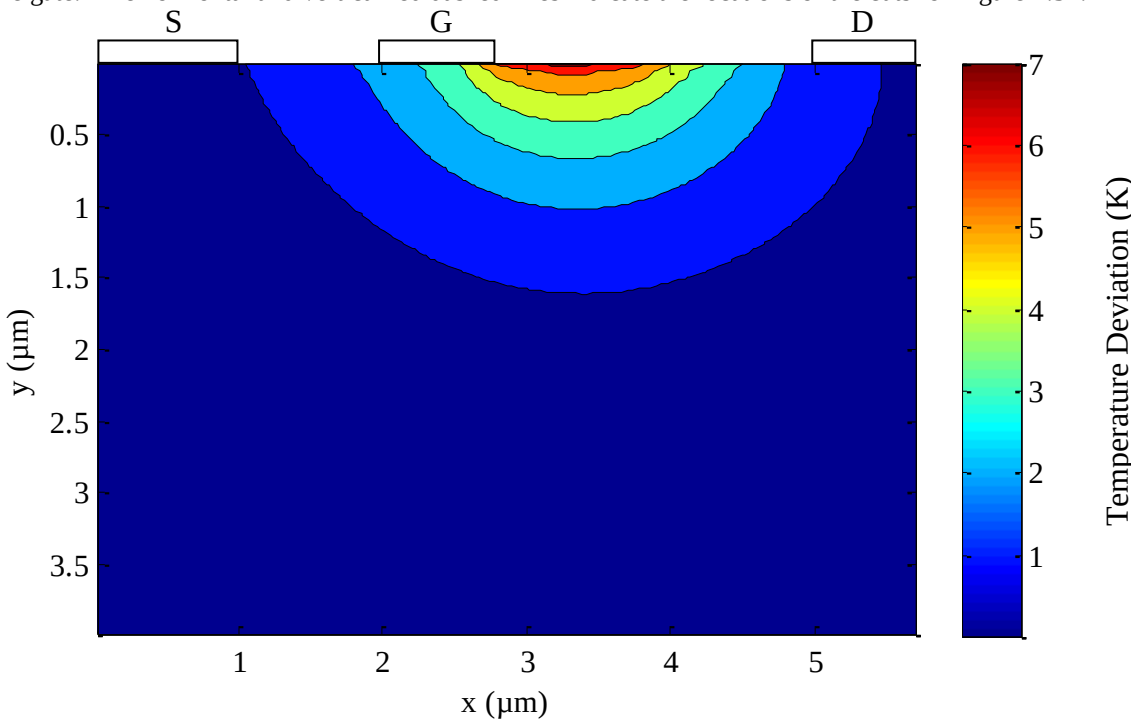

Figure 4.31. Contour plot of the temperature deviation from a traditional Fourier model.

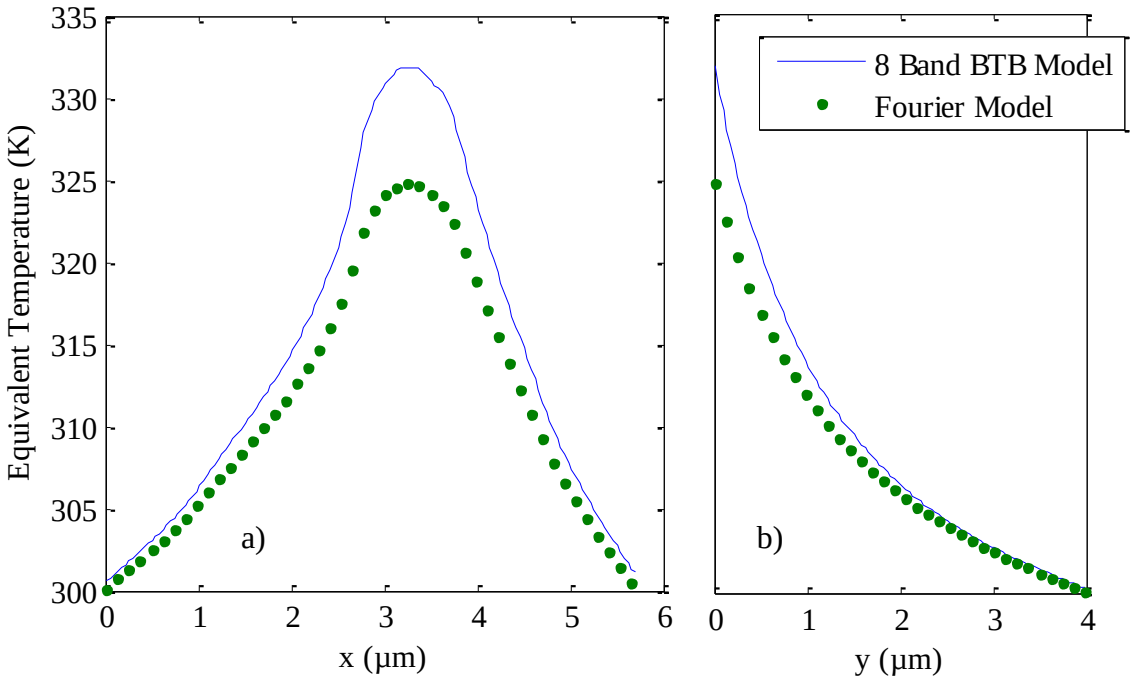

Figure 4.32. Comparison of a the equivalent temperatures in a band to band model and a Fourier model shown along a) a horizontal line and b) a vertical cut going through the hot spot as shown by dashed red lines in Figure 4.30.

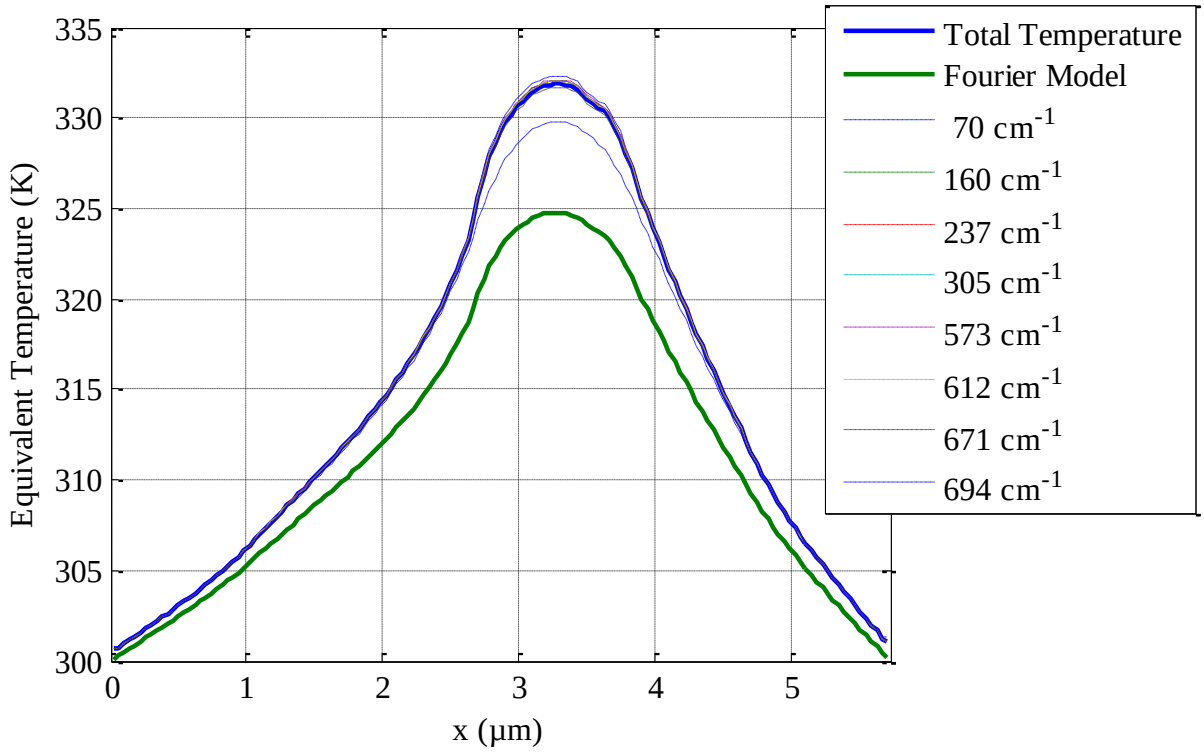

Figure 4.33. Comparison of phonon temperatures within the 8 band BTB Model.

The Band to Band model predicts an equivalent lattice temperature approximately 7 K hotter than the Fourier model in this case, which was a domain of limited size, which suggests temperature deviation will be less than in a real domain. Computational methods need to be improved before larger domains can be run, which will give a more accurate representation of the deviation from the Fourier model. Interestingly, the high energy optical phonon modes contained within the 694 cm^{-1} band do not deviate a large amount above the lattice temperature. The reason for this is unknown since some literature suggests optical phonons should be on the order of 1000 K[10]. There are a few possibilities. The band resolution may need to be improved beyond an 8 band model to further refine the transmission of phonons from the high energy optical phonons. For this to happen, more detailed information about the interaction between the electrons and phonons will be required as an input, since at this point energy is just assumed to be dumped entirely into the highest energy band. Another possibility for this deviation is

the process used in correcting the interaction matrix, which might destroy some important information about the band-to-band transitions. It might be necessary to implement information about the third derivative of energy in order to accurately capture the interaction strengths and bring the raw scattering matrix into better compliance with the conservation and physicality conditions.

CHAPTER 5

SUMMARY AND FUTURE WORK

This work investigates temperature and stress effects within AlGaN/GaN HEMT and MOSHEMTs during operation in order to better understand the lack of reliability currently seen in this material system. AlGaN/GaN HEMT and MOSHEMTs are currently limited by their reliability. The causes of degradation are still under investigation. It is believed that peak temperatures, non-equilibrium effects, ballistic effects, CTE mismatch stress, thermal gradient stress, epitaxial residual stress, and inverse-piezoelectric stresses all play a role in degradation. This complicated system is studied in some detail in this work, specifically peak temperatures as a result of changing generation region size and shape, thermal stresses as a result of CTE mismatch and thermal gradients, and ballistic and non-equilibrium effects.

Key contributions of this work include the following:

- The effect of including an accurate heat generation region extracted from an electrical device simulator was studied in detail. This avoids the issue of having to fit the size of the generation region to experimental peak temperature data and can offer insight onto why certain bias conditions yield higher temperatures than others at the same power conditions: specifically, higher pinch-off conditions imply a smaller generation region near the drain side of the gate and higher peak temperatures as a result.
- It was determined that using temperature dependent thermal conductivities in GaN modeling can significantly affect peak temperatures, especially at higher powers.
- Thermal gradient stress modeling indicates that the stresses within AlGaN/GaN HEMTs are not biaxial in nature, and that stresses along the width of the channel are higher than stresses along with length of the

channel due to the aspect ratio of the channel and the cylindrical shape of the heat generation region.

- A method for solving the DOM in COMSOL was developed, and furthermore, a method for coupling this model to a Fourier domain was formulated.
- A method was developed to compress phonon interaction strength data into a usable form for calculation of phonon group energies within devices. In addition a method was developed to ensure that these compressed interaction strengths represent energy conserving and physically reasonable processes.
- Good agreement was seen between phonon lifetimes as measured by Raman Spectroscopy and those determined from the phonon interaction expressions and energies calculated from the *ab. initio* program ABINIT with only one fitting parameter for the thermal conductivity of GaN.
- A semi-implicit solver method was developed that allows for rapid solution times for reasonably sized problems. In addition, an iterative solver method based on this implicit method was developed which allows for rapid solution of even large problems thanks to the Generalized Minimum Residual method (GMRES).
- It was determined that for the Finite Volume Discrete Ordinates Method, for reasonable grid spacings, it is necessary to use a third order scheme in order to accurately capture heat generation without losses.
- Ballistic and diffusive results in the Discrete Ordinates Method are dependent on quadrature selection. Quadrature sets that do not conserve the second moment (eq. 4.140) do not match diffusive heat flux results accurately. In addition, quadrature selection affects ballistic transport accuracy.

- A correction procedure for control angle quadratures was proposed which forces them to conserve the second moment and also performs as well as uncorrected quadrature sets in the ballistic regime. This has never been shown before, and allows for CADOM quadratures to be used with accuracy at all scales and any arbitrary number of directions.
- A Fourier model does not accurately capture the peak temperatures within a GaN HEMT, in fact, due to ballistic and dispersive effects, peak temperatures are likely to be significantly higher than predicted by a Fourier model. The test cases presented here confirm that and the model here builds a framework upon which to begin studying in more detail the temperature deviation of the high energy modes and the added thermal resistance from the phonon bottleneck between high energy optical modes and lower energy acoustic modes.

In addition to these contributions, this work also introduces a few items for further consideration which will develop the model further and facilitate better predictions and matching with experimental data. These include:

- Further work needs to be done to form a complete stress model of these devices. A partially developed residual stress model is currently under development. That model, which was not presented here, includes the effect of the thermal contraction of layers after growth. However, the total residual stress is likely to include many more contributions than just from thermal expansion mismatches between layers. In addition, an inverse-piezoelectric model was partially developed in COMSOL which used the electric field data from the device simulator Sentaurus as input to determine inverse piezoelectric stress and strain. These models should be developed further in order to gain insight into the complete stress state of AlGaN/GaN HEMTs.

- The phonon interaction strength model could be significantly improved. While this work included advances suggested by previous work such as inclusion of high energy optical modes thanks to the use of *ab. initio* lattice dynamics program ABINIT, it still leaves out the third order derivative of energy which give the interaction strengths of the three phonon processes and would provide orders of magnitude more accuracy in determining band-to-band interactions.
- Anisotropy of the phonon model could be included. At this point the model is isotropic meaning that the phonon interactions directional biases are washed out. The dispersion curves do not mandate this, so extension of the model to include anisotropic interactions would not be a large piece of work. Then the effect of the anisotropy in thermal conductivity in GaN could be more fully investigated.
- Temperature dependence of the band-to-band model should be included. The current work does not include temperature dependence of the phonon interaction strengths or any phonon properties, some of which are fairly strong functions of temperature. There are several ways this could be included, but it adds a significant amount of complexity to the model and was not attempted at this time. However, for accurate results, and probably for matching to experiments for higher power results, this is critical.
- Boundary scattering effects at the interfaces within AlGaN/GaN HFETs should be considered. Probably the most important interface is the GaN/AlN/SiC interface, but the AlGaN/GaN interface could also be important.

- Defect scattering and isotope scattering was not included in this model, but could be included in the future to more fully capture the temperature dependence of the thermal conductivity.
- The band to band model could include coupling with a Monte Carlo electrical simulator which would give detailed information about the generation of high energy optical phonon modes. This will be necessary before the band resolution can be increased.
- The computational model could be modified to include the SMART scheme which is a third flux limiting scheme which will hopefully eliminate oscillations near boundaries and generation discontinuities. This is not trivial as the flux limiters are nonlinear and the current equations are strictly linear. The solution method would need to be significantly modified and probably re-developed in a lower level code in order to improve speed. Implementing a non-linear solver in MATLAB of comparable speed to the current solver is probably difficult or impossible.
- The Finite Volume DOM model at this point cannot handle large domains due to memory and computational time constraints. Extending the model so that it couples to a Fourier style domain at the boundaries has been proposed and executed with a Lattice Boltzmann Method code by Christensen[2] and a similar method could be used with the DOM code in order to increase the domain size. This has already been partially developed for the COMSOL portion of the code but needs to be developed for the Finite Volume discretization code.

APPENDIX A

MATERIALS PROPERTIES

Table A.1. Gallium Nitride Material Properties

Thermal Conductivity[2]	$k = 762.78[\mathrm{W/cm}] \cdot T^{-1}$
Specific Heat Capacity[2]	$C_p = 2.2[\mathrm{J/cm^3 \cdot K}]$
Elastic Moduli[111]	$C_{11} = 390[\mathrm{GPa}]\ C_{33} = 398[\mathrm{GPa}]\ C_{44} = 105[\mathrm{GPa}]$ $C_{66} = 123[\mathrm{GPa}]\ C_{12} = 145[\mathrm{GPa}]\ C_{13} = 106[\mathrm{GPa}]$
Thermal Expansion Coefficients[112]	$\alpha_{\parallel \text{c-axis}} = 3.17 \times 10^{-6}[1/\mathrm{K}]$ $\alpha_{\perp \text{c-axis}} = 5.59 \times 10^{-6}[1/\mathrm{K}]$
Density[113]	$\rho = 6.15[\mathrm{g/cm^3}]$
Piezoelectric Constants[114]	$e_{15} = -0.3[\mathrm{C/m^2}]\ e_{31} = -0.33[\mathrm{C/m^2}]\ e_{33} = 0.65[\mathrm{C/m^2}]$
Dielectric Constants[115]	$\epsilon_{\parallel \text{c-axis}} = 10.4\ \epsilon_{\perp \text{c-axis}} = 9.5$

Table A.2. Silicon Carbide Material Properties

Thermal Conductivity[116]	$k = 61100[\mathrm{W/cm}]/(T - 155[\mathrm{K}])$
Specific Heat Capacity[113]	$C_p = 2.21[\mathrm{J/cm^3 \cdot K}]$
Elastic Moduli[117]	$C_{11} = 501[\mathrm{GPa}]\ C_{33} = 553[\mathrm{GPa}]\ C_{44} = 163[\mathrm{GPa}]$ $C_{66} = 195[\mathrm{GPa}]\ C_{12} = 111[\mathrm{GPa}]\ C_{13} = 52[\mathrm{GPa}]$
Thermal Expansion Coefficients[118]	$\alpha_{\parallel \text{c-axis}} = 3.14 \times 10^{-6}[1/\mathrm{K}]$ $\alpha_{\perp \text{c-axis}} = 3.28 \times 10^{-6}[1/\mathrm{K}]$
Density[119]	$\rho = 3.211[\mathrm{g/cm^3}]$
Piezoelectric Constants[120]	$e_{15} = -0.198[\mathrm{C/m^2}]\ e_{31} = -0.2[\mathrm{C/m^2}]\ e_{33} = 0.399[\mathrm{C/m^2}]$
Dielectric Constants[121]	$\epsilon_{\parallel \text{c-axis}} = 10.03\ \epsilon_{\perp \text{c-axis}} = 9.66$

Table A.3. Copper Tungsten Material Properties

Thermal Conductivity[122]	$k = -3.2501 \times 10^{-4} T^2 + 0.4220413T + 221.45450438\ [\mathrm{W/m/K}]$
Specific Heat Capacity	$C_p = 175[\mathrm{J/kg \cdot K}]$
Elastic Moduli	$E = 357[\mathrm{GPa}], \nu = 0.289$
Thermal Expansion Coefficients	$\alpha = 6.16 \times 10^{-6}[K^{-1}]$
Density	$\rho = 16.31[\mathrm{g/cm^3}]$

REFERENCES

[1] R. Gaska, *et al.*, "Electron mobility in modulation-doped AlGaN--GaN heterostructures," *Applied Physics Letters,* vol. 74, pp. 287-289, 1999.

[2] A. Christensen, "Multiscale Modeling of Thermal Transport in Gallium Nitride Microelectronics," Mechanical Engineering Dissertation, Mechanical Engineering, Georgia Institute of Technology, Atlanta, 2009.

[3] I. Daumiller, *et al.*, "Evaluation of the temperature stability of AlGaN/GaN heterostructure FETs," *Electron Device Letters, IEEE,* vol. 20, pp. 448-450, 1999.

[4] U. K. Mishra, *et al.*, "GaN-Based RF Power Devices and Amplifiers," *Proceedings of the IEEE,* vol. 96, pp. 287-305, 2008.

[5] P. Burgaud, *et al.*, "Preliminary reliability assessment and failure physical analysis on AlGaN/GaN HEMTs COTS," *Microelectronics Reliability,* vol. 47, pp. 1653-1657.

[6] Y. F. Wu, *et al.*, "40-W/mm Double Field-plated GaN HEMTs," in *Device Research Conference, 2006 64th*, 2006, pp. 151-152.

[7] S. Sridharan, *et al.*, "Electrothermal analysis of AlGaN/GaN high electron mobility transistors," *Journal of Computational Electronics,* vol. 7, pp. 236-239, 2008.

[8] Y. Ando, *et al.*, "10-W/mm AlGaN-GaN HFET with a field modulating plate," *Electron Device Letters, IEEE,* vol. 24, pp. 289-291, 2003.

[9] J. Li, *et al.*, "High breakdown voltage GaN HFET with field plate," *Electronics Letters,* vol. 37, pp. 196-197, 2001.

[10] J. W. Pomeroy, *et al.*, "Direct optical measurement of hot-phonons in active AlGaN/GaN devices," *physica status solidi (b),* vol. 245, pp. 910-912, 2008.

[11] A. R. Vasconcellos, *et al.*, "Hot-phonon bottleneck in the photoinjected plasma in GaN," *Applied Physics Letters,* vol. 82, pp. 2455-2457, 2003.

[12] G. Meneghesso, *et al.*, "Reliability of GaN High-Electron-Mobility Transistors: State of the Art and Perspectives," *Device and Materials Reliability, IEEE Transactions on,* vol. 8, pp. 332-343, 2008.

[13] H. Kim, *et al.*, "Reliability Evaluation of High Power AlGaN/GaN HEMTs on SiC Substrate," *physica status solidi (a),* vol. 188, pp. 203-206, 2001.

[14] J. A. del Alamo and J. Joh, "GaN HEMT reliability," *Microelectronics Reliability,* vol. 49, pp. 1200-1206, 2009.
[15] A. Sozza, *et al.*, "A 3000 hours DC Life Test on AlGaN/GaN HEMT for RF and microwave applications," *Microelectronics Reliability,* vol. 45, pp. 1617-1621, 2005.
[16] Y. C. Chou, *et al.*, "Degradation of AlGaN/GaN HEMTs under elevated temperature lifetesting," *Microelectronics and Reliability,* vol. 44, pp. 1033-1038, 2004.
[17] A. M. Conway, *et al.*, "Accelerated RF life Testing of Gan Hfets," in *Reliability physics symposium, 2007. proceedings. 45th annual. ieee international*, 2007, pp. 472-475.
[18] J. Joh and J. A. del Alamo, "Mechanisms for Electrical Degradation of GaN High-Electron Mobility Transistors," in *Electron Devices Meeting, 2006. IEDM '06. International*, 2006, pp. 1-4.
[19] J. Jungwoo and J. A. del Alamo, "Critical Voltage for Electrical Degradation of GaN High-Electron Mobility Transistors," *Electron Device Letters, IEEE,* vol. 29, pp. 287-289, 2008.
[20] P. Makaram, *et al.*, "Evolution of structural defects associated with electrical degradation in AlGaN/GaN high electron mobility transistors," *Applied Physics Letters,* vol. 96, pp. 233509-3, 2010.
[21] E. T. Yu, *et al.*, "Measurement of piezoelectrically induced charge in GaN/AlGaN heterostructure field-effect transistors," *Applied Physics Letters,* vol. 71, pp. 2794-2796, 1997.
[22] B. Jogai, *et al.*, "Effect of electromechanical coupling on the strain in AlGaN/GaN heterojunction field effect transistors," *Journal of Applied Physics,* vol. 94, pp. 3984-3989, 2003.
[23] A. Christensen and S. Graham, "Multiscale Lattice Boltzmann Modeling of Phonon Transport in Crystalline Semiconductor Materials," *Numerical Heat Transfer: Part B -- Fundamentals,* vol. 57, pp. 89-109, 2010.
[24] S. Mazumder and A. Majumdar, "Monte Carlo Study of Phonon Transport in Solid Thin Films Including Dispersion and Polarization," *Journal of Heat Transfer,* vol. 123, pp. 749-759, 2001.
[25] P. G. Klemens, "Thermal Conductivity and Lattice Vibrational Modes," in *Solid State Physics*. vol. Volume 7, S. Frederick and T. David, Eds., ed: Academic Press, 1958, pp. 1-98.
[26] M. G. Holland, "Analysis of Lattice Thermal Conductivity," *Physical Review,* vol. 132, pp. 2461-2471, 1963.
[27] M. G. Holland, "Phonon Scattering in Semiconductors From Thermal Conductivity Studies," *Physical Review,* vol. 134, p. A471, 1964.

[28] S. V. J. Narumanchi, *et al.*, "Submicron Heat Transport Model in Silicon Accounting for Phonon Dispersion and Polarization," *J. Heat Transfer*, vol. 126, pp. 946-955, December 2004 2004.
[29] S. V. J. Narumanchi, *et al.*, "Comparison of Different Phonon Transport Models for Predicting Heat Conduction in Silicon-on-Insulator Transistors," *Journal of Heat Transfer*, vol. 127, pp. 713-723, 2005.
[30] J. W. Johnson, *et al.*, "12 W/mm AlGaN-GaN HFETs on silicon substrates," *Electron Device Letters, IEEE*, vol. 25, pp. 459-461, 2004.
[31] W. Saito, *et al.*, "High breakdown Voltage undoped AlGaN-GaN power HEMT on sapphire substrate and its demonstration for DC-DC converter application," *Electron Devices, IEEE Transactions on*, vol. 51, pp. 1913-1917, 2004.
[32] G. H. Jessen, *et al.*, "AlGaN/GaN HEMT on Diamond Technology Demonstration," in *Compound Semiconductor Integrated Circuit Symposium, 2006. CSIC 2006. IEEE*, 2006, pp. 271-274.
[33] A. Vescan, *et al.*, "AlGaN/GaN HFETs on 100 mm Silicon Substrates for Commercial Wireless Applications," *physica status solidi (c)*, vol. 0, pp. 52-56, 2003.
[34] J. G. Felbinger, *et al.*, "Comparison of GaN HEMTs on Diamond and SiC Substrates," *Electron Device Letters, IEEE*, vol. 28, pp. 948-950, 2007.
[35] L. F. Eastman, *et al.*, "Undoped AlGaN/GaN HEMTs for microwave power amplification," *Electron Devices, IEEE Transactions on*, vol. 48, pp. 479-485, 2001.
[36] A. Sarua, *et al.*, "Thermal Boundary Resistance Between GaN and Substrate in AlGaN/GaN Electronic Devices," *Electron Devices, IEEE Transactions on*, vol. 54, pp. 3152-3158, 2007.
[37] J. Kuzmik, *et al.*, "Investigation of the thermal boundary resistance at the III-Nitride/substrate interface using optical methods," *Journal of Applied Physics*, vol. 101, pp. 054508-6, 2007.
[38] V. O. Turin and A. A. Balandin, "Performance degradation of GaN field-effect transistors due to thermal boundary resistance at GaN/substrate interface," *Electronics Letters*, vol. 40, pp. 81-83, 2004.
[39] O. Ambacher, *et al.*, "Two-dimensional electron gases induced by spontaneous and piezoelectric polarization charges in N- and Ga-face AlGaN/GaN heterostructures," *Journal of Applied Physics*, vol. 85, pp. 3222-3233, 1999.

[40] S. Ruvimov, *et al.*, "Microstructure of Ti/Al ohmic contacts for n-AlGaN," *Applied Physics Letters*, vol. 73, pp. 2582-2584, 1998.
[41] B. Jacobs, *et al.*, "Optimisation of the Ti/Al/Ni/Au ohmic contact on AlGaN/GaN FET structures," *Journal of Crystal Growth*, vol. 241, pp. 15-18, 2002.
[42] A. Sozza, *et al.*, "Evidence of traps creation in GaN/AlGaN/GaN HEMTs after a 3000 hour on-state and off-state hot-electron stress," in *Electron Devices Meeting, 2005. IEDM Technical Digest. IEEE International*, 2005, pp. 4 pp.-593.
[43] D. Qiao, *et al.*, "Ta-based interface ohmic contacts to AlGaN/GaN heterostructures," *Journal of Applied Physics*, vol. 89, p. 5543, 2001.
[44] M. A. Khan, *et al.*, "AlGaN/GaN metal oxide semiconductor heterostructure field effect transistor," *Electron Device Letters, IEEE*, vol. 21, pp. 63-65, 2000.
[45] B. M. Green, *et al.*, "The effect of surface passivation on the microwave characteristics of undoped AlGaN/GaN HEMTs," *Electron Device Letters, IEEE*, vol. 21, pp. 268-270, 2000.
[46] M. A. Mastro, *et al.*, "Simulation on the effect of non-uniform strain from the passivation layer on AlGaN/GaN HEMT," *Microelectronics Journal*, vol. 36, pp. 705-711, 2005.
[47] M. Kuball, *et al.*, "Measurement of temperature in active high-power AlGaN/GaN HFETs using Raman spectroscopy," *Electron Device Letters, IEEE*, vol. 23, pp. 7-9, 2002.
[48] N. Shigekawa, *et al.*, "Optical study of high-biased AlGaN/GaN high-electron-mobility transistors," *Journal of Applied Physics*, vol. 92, pp. 531-535, 2002.
[49] J. Kuzmik, *et al.*, "Determination of channel temperature in AlGaN/GaN HEMTs grown on sapphire and silicon substrates using DC characterization method," *Electron Devices, IEEE Transactions on*, vol. 49, pp. 1496-1498, 2002.
[50] L. Bergman, *et al.*, "Raman analysis of phonon lifetimes in AlN and GaN of wurtzite structure," *Physical Review B*, vol. 59, p. 12977, 1999.
[51] T. Beechem, *et al.*, "Assessment of stress contributions in GaN high electron mobility transistors of differing substrates using Raman spectroscopy," *Journal of Applied Physics*, vol. 106, pp. 114509-9, 2009.
[52] M. Kuball, *et al.*, "Measurement of temperature distribution in multifinger AlGaN/GaN heterostructure field-effect transistors using

micro-Raman spectroscopy," *Applied Physics Letters,* vol. 82, pp. 124-126, 2003.

[53] T. Batten, *et al.*, "Temperature analysis of AlGaN/GaN based devices using photoluminescence spectroscopy: Challenges and comparison to Raman thermography," *Journal of Applied Physics,* vol. 107, p. 074502, 2010.

[54] J. Jungwoo, *et al.*, "Gate Current Degradation Mechanisms of GaN High Electron Mobility Transistors," in *Electron Devices Meeting, 2007. IEDM 2007. IEEE International*, 2007, pp. 385-388.

[55] A. Sarua, *et al.*, "Thermal and Piezoelectric Stress in Operating AlGaN/GaN HFET Devices and Effect
of the Fe Doping in the GaN Buffer Layer," presented at the CS MANTECH Conference, Tampa, FL, 2009.

[56] A. Sarua, *et al.*, "Piezoelectric strain in AlGaN/GaN heterostructure field-effect transistors under bias," *Applied Physics Letters,* vol. 88, pp. 103502-3, 2006.

[57] Y. F. Wu, *et al.*, "High power AlGaN/GaN HEMTs for microwave applications," *Solid-State Electronics,* vol. 41, pp. 1569-1574, 1997.

[58] A. M. Darwish, *et al.*, "Thermal resistance calculation of AlGaN-GaN devices," *Microwave Theory and Techniques, IEEE Transactions on,* vol. 52, pp. 2611-2620, 2004.

[59] J. Park, *et al.*, "Thermal modeling and measurement of GaN-based HFET devices," *Electron Device Letters, IEEE,* vol. 24, pp. 424-426, 2003.

[60] E. R. Heller, "Simulation of Life Testing Procedures for Estimating Long-Term Degradation and Lifetime of AlGaN/GaN HEMTs," *Electron Devices, IEEE Transactions on,* vol. 55, pp. 2554-2560, 2008.

[61] E. R. Heller and A. Crespo, "Electro-thermal modeling of multifinger AlGaN/GaN HEMT device operation including thermal substrate effects," *Microelectronics Reliability,* vol. 48, pp. 45-50, 2008.

[62] Z. Wang, *et al.*, "Atomistic simulation of the size and orientation dependences of thermal conductivity in GaN nanowires," *Applied Physics Letters,* vol. 90, pp. 161923-3, 2007.

[63] A. Majumdar, "Microscale Heat Conduction in Dielectric Thin Films," *J. Heat Transfer,* vol. 115, pp. 7-16, February 1993 1993.

[64] Y. J. Han and P. G. Klemens, "Anharmonic thermal resistivity of dielectric crystals at low temperatures," *Physical Review B,* vol. 48, pp. 6033-6042, 1993.

[65] F. Gao, "Degradation Study of AlGaN/GaN HEMT through Electro-Thermo-Mechanical Calculations and Thermo-Reflectance Measurements," Materials Science and Engineering Masters, Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge, MA, 2010.
[66] *COMSOL Multiphysics User's Guide*. Burlington, MA: COMSOL, Inc., 2008.
[67] I. Newton, "Scala graduum Caloris," *Philosophical Transactions,* vol. 22, pp. 824-829, 1701.
[68] J. B. J. Fourier, *Théorie Analytique de la Chaleur*, 1822.
[69] K. C. Cheng and T. Fujii, "heat in history Isaac Newton and Heat Transfer," *Heat Transfer Engineering,* vol. 19, pp. 9-21, 1998.
[70] M. Kuball, *et al.*, "Self-Heating Effects in High-Power AlGaN/GaN HFETs," *Mat. Res. Soc. Symp. Proc.*, vol. 693, 2002.
[71] D. Gregušová and et al., "AlGaN/GaN metal–oxide–semiconductor heterostructure field-effect transistors with 4 nm thick Al 2 O 3 gate oxide," *Semiconductor Science and Technology,* vol. 22, p. 947, 2007.
[72] Khan M A, *et al.*, "AlGaN/GAN metal oxide semiconductor heterostructure field effect transistor," *IEEE Electron Device Lett.,* vol. 21, pp. 63-65, 2000.
[73] Adivarahan V, *et al.*, "Submicron Gate Si_3N_4/AlGaN/GaN-metal-insulator-semiconductor heterostructure field-effect transistors," *IEEE Electron Device Lett.,* vol. 24, pp. 541-543, 2003.
[74] Mehandru R, *et al.*, "AlGaN/GaN metal-oxide-semiconductor high electron mobility transistors using Sc_2O_3 as the gate oxide and surface passivation," *Appl. Phys. Lett.,* vol. 82, pp. 2530-2532, 2003.
[75] Gregusova D, *et al.*, "AlGaN/GaN metal-oxide-semiconductor heterostructure field-effect transistors with 4nm thick Al_2O_3 gate oxide," *Semicond. Sci. Technol.,* vol. 22, pp. 947-951, 2007.
[76] Rai S, *et al.*, "Low threshold-14W/mm ZrO_2/AlGaN/GaN metal-oxide-semiconductor heterostructure field effect transistors," *Jpn. J. Appl. Phys.,* vol. 45, pp. 4985-4987, 2006.
[77] N. Maeda, *et al.*, "Systematic Study of Insulator Deposition Effect (Si3N4, SiO2, AlN, and Al2O3) on Electrical Properties in AlGaN/GaN Heterostructures," *Japanese Journal of Applied Physics,* vol. 46, p. 547, 2007.
[78] M. T. Dove, *Introduction to Lattice Dynamics*. Cambridge, UK: Cambridge University Press, 1993.
[79] A. Greenstein, "Analysis of Thermal Conductivity Models with an Extension to Complex Crystalline Materials," Mechanical

Engineering Dissertation, George W. Woodruff School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, GA, 2008.
[80] J. A. Pascual-Gutiérrez, *et al.*, "Thermal Conductivity and phonon transport properties of silicon using perturbation theory and the environment-dependent interatomic potential," *Journal of Applied Physics,* vol. 106, 2009.
[81] X. Gonze, *et al.*, "First-principles computation of material properties: the ABINIT software project," *Computational Materials Science,* vol. 25, pp. 478-492, 2002.
[82] X. Gonze, *et al.*, "ABINIT: First-principles approach to material and nanosystem properties," *Computer Physics Communications,* vol. 180, pp. 2582-2615, 2009.
[83] X. Gonze, "A brief introduction to the ABINIT software package," *Zeitschrift für Kristallographie,* vol. 220, pp. 558-562, 2005/05/01 2005.
[84] L. S. Pereira, *et al.*, "Dynamical and thermodynamic properties of III-nitrides," *Microelectronics Journal,* vol. 34, pp. 655-657, 2003.
[85] C. Stampfl and C. G. Van de Walle, "Density-functional calculations for III-V nitrides using the local-density approximation and the generalized gradient approximation," *Physical Review B,* vol. 59, p. 5521, 1999.
[86] Y.-N. Xu and W. Y. Ching, "Electronic, optical, and structural properties of some wurtzite crystals," *Physical Review B,* vol. 48, p. 4335, 1993.
[87] H. J. Monkhorst and J. D. Pack, "Special points for Brillouin-zone integrations," *Physical Review B,* vol. 13, p. 5188, 1976.
[88] S. R. Sihn, Ajit K. , "Nanoscale Heat Transfer using Phonon Boltzmann Transport Equation," presented at the COMSOL Conference, Boston, 2009.
[89] J. Y. Murthy and S. R. Mathur, "Computation of Sub-Micron Thermal Transport Using an Unstructured Finite Volume Method," *J. Heat Transfer,* vol. 124, December 2002 2002.
[90] J. Y. Murthy and S. R. Mathur, "Finite Volume Method for Radiative Heat Transfer Using Unstructured Meshes," *J. of Thermophysics and Heat Transfer,* vol. 12, pp. 313-321, July 1998 1998.
[91] *COMSOL Multiphysics Reference Guide*. Burlington, MA: COMSOL, Inc., 2008.
[92] M. I. Flik and C. L. Tien, "Size Effect on the Thermal Conductivity of High-Tc Thin-Film Superconductors," *J. Heat Transfer,* vol. 112, pp. 872-881, November 1990 1990.

[93] G. P. Srivastava, *The Physics of Phonons*. Bristol: IOP Publishing Ltd, 1990.
[94] K. T. Tsen, *et al.*, "Subpicosecond time-resolved Raman studies of LO phonons in GaN: Dependence on photoexcited carrier density," *Applied Physics Letters*, vol. 89, pp. 112111-3, 2006.
[95] D. Y. Song, *et al.*, "Decay of zone-center phonons in GaN with A[sub 1], E[sub 1], and E[sub 2] symmetries," *Journal of Applied Physics*, vol. 101, pp. 053535-6, 2007.
[96] J. C. Chai and S. V. Patankar, *Discrete-Ordinates and Finite-Volume Methods for Radiative Heat Transfer*: John Wiley & Sons, Inc., 2009.
[97] P. Heino, "Lattice-Boltzmann finite-difference model with optical phonons for nanoscale thermal conduction," *Computers and Mathematics with Applications*, vol. 59, pp. 2381-2359, 2010.
[98] J. C. Chai, *et al.*, "Ray Effect and False Scattering in the Discrete Ordinates Method," *Numerical Heat Transfer, Part B: Fundamentals: An International Journal of Computation and Methodology*, vol. 24, pp. 373 - 389, 1993.
[99] D. Lemonnier, "Solution of the Boltzmann Equationfor Phonon Transport," in *Microscale and Nanoscale Heat Transfer*. vol. 107, S. Volz, Ed., ed: Springer Berlin / Heidelberg, 2007, pp. 77-106.
[100] R. Koch and R. Becker, "Evaluation of quadrature schemes for the discrete ordinates method," *Journal of Quantitative Spectroscopy and Radiative Transfer*, vol. 84, pp. 423-435, 2004.
[101] B. G. Carlson and L. Los Alamos Scientific, *Tables of equal weight quadrature EQn over the unit sphere*. Los Alamos, N.M.: Los Alamos Scientific Laboratory of the University of California, 1971.
[102] V. I. Lebedev, "Values of the nodes and weights of ninth to seventeenth order gauss-markov quadrature formulae invariant under the octahedron group with inversion," *USSR Computational Mathematics and Mathematical Physics*, vol. 15, pp. 44-51, 1975.
[103] R. Koch, *et al.*, "Discrete ordinates quadrature schemes for multidimensional radiative transfer," *Journal of Quantitative Spectroscopy and Radiative Transfer*, vol. 53, pp. 353-372, 1995.
[104] B. P. Leonard, "A stable and accurate convective modelling procedure based on quadratic upstream interpolation," *Computer Methods in Applied Mechanics and Engineering*, vol. 19, pp. 59-98, 1979.
[105] S. V. Patankar, *Numerical Heat Transfer and Fluid Flow*: Hemisphere Publishing Corporation, 1980.
[106] C. T. Kelley, *Iterative Methods for Linear and Nonlinear Equations*: SIAM, 1995.

[107] M. A. Heaslet and R. F. Warming, "Radiative transport and wall temperature slip in an absorbing planar medium," *International Journal of Heat and Mass Transfer,* vol. 8, pp. 979-994, 1965.
[108] P. H. Gaskell and A. K. C. Lau, "Curvature-compensated convective transport: SMART, A new boundedness- preserving transport algorithm," *International Journal for Numerical Methods in Fluids,* vol. 8, pp. 617-641, 1988.
[109] K. J. Yee, *et al.*, "Coherent Optical Phonon Oscillations in Bulk GaN Excited by Far below the Band Gap Photons," *Physical Review Letters*, vol. 88, p. 105501, 2002.
[110] T. Beechem and S. Graham, "Temperature and doping dependence of phonon lifetimes and decay pathways in GaN," *Journal of Applied Physics*, vol. 103, pp. 093507-8, 2008.
[111] A. Polian, *et al.*, "Elastic constants of gallium nitride," *Journal of Applied Physics,* vol. 79, pp. 3343-3344, 1996.
[112] W. Qian, *et al.*, "Structural defects and their relationship to nucleation of GaN thin films," *Mat. Res. Soc. Symp. Proc.*, vol. 423, pp. 475-486, 1996.
[113] M. E. Levinshtein, *et al.*, *Properties of Advanced Semiconductor Materials: GaN, AlN, InN, BN, SiC, SiGe*: John Wiley & Sons, 2001.
[114] A. D. Bykhovski, *et al.*, "Elastic strain relaxation and piezoeffect in GaN-AlN, GaN-AlGaN and GaN-InGaN superlattices," *Journal of Applied Physics,* vol. 81, p. 6332, 1997.
[115] A. S. Barker and M. Ilegems, "Infrared Lattice Vibrations and Free-Electron Dispersion in GaN," *Physical Review B,* vol. 7, p. 743, 1973.
[116] O. Nilsson, *et al.*, "Determination of the thermal diffusivity and conductivity of monocrystalline silicon carbide (300-2300 K)," *High Temperatures High Pressures,* vol. 29, pp. 73-79, 1997.
[117] K. Kamitani and M. Grimsditch, "The elastic constants of silicon carbide: A Brillouin-scattering study of 4H and 6H SiC single," *Journal of Applied Physics,* vol. 82, p. 3152, 1997.
[118] Z. Li and R. C. Bradt, "Thermal expansion of the hexagonal (4H) polytype of SiC," *Journal of Applied Physics*, vol. 60, pp. 612-614, 1986.
[119] A. Gomes, "Refinement of the crystal structure of SiC type 6H," *Acta Crystallographica %@ 0365-110X*, vol. 23, pp. 610-617, 1967.
[120] "Silicon carbide (SiC) piezoelectric and related constants." vol. 41A1a, O. Madelung, *et al.*, Eds., ed.
[121] L. Patrick and W. J. Choyke, "Static Dielectric Constant of SiC," *Physical Review B,* vol. 2, p. 2255, 1970.

[122] Y. D. Kim, *et al.*, "Thermal conductivity of W-Cu composites at various temperatures," *Materials Letters*, vol. 51, pp. 420-424, 2001.