Device-level Thermal Analysis of GaN-based Electronics

by

Kevin Robert Bagnall

B.S. Mechanical Engineering (2009)
University of Oklahoma

Submitted to the Department of Mechanical Engineering in
Partial Fulfillment of the Requirements for the
Degree of Master of Science in Mechanical Engineering
at the
Massachusetts Institute of Technology

June 2013

Signature of Author: ………………………………………………………………………….
Department of Mechanical Engineering
May 10, 2013

Certified by: ……………………………………………………………………………………..
Evelyn N. Wang
Associate Professor
Department of Mechanical Engineering
Thesis Supervisor

Accepted by: …………………………………………………………………………………….
David E. Hardt
Chairman, Department Committee on Graduate Theses

Device-level Thermal Analysis of GaN-based Electronics

by

Kevin Robert Bagnall

Submitted to the Department of Mechanical Engineering on May 10, 2013 in Partial Fulfillment of the Requirements for the Degree of Master of Science

Abstract

Gallium nitride (GaN)-based microelectronics are one of the most exciting semiconductor technologies for high power density and high frequency electronics. The excellent electrical properties of GaN and its related alloys (high critical electric field, carrier concentration, and carrier mobility) have enabled record-breaking performance of GaN-based high electron mobility transistors (HEMTs) for radio-frequency (RF) applications. However, the very high power density in the active region of GaN HEMTs leads to significant degradation in performance as the device temperature increases. Thus, effective thermal management of GaN-based electronics is a key to enabling the technology to reach its full potential.

Despite the vast amount of research into thermal issues in GaN-based electronics, including both modeling and experimental studies, there are a number of poorly understood issues. For instance, the heat source distribution in GaN HEMTs for RF applications has not been quantified nor have metrics been published for the heat flux in the near-junction region. Often, device engineers neglect the importance of thermal boundary conditions, which play a major role in shaping the temperature distribution in the device. Temperature rise in GaN HEMTs is typically modeled using computationally expensive numerical methods; analytical methods that are more computationally efficient are often quite limited.

In this thesis, a literature review is given that discusses previous research in thermal issues in GaN-based electronics and that provides a perspective on the important factors to consider for thermal management. Electro-thermal modeling tools validated with test devices were used to derive quantitative information about the heat source distribution in GaN HEMTs. Both numerical and analytical thermal models were developed that provide helpful insight into the dominant factors in the formation of highly localized hotspots in the near-junction region. The Kirchhoff transformation, a technique for solving the heat conduction equation for situations in which the thermal conductivity of a material depends on temperature, was extended and applied to GaN HEMTs. The research described in this thesis provides critical information in understanding thermal issues in GaN-based electronics required to develop next generation near-junction thermal management technologies.

Thesis Supervisor: Evelyn N. Wang
Title: Associate Professor, Department of Mechanical Engineering

Acknowledgements

First of all, I would like to thank and praise my Redeemer and Savior, the Lord Jesus Christ, to Whom I am so appreciative. While scientific and engineering research provides a very fulfilling career, I am realizing more and more that only the Lord Jesus is the real satisfaction. When we realize that everything under the sun is vanity of vanities (according to Solomon in the book of Ecclesiastes), the Lord ushers us into the divine romance as portrayed in the Song of Songs. The Christ we can gain in our education or career is much more precious than any groundbreaking scientific discovery, no matter how significant it appears. As I finish my Master's degree and continue for a Ph.D., I can truly testify: "I've found the One of peerless worth" (*Hymns* #509).

I would very much like to thank my advisor, Prof. Evelyn N. Wang, who has always been a support and encouragement during my time at MIT. She has fostered a creative and collaborative environment in our research group, which has made my graduate school experience overall enjoyable.

I would like to thank and acknowledge all of the students, post-docs, and visitors of the Device Research Laboratory, including Prof. Youngsuk Nam of Kyung-Hee University (formerly at post-doc at MIT) for their support and insightful discussions. I particularly appreciate the help of Daniel Hanks, who machined parts for me. I also really appreciated the help of Prof. Tomás Palacios and several members of his group (Dr. Tatsuya Fujishima, Omair Saadat, Yuhao Zhang and Daniel Piedra) who have helped me with the test devices, measurements, and electro-thermal modeling. I often felt like an honorary member of their group.

I appreciate all the advice and helpful discussions from a number of professors and researchers in industry, including Prof. Jesus del Alamo and Prof. David Perreault (MIT); Prof. Yuri S. Muzychka (Memorial University of Newfoundland); Prof. Eric Pop (Stanford); Dr. Jim Fiorenza (Analog Devices), Dr. Jungwoo Joh, Matt Romig and Dr. K. K. Koduri (Texas Instruments); Dr. Wataru Saito (Toshiba Semiconductor Co., Japan); and Dr. Paul Saunier and Dr. U.C. Chowdhury (TriQuint Semiconductor). I would like to thank my funding sources, the Department of Mechanical Engineering at MIT and the National Defense Science and Engineering Graduate Fellowship Program for graciously providing graduate fellowships. I would also like to acknowledge the project funding by the MIT GaN Energy Initiative and the MIT-SMART LEES Program.

My family and friends have also been a source of unending support and supply, which encourage me to press on even when things are difficult. Although I often felt academically challenged at school, they were faithful to convince me that I really do know a few things and just need to persevere.

Finally, I would like to thank my fiancée, Josephine Shaw, whom I love very much and am so happy to have as a part of my life. She is the most wonderful girl I have ever met and the love of my life.

Table of Contents

List of Figures

9

List of Tables

1. Introduction

The development of gallium nitride (GaN)-based electronics is one of the most exciting recent advances in electronics for high power and high voltage applications. Although most solid-state devices for digital logic and memory are based on silicon technologies, other semiconductors, particularly the III-V compound semiconductor group, have favorable properties in other applications. GaN, which was first developed for blue light emitting diodes (LEDs) in Japan, has demonstrated excellent properties for electronic devices, such as transistors. GaN transistors may either be high electron mobility transistors (HEMTs) or metal-insulator-semiconductor field effect transistors (MISFETs) depending upon how the electron channel is formed. HEMTs are currently the most common device structure for GaN-based electronics; development of vertical MISFETs in an important area of on-going research for the development of very high voltage devices. GaN-based power amplifiers (PAs) have demonstrated record-breaking radio frequency (RF) power output levels in excess of 40 W/mm [1] and GaN power-switching devices have shown breakdown voltages above 10 kV [2]. Sample images of a packaged GaN high frequency module for defense or communications applications and a very wide GaN power device for voltage conversion applications are shown in Figure 1.
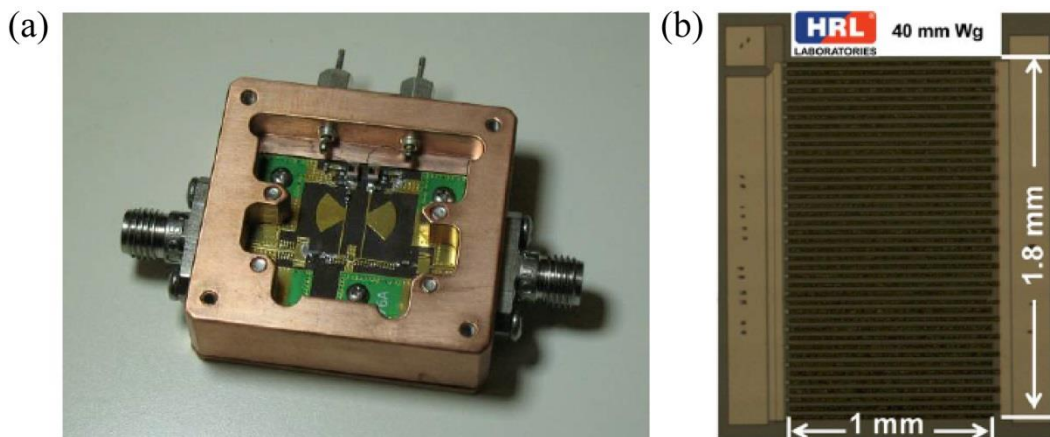
Figure 1: Commercial GaN devices
(a) High frequency GaN module [3] and (b) GaN power transistor [4]

These capabilities are due to the very high critical electrical field and electron concentration and mobility in the two dimensional electron gas (2DEG). The breakdown voltage is often defined as the maximum voltage a transistor can sustain between the drain and source terminals before an unacceptably large amount of current (typically 1 mA/mm) flows in the off-state [5]-[6]. The cutoff frequency, used as a metric for evaluating high frequency electronic devices is the frequency at which the current gain of a power amplifier is unity [5]. As shown in the plot of breakdown voltage versus cutoff frequency in Figure 2, the material limits of GaN far exceed those of other semiconductors used for high power, high frequency devices, such as silicon (Si), indium phosphide (InP), and gallium arsenide (GaAs).
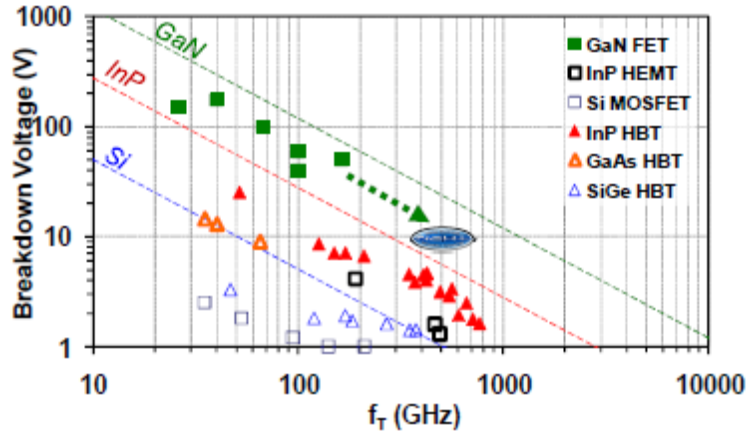
Figure 2: Breakdown voltage versus cutoff frequency for various semiconductor devices [7]

Very high power densities are possible in GaN-based electronics due to the simultaneous high voltage and high current density. As current flows through the device, power dissipation due to Joule heating effects causes the device temperature to rise, which reduces the device performance, reliability, and lifetime. It is well known that the failure rate of electronic components (a metric of reliability) increases with increasing temperature due to the variety of thermally-activated degradation mechanisms that occur more rapidly at higher temperatures [8]. As part of a recent Defense Advanced Research Projects Agency (DARPA) program aiming to provide high performance and reliable GaN HEMTs for high frequency applications, several defense contractors have provided accelerated lifetime test data for GaN HEMTs. In Figure 3 the mean time to failure (MTTF) for GaN HEMTs is plotted as a function of the device temperature from devices developed by TriQuint Semiconductor, Inc. [9].
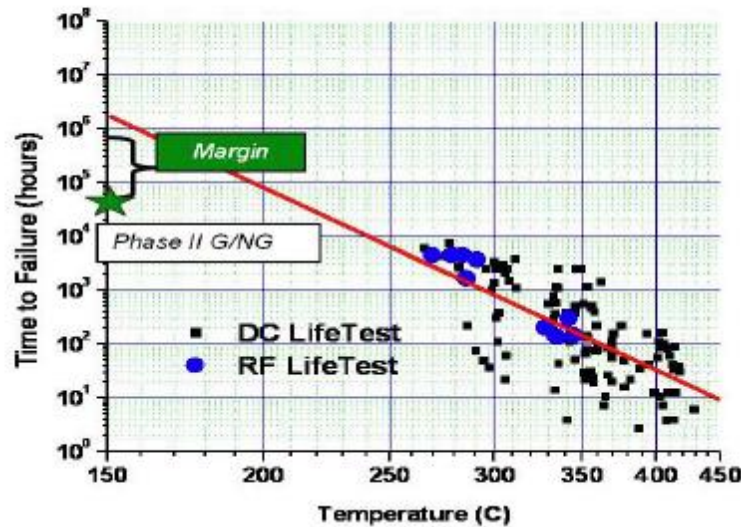
Figure 3: Mean time to failure (MTTF) for TriQuint GaN PAs [9]

The approximately exponential dependence of time to failure on the device temperature implies that small reductions in device temperature can result in greatly increased device lifetime. As will be discussed later, the peak device temperature is known to be much higher than the junction

14

or average channel temperature typically measured and reported by device manufacturers. It appears that the temperature on the horizontal axis of Figure 2 is the average channel temperature, which may be measured directly via the pulse method developed by *Joh et al. (2009)* [10] or estimated from thermal models. This data highlights the importance of keeping the channel temperature below an acceptable level in order to provide reliable and long-lasting devices.

1.1 Thermal and Electro-thermal Modeling

Since GaN-based electronics have grown in popularity, many researchers and commercial developers have recognized the importance of thermal issues and predicting the safe operating temperature of GaN-based devices with thermal models. The typical methods include thermal modeling with finite element method (FEM) and finite difference method (FDM) software packages and electro-thermal modeling tools such as Silvaco ATLAS [11], Synopsys Sentaurus Device [12], or Minimos-NT [13].

While those in traditional thermal management are more familiar with thermal modeling tools based on the FEM, there are some significant disadvantages of thermal-only finite element analysis (FEA) models. When using FEA tools to solve for the temperature field as the solution to the steady state heat conduction equation

$$\nabla \cdot (k\nabla T) + H = 0 \qquad (1)$$

where k is the thermal conductivity of the medium, T is the temperature, and H is the volumetric heat generation rate, the distribution of heat generation associated with Joule heating must be assumed. Although there are many different assumptions present in the literature, as will be discussed later, there is no general rule on how to choose this value. Joule heat generation may also be modeled as a surface heat flux at the top of the GaN layer because most of the heat generation occurs within ~20 to 30 nm of the AlGaN/GaN heterojunction. Replacement of the volumetric heat source with a surface heat flux in semiconductor devices has been practiced for several decades to reduce the mathematic complexity of solving the heat conduction equation [14]. The following from the literature only represent a sampling of numerical thermal models that have been reported for GaN-based electronics.

Garven and Calame (2009) [15] developed a 3D FDM code and compared it to the commercial FEM software ANSYS in order to optimize the gate spacing for high-power GaN MMICs. In the past, *Calame et al. (2007)* [16] at the U.S. Naval Research Laboratories have developed micro-channel liquid cooling devices based on single-phase forced convection. Their GaN MMIC models include a Tin-Silver (SnAg) solder die-attach layer and finite conductance of 6.5 x 10^4 W/m^2-K at the bottom of the chip associated with spreading and convection to the working fluid. The authors modeled the power dissipation as a uniform surface heat flux under the gate area and demonstrated that the local flux in that region is on the order of 10^6 W/cm^2. One of the key insights of the authors' analysis is highlighting the importance of the substrate base boundary condition (finite base conductance or fixed temperature) and of the SnAg die attach layer in the device-level analysis of packaged GaN MMICs. The effects of temperature-dependent thermal conductivity, finite base conductance, and the SnAg die-attach layer greatly increase the peak

temperature of the device and increase the non-uniformity in temperature profile between gate fingers.

Bertoluzza et al. (2009) [17] discussed a number of issues related to thermal FEM simulations of GaN HEMTs for RF power amplifier applications. The authors modeled a six finger GaN HEMT with a 256 μm thermal spreading region on each side of the outer most fingers. It was shown that the channel temperature increase for 3D simulations may be as much as 75% less than that for equivalent 2D simulations, depending upon the width of the gate fingers or corresponding heat sources. Transient thermal simulations were performed in which the power dissipation was modeled as a surface flux with the remainder of the top surface treated as adiabatic. The authors used a 1 μm long heat source corresponding to uniform power dissipation under the gate and stated that use of a 0.5 μm only changed the peak temperature by ± 5%. While the length of the heat source did not significantly affect their results, it may be a significant factor when more dramatic spreading associated with a large thermal spreading region is appropriate for the device layout. The importance of considering thermal spreading and sinking from the top side through the source and drain metallization layers is discussed and shown to be associated with a reduced peak temperature. Finally, reasonable agreement between previous experimental temperature measurements and FEM simulations is shown.

Douglas et al. (2011) [18] performed 2D and 3D thermal finite element simulations with the power dissipation modeled as a surface heat flux of length 1 μm, a constant substrate base temperature of 300 K, temperature-dependent thermal conductivity values, and no interfacial resistance. The authors conduct a few parametric studies that show the dependence of the maximum temperature on the length of a thermal spreading region, substrate choice, and number of gate fingers. They also discuss the difference in channel temperature predicted by 2D or 3D models and show that 2D models predict a significantly higher temperature than 3D models. However, this work is indicative of the kind of numerical thermal simulations of GaN devices that do not capture many of the key model details and that do not provide helpful physical insight.

Nochetto et al. (2011) [19] studied the effect of thermal boundary resistance between the GaN epitaxial layer and the substrate in the context of understanding the potential benefits of diamond substrates in GaN electronics. Many parametric studies were conducted and the importance of thermal spreading resistance was discussed but the results have limited value because the simulations were conducted in 2D, only single finger devices were studied, and constant thermal conductivity values were used. It is difficult to transfer the conclusions from this work to actual device layouts that are commonly used in commercial and military applications.

Electro-thermal device simulations are more comprehensive than thermal simulations in that they are capable of computing the heat source distribution from the electron transport equations and can provide an understanding of how increased device temperature affects electron transport. Locally, the heat generation rate due to the Joule heating mechanism (the transfer of energy from the electrons in a semiconductor to the lattice) is given by

$$H = \vec{J} \cdot \vec{E} \tag{2}$$

where H is the Joule heat generation rate (W/cm^3), $\vec{J}$ is the current density (A/cm^2), and $\vec{E}$ is the electric field (V/cm) [20]. Electrons gain energy from the external electric field created by the application of voltages to the contacts of a transistor or other semiconductor device; electrons then lose this energy as they collide with lattice while drifting and diffusing throughout the semiconductor lattice. Heat generation (H) with respect to the heat equation for the lattice represents the energy gained by the lattice as electrons collide with lattice atoms and generate phonons, thus raising the lattice temperature. Electro-thermal device simulation tools compute the current density and electric field distributions from the electron transport equations and electrical boundary conditions and then accurately compute the heat generation distribution. If solved in an iterative manner with the lattice heat equation, one can simulate the mutual interaction between electron transport and self-heating.

Since electro-thermal device simulation tools must solve at least two partial differential equations (Poisson's equation for electrostatics and the current continuity equation) in addition to the heat conduction equation, they are much more computationally expensive than numerical thermal models. As a result, most electro-thermal models are limited to 2D geometries with restrictive thermal boundary conditions. There can also be numerical convergence issues associated with rapid changes in electron concentration (which may vary by 15 orders of magnitude in a device) and sharp geometry changes. These factors often mean that it is not practical to simulate a large device with electro-thermal modeling tools. As noted by *Bertoluzza et al. (2009)* [17] and others, many fitting parameters are needed to match the measured electrical characteristics to actual devices, leading to skepticism about the validity of the results. In spite of these drawbacks, electro-thermal modeling tools can be a valuable source of insight into GaN HEMT device operation, in particular in understanding the role of the heat source distribution. Without electro-thermal modeling tools, one can only estimate what the heat source shape and size may be and how it should vary with device geometry and operating conditions.

In the past, electro-thermal device simulations have mostly been performed by researchers in the fields of electrical engineering and physics who have more experience with the details of semiconductor devices than those in mechanical engineering. This is indicated by the limited physical insight into the heat transfer phenomena associated with GaN HEMT self-heating and the lack of traditional heat transfer terminology.

Turin and Balandin (2006) [21] used the commercial device simulation ISE-DESSIS with the drift-diffusion model (now Synopsys Sentaurus Device [12]) to simulate the electrical characteristics and device temperature profile for a GaN MESFET (not HEMT) taken from the literature. The authors reported the electrical output characteristics and demonstrated the reduction in drain current associated with increased device temperature. The effect of an uniform base temperature of 300 K and a base resistance of 0.1 cm^2-K/W was briefly discussed, demonstrating that a resistance between the substrate base and the ambient results in a significantly increased maximum temperature. A few 2D device temperature profiles are shown and the temperature dependence on substrate thickness is discussed but little physical insight is given that can be generally applied.

Heller and Crespo (2007) [22] performed a very helpful analysis of GaN HEMTs with ISE-DESSIS for 2D electro-thermal modeling with the drift-diffusion equations coupled to 3D

thermal simulations in ANSYS. The authors used experimental electrical measurements from double finger GaN HEMTs to validate their device simulations and showed very good agreement between measured and modeled electrical curves. The thermal resistance (temperature rise divided by the power dissipation) was compared between the FEM results and an analytical model developed by *Darwish et al. (2005)* [23]. The authors observed good agreement between the analytical model and their FEM results with respect to the dependence of thermal resistance on substrate thickness and gate width. However, the thermal resistance did not vary as expected with the gate length (swept from 0.1 to 0.8 µm) in comparison to the analytical model. The analytical model developed by *Darwish et al. (2005)* assumes uniform power dissipation under the gate such that the thermal resistance increases significantly with decreasing gate length because of increasing thermal spreading resistance. *Heller and Crespo (2007)* found that the thermal resistance did not change significantly with increasing gate length, indicating that the heat source and temperature rise must depend primarily on a different geometric parameter. The authors also provided some discussion of reliability and lifetime estimates based on thermal issues in GaN HEMTs.

Benbakhti et al. (2009) [24] modeled temperature rise in ungated AlGaN/GaN structures *via* the hot electron model (energy balance equations) and compared to experimental temperature measurements acquired from electro-thermal modeling. The authors demonstrated the capability of reproducing experimental electrical and temperature measurements for the ungated structures. However, the limitation of this work is that it does not directly apply to GaN HEMTs. The presence of the gate and the depletion region in the channel has a strong effect on the electron transport between two ohmic contacts and makes electro-thermal modeling using energy balance equations more challenging.

Vitanov et al. (2010) [25] modeled complete AlGaN/GaN HEMTs with the energy balance equations in the 2D device simulation software Minimos-NT. The authors demonstrated excellent agreement between measured and modeled direct current (DC) output and transfer characteristics as well as alternating current (AC) parameters. The paper did not include any experimental temperature measurements to validate the electro-thermal model. However, it is interesting that the authors found such good agreement between their measured and modeled electrical characteristics in light of the importance of 3D thermal spreading effects, *i.e.,* the 2D temperature distribution predicted is higher than 3D temperature distribution and should result in a reduced drain current.

1.2 Experimental Temperature Measurement Techniques

Measurement of the temperature distribution in GaN-based electronics is particularly challenging and has been the subject of much investigation by researchers in the field of device reliability. Typically, device designers and manufacturers characterize the thermal resistance of an electronic device such as a power transistor by electrical methods [26]. In this technique, the device temperature is calculated indirectly from the change in an electrical parameter, such as the drain to source on-resistance or $R_{DS(ON)}$, with temperature. First, a calibration curve is constructed by externally raising the temperature of the entire device and measuring the on-resistance as a function of temperature. After a calibration correction for the dependence of on-

resistance on drain current, the device junction temperature when the device is in use can be determined from the equation

$$\Delta T_j = \left[\frac{\ln\left(R_{DS(ON)}\right) - A}{B} \right] - T_\infty \tag{3}$$

where ΔT_j is the temperature rise of the junction, A and B are constants obtained from calibration, and T_∞ is the ambient temperature [26]. Although this characterization procedure has been well-established for silicon-based power devices, it only provides a measure of the average temperature over which electron transport occurs. Therefore, it effectively measures the average channel temperature and cannot give information about the temperature distribution associated with non-uniform heat generation.

In GaN power amplifiers (PAs) for RF and high frequency applications, a large amount of heat is generated in a small region around the gate due to the simultaneous high current density and electric field. As an example, the temperature distribution from a 3D steady-state thermal model of a typical GaN PA layout is shown in Figure 4. The sharp peaks with temperatures of ~180 °C are associated with regions of high thermal spreading resistance around the intense heat sources of 0.5 μm x 150 μm area corresponding to the area of the gate. Within a few microns, the temperature of the device can change ~40 °C.
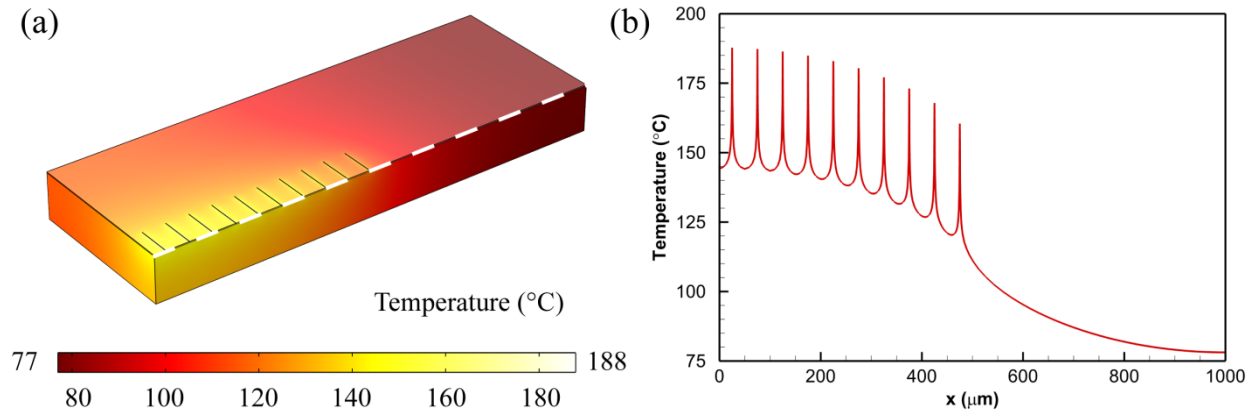
Figure 4: Example temperature distribution in GaN RF PA from thermal model
(a) **3D quarter-model and (b) 1D temperature slice along center of device (dotted white line)**

As seen in Figure 4(a), the temperature changes sharply in the region around the heat source, *i.e.,* as much as 40 °C over a 5 μm spatial region around the heat source. A measurement of the average temperature would yield a value of ~160 °C, far underestimating the peak temperature of 188 °C in the device. Thus, it has been recognized that experimental temperature measurement techniques with very high spatial resolution ($\leq$ 1 μm) are required to resolve the true temperature distribution in GaN-based electronics.

Micro-Raman thermometry, based on Raman spectroscopy, has been developed as one of the most accurate temperature measurement techniques for resolving the device temperature with high spatial resolution. Raman spectroscopy is an optical material characterization technique that

excites the electronic and lattice vibration energy modes in a solid with a laser and measures the light signal emitted by the solid. The response of the material, the Raman spectrum, provides information about its structure, quality, and other properties. A sample spectrum for a GaN-on-SiC wafer measured on a confocal Raman microscope at the MIT Spectroscopy Laboratory with 532 nm laser excitation source is shown in Figure 5.
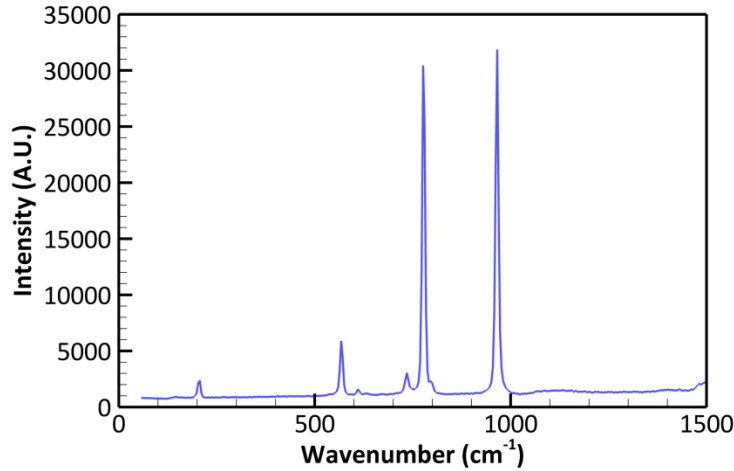
Figure 5: Raman spectrum of Cree HPSI GaN-on-SiC wafer

The peaks shown in Figure 5 are associated with phonon resonance modes in GaN and SiC and have been characterized well in the literature with theoretical calculations and experimental measurements. For the spectrum shown in Figure 5, the peaks at 567.7 cm^{-1} and 735.2 cm^{-1} are associated with the E2 high and A1 longitudinal optical (LO) lines in GaN and the peak at 776.5 cm^{-1} is associated with the E2 line in SiC. The wavenumber at which each of the peaks occur is temperature dependent [27]; thus, the local device temperature can be measured by detecting the peak associated with different phonon modes when the device is under bias [28]-[29].

Micro-Raman thermometry has been one of the most widely used temperature measurement techniques in high-power GaN electronics to validate thermal models [24], to demonstrate the effectiveness of advanced thermal management strategies [30], and to aid in understanding the influence of device structure on temperature rise [31]-[33]. The technique has been extended more recently to achieve spatial resolutions of 0.5 µm, temporal resolutions of 10 ns, and temperature resolutions of ± 5 °C [34]. However, micro-Raman thermometry does require a high capital cost to set up (~$500,000) and a moderate cost to maintain. In addition, it can only measure the temperature of one spot at a time; the sample must be scanned via a motorized stage to complete a thermal map. Therefore, micro-Raman thermometry is most suitable as a research tool when spatial resolution is one of the most important concerns.

Infrared (IR) thermometry has also been investigated as a temperature measurement technique in GaN-based electronics [35]-[36]. Since all bodies at a finite temperature emit a spectrum of radiation described by Planck's law, the temperature of the surface of a body can be measured by detecting the infrared radiation leaving the body. In infrared thermometry, infrared radiation leaving the surface is focused by a microscope objective onto the detector array (typically a low

bandgap semiconductor such as Indium antimonide). By evaluating the radiosity and accounting for variable surface emissivity, the temperature at each pixel in the detector array can be determined with a temperature resolution of ± 0.1 K and spatial resolution of ~3 µm. Because the blackbody emissive power depends on the absolute temperature raised to the fourth power, a higher signal-to-noise ratio (SNR) is achieved when the device is held at a temperature above the ambient [35]. The blackbody spectrum calculated from Planck's law for blackbody radiation and a sample temperature map of a GaN HEMT under ~23 W/mm bias obtained from the Quantum Focus Instruments Infrascope II at MIT is shown in Figure 6.
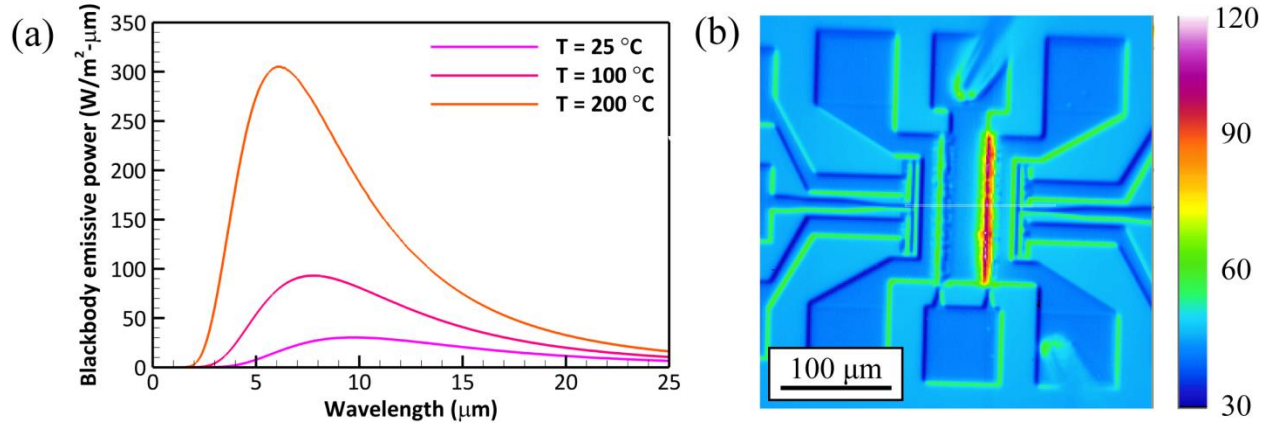
Figure 6: (a) Blackbody infrared spectrum and (b) Sample infrared temperature (°C) map of GaN HEMT under bias

The main disadvantage of IR thermometry is that the horizontal and vertical (depth of field) resolutions are limited by the wavelength of the radiation in the mid-wave IR (MWIR) range of 2 to 5.5 µm. Thus, the technique is not capable of resolving the temperature distribution with the same accuracy as micro-Raman spectroscopy. The strength of the signal also depends upon the emissivity of the materials; low emissivity materials such as metals can lead to inaccurate temperature values unless proper calibration and data analysis is used [37]-[38]. However, one advantage is that an IR microscope measures the temperature of the entire field of view at the same time, as it is a wide field microscopy technique.

Other measurement techniques based on electrical characterization [10], [39], optical interferometry [40]-[41], and scanning thermal microscopy [42] have been reported in the investigation of thermal issues in GaN-based electronics. For research applications, micro-Raman spectroscopy provides the highest spatial accuracy for fine resolution of the temperature profile. However, in deployed and packaged devices, electrical methods are more practical for measuring an average temperature that can be monitored for reliability and performance evaluation. Despite all of the work that has been done in this field, there are still opportunities for developing new techniques and utilizing the existing techniques to characterize thermal issues in GaN-based electronics in new ways.

1.3 Thermal Management

GaN-based electronics have been the target of many investigations into the improvement of thermal management on a system, package, and device-level. Thermal management of commercial GaN RF and high-frequency PAs fabricated by companies such as Raytheon, Northrup Grumman, and TriQuint Semiconductor is a standardized process similar to that developed for high power GaAs devices in the 1980s and 1990s. Devices are typically packaged in ceramic packages whose materials have moderate thermal conductivity and have similar coefficients of thermal expansion (CTE) to SiC, the dominant substrate material for high power GaN devices. The typical cost for a GaN RF PA is $1 to $2 per watt of output power; thus, a 40 W PA would likely cost $40 to $80 if purchased from a commercial supplier. This relatively high cost, compared to Si-based devices, allows the manufacturer to use a package with higher thermal performance but also makes GaN-based electronics not as cost-competitive as Si-based electronics. The high cost of GaN PAs has made their market share limited in civilian communications applications such as cell phone base stations in the United States. It is interesting to note that GaN PAs are widely used in civilian communications systems in Japan, where efficiency is often valued as much or more than cost. GaN devices for power electronics are often packaged in more standard analog packages, such as the TO-220, which have poorer thermal performance but a significantly lower cost needed for market penetration. The following references provide a representative view of some of the thermal management strategies that have been investigated.

Researchers at the Naval Research Laboratory have studied micro-channel cooling devices based on single-phase convection with modeling and experimental approaches for almost a decade [16]. Development of a micro-channel cooler represents a package-level thermal management device, in which an unmodified GaN-based device is attached to a cooling system. *Calame et al. (2007)* [16] demonstrated that single phase cooling with water in SiC micro-channel coolers could remove 3000-4000 W/cm^2 of heat flux from the backside of a GaN device. This very high heat flux value and good performance is due to the high thermal conductivity of SiC, leading to good spreading through the channel walls, and the excellent thermophysical properties of water. Such as system, however, with all of the required components is not practical for many applications because of the overhead and power consumption associated with the system. In very niche defense applications in which high power is the most important concern, such micro-channel cooling systems may be possible to implement. The desire for passive or less complicated active cooling systems with similar performance metrics (removal of heat fluxes greater than 1000 W/cm^2) is the motivation for the DARPA Microsystems Technology Office ICECool Fundamentals Program [43].

Because the thermal management community has recognized that the highly localized hot spots in GaN-based electronics lead to high thermal spreading resistances on a device-level, many researchers have targeted thermal management approaches much closer to the region where heat is generated. One such approach is spreading or sinking from the top side of the device near the electrical contacts by adding a high thermal conductivity solid. *Tsurumi et al. (2010)* [44] introduced sputtered AlN in place of the usual passivation layer around the gate and showed a reduction in thermal resistance via electro-thermal modeling and experimental measurements by change in the electrical parameters. While a reduction in thermal resistance is clearly

demonstrated experimentally, it is difficult to interpret the modeling results presented in the paper because the bulk value of AlN thermal conductivity (285 W/m-K) was used. The authors claimed that the crystallinity of the AlN was very good from X-ray diffraction (XRD) measurements, implying a thermal conductivity approaching the bulk value. However, it is well known that polycrystalline materials with relatively small grain sizes have a thermal conductivity significantly smaller than the bulk value. Industrial producers of high power GaN RF PAs are unlikely to substitute traditional passivation materials, such as SiN and SiO_2, for a high thermal conductivity but much less electrically characterized material like sputtered AlN.

Tadjer et al. (2012) [45] investigated the use of nanocrystalline diamond films deposited above a thin SiO_2 passivation layer in AlGaN/GaN HEMTs on Si substrates. The authors reported that the thermal resistance of the device with nanocrystalline diamond was 3.75 times lower than the device with traditional passivation layers. The device with nanocrystalline diamond also exhibited moderate device performance. *Yan et al. (2012)* [30] attached few layer graphene/graphite sheets to the top of the drain contact in order to provide an additional spreading or sinking path through the high thermal conductivity graphene (2000 to 4000 W/m-K). The authors demonstrated a reduction of 20 °C in the localized device temperature at a dissipated power of 13 W/mm via micro-Raman thermometry. There are a number of issues, however, that are of a concern in introducing graphene spreaders as proposed in this paper. First, the contact resistance between graphene and the drain pad may be prohibitively high and may outweigh the benefit of graphene's high thermal conductivity. Actual sinking of the heat, rather than merely spreading, requires that one end of the graphene be maintained at a temperature near the ambient. It is not clear from the thermal simulations in this work why the heat source was chosen as a value larger than the gate length and how the thermal boundary conditions were set with respect to the graphene spreader. The local temperature was also only measured at a single location with the micro-Raman temperature probe.

Because the resistance to multi-dimensional conduction contributes so greatly to the device-level thermal resistance of GaN HEMTs, many groups have pursued the development of GaN devices on very high thermal conductivity substrates. Electrically semi-insulating SiC with a thermal conductivity of approximately 400 W/m-K is the standard substrate for high power GaN PAs. Diamond substrates, however, have demonstrated thermal conductivities of 1200 to 2000 W/m-K, depending on the crystal structure, growth method, and quality. These very high thermal conductivity diamond substrates, although expensive, provide excellent spreading from the localized hot spots in GaN HEMTs to the device package. Group4 Labs of Fremont, CA in collaboration with university partners has developed a process to transfer GaN epilayers to 2 in. diamond wafers and has demonstrated functional GaN-on-diamond PAs with reduced thermal resistance [3], [42]. In the epilayer transfer and wafer bonding process, it is critical to minimize the interfacial resistance associated with the bonding material by keeping it as thin as possible. Epitaxial growth of GaN epilayers on single crystal diamond substrates has also been reported [46]. However, the use of superlattice or other multi-layered structures required to manage the stress from GaN-diamond lattice mismatch makes the growth process quite complicated and may create an prohibitively large thermal resistance.

Otsuka et al. (2011) [47] at Panasonic Corporation in Japan developed a phase change, package-level cooling solution for GaN power transistors with far superior performance compared to

traditional TO-220 type packages. The system consisted of a hermetically sealed package similar to a heat pipe in which ethanol was boiled on top of the GaN device, condensed on a radiator, and returned as a liquid through a wicking structure. The authors showed a temperature reduction of 55 °C or 100% increase in allowable power for the phase change cooling solution compared to a TO-220 package. While this represents the potential for phase change solutions in GaN thermal management, liquid contact with the actual device may cause reliability concerns, particularly on the top side of the device.

1.4 Thesis Outline and Objective

As noted before, thermal issues in GaN-based electronics have been thoroughly investigated by researchers in academia and industry with modeling and experimental approaches. GaN HEMTs are widely used in defense, aerospace, and communications applications in the United States, Europe, Japan, and other countries. The commercial success of GaN in some application areas might lead one to think that there is no need for further research in this topic. However, the constant demand for increasing power density of GaN-based electronics and improved reliability makes thermal management of GaN HEMTs an important area of on-going research. This thesis seeks to provide a clearer understanding of the relevant factors in thermal issues in GaN-based devices, to describe best practices for thermal modeling in order to evaluate thermal management strategies, and to propose new physical insight needed to design the next generation of advanced thermal management technology.

Important research opportunities that need further investigation in this field include:

- Simpler thermal metrology and monitoring for packaged devices without the complication of high spatial resolution optical methods
- A quantitative description of the heat source and its dependence on device geometry and electrical bias condition
- Improved analytical models for temperature rise that are much more computationally efficient than FEA models
- Discussion of the distinction between GaN-based electronics in RF/high frequency and power switching applications with respect to thermal issues.

In *Chapter 1*, an introduction to GaN-based electronics is given and previous work in the fields of electro-thermal modeling, thermal modeling, experimental temperature measurement techniques, and thermal management approaches is provided.

In *Chapter 2*, electro-thermal device simulation software is used to provide **a quantitative analysis of the heat source distribution in GaN HEMTs**, in particular with respect to the gate and drain bias for RF applications. The details of the electro-thermal model are given and validation with a 2 x 250 µm AlGaN/GaN HEMT is demonstrated with experimental electrical and temperature measurements.

In *Chapter 3*, thermal modeling tools are utilized to **provide helpful physical insight into the key factors to consider in thermal modeling of GaN-based electronics**. An analytical model is developed to calculate the temperature rise in GaN HEMTs and its computational efficiency for

multi-finger HEMT structures is demonstrated. Best practices and the validity of thermal boundary conditions are discussed.

In *Chapter 4*, the **Kirchhoff transform is applied to analytical models for temperature rise in GaN HEMTs to illustrate the impact of temperature-dependent thermal conductivity**. The decrease in thermal conductivity with increasing temperature can significantly increase the peak device temperature in multi-finger HEMTs with high power dissipation. The Kirchhoff transform is extended for problems with a finite heat transfer coefficient at the substrate base for the first time.

In *Chapter 5*, the work is summarized and the future direction for next-generation thermal management of GaN-based electronics is discussed. Various possible thermal management strategies are suggested from the insight gained in this thesis.

2. Electro-thermal Modeling

Electro-thermal modeling is a valuable tool to gain insight into thermal issues in GaN-based electronics because it provides information on how electrical and thermal transport phenomena influence one another. For instance, electro-thermal modeling software computes the heat source distribution required to solve the heat conduction equation from the electric field and current density distributions. One can also study how the temperature distribution affects the electron transport in a device since electron mobility in GaN and related semiconductor materials decreases strongly with increasing temperature. There are several disadvantages, however, in using electro-thermal modeling tools. Actual devices fabricated in the cleanroom may differ significantly from ideal behavior due to variability among devices, contamination issues, material defects, device degradation over time, and phenomena such as charge trapping effects that are difficult to account for. One also needs to adjust a number of parameters involved in the electron transport that vary among GaN devices in order to obtain a good agreement between measured and modeled electrical characteristics. Finally, many device engineers and researchers question the validity of electro-thermal modeling results because of the complexity of solving several non-linear, partial differential equations.

2.1 Basics of GaN HEMTs

A GaN high electron mobility transistor (HEMT) or heterostructure field effect transistor (HFET) is a semiconductor device that acts essentially as an electrical switch. GaN (or other material) HEMTs can be used in power conversion (DC/DC voltage converters, AC/DC rectifiers, or DC/AC inverters), RF power amplification, or logic applications. While a Si-based metal oxide semiconductor field effect transistor (MOSFET) uses the MOS structure to form the electron channel (inversion layer), GaN HEMTs make use of the heterojunction between a related ternary alloy and GaN to form the two dimensional electron gas (2DEG). A simple single finger GaN HEMT under bias with electrical probes and a side view schematic of the GaN HEMT structure are shown in Figure 7.
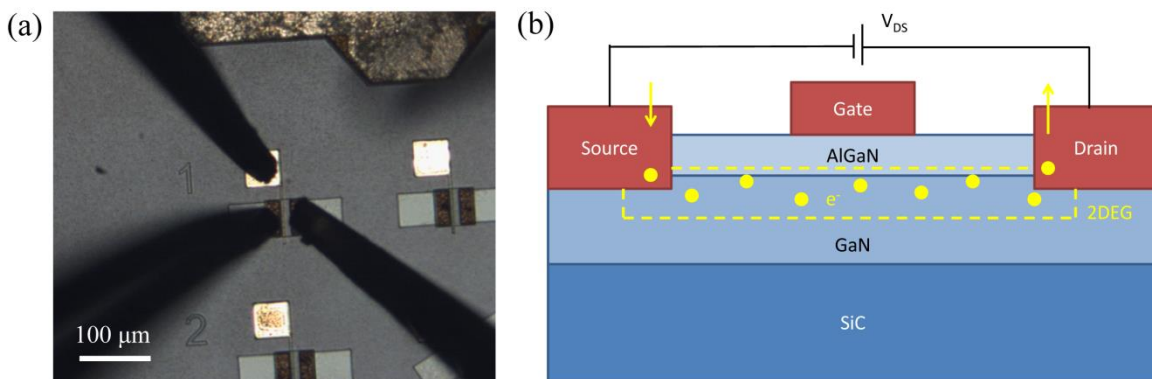
Figure 7: (a) Optical image of GaN HEMT under bias and (b) cross-section schematic of HEMT structure Optical microscope image is courtesy of Daniel Piedra (Ph. D. candidate, EECS, MIT)

GaN HEMTs are fabricated from the epitaxial growth of GaN and other layers on top of a foreign substrate material, such as SiC, sapphire, and Si. Although GaN substrates are available,

they have been historically difficult to grow in large crystals; thus, it is necessary to epitaxially grow GaN on top of a foreign substrate. The substrate should be electrically semi-insulated to prevent vertical breakdown between the channel and the transistor body and highly thermally conductive to allow for efficient removal of the heat generated in the channel. The lattice constants of the substrate should also be as close as possible to those of GaN in order to provide low stress and high quality epitaxial GaN films for device layers.

In the standard growth process by metal organic chemical vapor deposition (MOCVD), a thin AlN nucleation layer of 20 to 200 nm thickness is grown from ammonia (NH_3) and tri-methyl-aluminum (TMA) on the substrate in preparation for the GaN growth process [48]. Because the nucleation layer is very thin, it adopts the lattice spacing of the substrate and experiences significant mechanical strain. Next, GaN is grown by introducing NH_3 and tri-methyl-gallium (TMG) at high temperature and pressure. GaN nucleates on the AlN nucleation layer surface and grows upward to a height of 1 to 2 μm, depending on the device application. The initial quality of the GaN at the nucleation layer-GaN interface is quite poor; however, the crystal quality improves as the growth proceeds and reaches a high enough quality for electron transport. A ternary alloy barrier layer, such as AlGaN, is grown next by flowing a mixture of TMA and TMG in the MOCVD reaction in order to produce the desired alloy $Al_xGa_{1-x}N$, where x denotes the mole fraction of Al in the barrier layer. The mole fraction of Al typically varies between 0.1 and 0.3 and is optimized by the wafer producer in order to optimize carrier concentration and transport properties. In a series of more detailed fabrication steps, the source, gate, and drain contacts are fabricated on the AlGaN layer and passivation around the gate is introduced.

The difference in electronic band parameters between GaN and AlN (and the associated ternary allow GaN) is responsible for the presence of a high concentration 2DEG layer at the heterojunction [5]. When AlGaN and GaN are in contact, as shown in the figure below, the Fermi levels must be aligned and flat in a situation of thermal equilibrium. The larger bandgap of AlGaN causes a bend in the conduction band of the structure that leads to part of the conduction band on the GaN side lying below the Fermi level. It is energetically favorable for electrons to occupy this region; thus, a 2DEG spontaneously exits on the GaN side of the heterojunction.
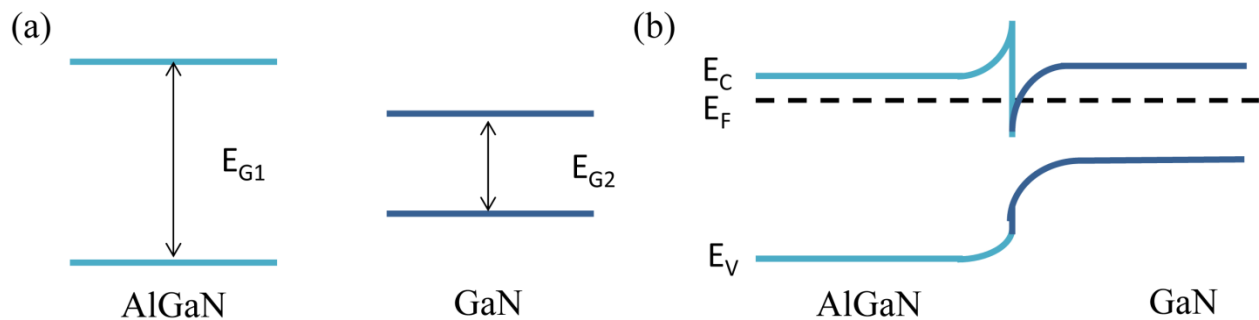
Figure 8: Schematic of heterojunction band structure
(a) semiconductors far apart and (b) semiconductors in contact in thermal equilibrium

Often, the barrier layer (in this case AlGaN) is intentionally n-type doped to introduce additional electrons into the system. The electrons from these shallow (low activation energy) donors move over to the GaN side of the heterojunction where they are trapped in the potential well formed by

the conduction band discontinuity as described above [5]. HEMTs exhibit very high electron mobility values because the excess carriers introduced by doping are physically separated from the donors themselves, leading to decreased ionized impurity scattering.

Metal contacts are used for transferring charge to and from the transistor and for controlling the charge in the device. Ohmic contacts are formed by high temperature annealing of metals to produce spikes with intermetallic compounds that penetrate through the thin barrier layer and make low resistance electrical contact with the 2DEG. The source and drain are formed with ohmic contacts and should have a linear current-voltage response with low resistance. Schottky contacts are formed by lower temperature deposition of metals on top of AlGaN layer with no penetration. The gate is formed with a Schottky contact and should have a logarithmic (similar to a diode) response with a large reverse bias blocking voltage. The ultimate performance of a GaN device depends greatly on the quality of the metal contacts, such as minimizing the ohmic contact resistance and gate leakage current. It is important to note that many GaN HEMTs use contact metallurgies involving gold (Au), which is not allowed in standard Si fabrication facilities.

The role of the drain and source contacts are to provide a path for current flow from the outside world to the 2DEG. By convention, the source is always grounded and the drain and gate voltages are always specified with respect to the source. Although positive current flows from the drain to the source, the source and drain are given their names because electrons enter the device at the source contact and exit at the drain contact. The gate contact controls the charge in the 2DEG by applying a negative (reverse bias) voltage to drive charges out of the channel or by applying a positive (forward bias) or zero voltage to allow charges in the channel. In the on-state, current flows through the device; in the off-state, one intends to restrict current from flowing in the transistor. However, due to the various non-idealities and leakage mechanisms, current may flow in the off-state.

When GaN HEMTs are used in RF or high frequency power amplification applications, DC voltages must be applied to the gate and drain contacts in order to place the device in its quiescent bias point in the saturation region. The basic circuit for a GaN PA is shown in Figure 9 [49].

28
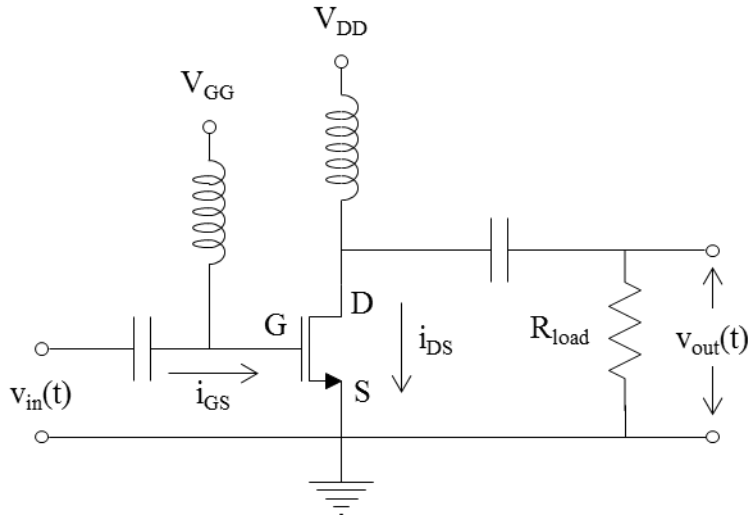
Figure 9: Schematic of basic GaN PA circuit

The input AC signal that one wishes to amplify, $v_{in}(t)$, is applied to the gate (G) and the drain (D) contact gives the amplified output AC signal, $v_{out}(t)$. From the point of energy conservation, the additional AC electrical power in the output signal comes from the DC power consumed in the process of placing the amplifier in its quiescent operating point. The large DC power dissipation associated with the quiescent drain bias is the source of thermal issues in GaN-based PAs. GaN-based PAs typically have efficiencies of 40% to 60%; as the desired RF output power increases, the DC power dissipated also increases. Thus, thermal issues in GaN-based PAs have become one of the major factors limiting the RF power density in commercial and military devices to 2 to 4 W/mm.

FETs and HEMTs are often characterized with respect to their electrical performance by two types of curves: output characteristics (drain current vs. drain voltage) and transfer characteristics (drain current vs. gate voltage). A sample of output and transfer characteristics for a 2 x 150 µm GaN-on-SiC HEMT fabricated by Omair I. Saadat (Ph.D. candidate, Department of Electrical Engineering and Computer Science, MIT) are shown in Figure 10 as an example.
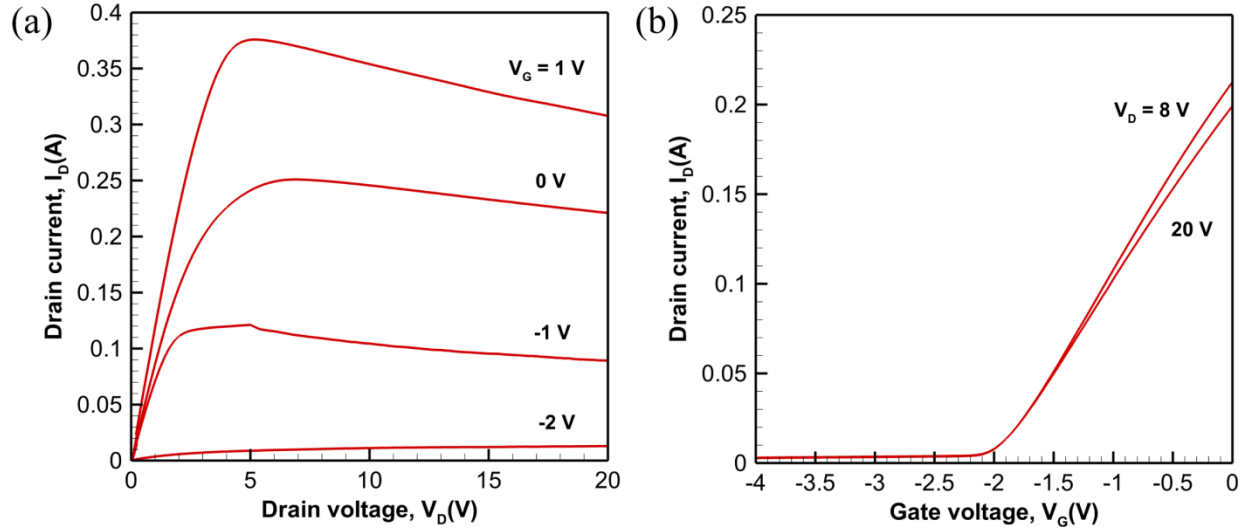
Figure 10: Sample electrical output and transfer characteristics for GaN-on-SiC HEMT
(a) Output (I_D-V_D) and (b) transfer (I_D-V_G) characteristics

The electrical output and transfer characteristics provide complimentary information about the performance of GaN HEMTs. Most electro-thermal modeling studies validate that the carrier transport described by the model is accurate by demonstrating good agreement between the modeled and measured electrical characteristics. It should be noted, however, that the device parameters measured by electrical characteristics, such as the drain current, are macroscopic quantities for the entire device. Researchers often assume that if the device parameters (material properties, electron mobility models, etc.) are as accurate as possible and the macroscopic measurements agree closely with the model, then the model captures the key physics of carrier transport. Unfortunately, it is often difficult to experimentally measure microscopic quantities, such as the electric field, electron concentration, and current density distributions, in order to more closely validate electro-thermal device models. Experimental temperature measurements also provide convincing evidence that the device model is correct.

2.2 Formulation of Equations

2.2.1 Electron Transport

The transport models used in this work are the non-isothermal drift-diffusion equation for electron transport with the classical heat equation for the lattice. These equations assume that the electrons are always in thermodynamic equilibrium with the lattice and do not distinguish between the temperature of the electrons and the temperature of the lattice. Although the equations are specified in the Silvaco ATLAS user manual for both electrons and holes (denoted with "n" and "p" subscripts, respectively), the equations given here are only for electrons [50]. Hole concentrations are very low ($\sim 10^6$ cm^{-3} or below) compared to electron concentrations and are neglected for simplicity in this work. The first basic equation governing electron transport is Poisson's equation for electrostatics, which governs the relationship between volumetric charge density and electrostatic potential

$$\nabla \cdot (\varepsilon \nabla \psi) = -\rho \tag{4}$$

where ε is the electric permittivity of the medium, ψ is the electrostatic potential, and ρ is the volumetric charge density (positive by convention). The electrical permittivity is kept inside the divergence operator to account for media in which the permittivity is anisotropic. The local volumetric charge density is due to the combination of electrons, holes, donors, and trapped charges which exist at the same location in space.

The second basic equation for carrier transport is the carrier continuity equation, which in the general case is given by

$$\frac{\partial n}{\partial t} = \frac{1}{q} \nabla \cdot \vec{J} + G - R \tag{5}$$

where n is the electron concentration per unit volume, q is the fundamental charge unit (1.6022×10^{-19} C), $\vec{J}$ is the current density (positive by convention), and G and R are the electron generation and recombination rates per unit volume, respectively. In steady-state, and the absence of significant carrier generation and recombination, the equation above reduces to

$$\nabla \cdot \vec{J} = 0. \tag{6}$$

Poisson's equation for electrostatics relates the electron concentration and electrostatic potential. In its general form, the charge conservation equation merely establishes that the current density field is divergence-free. Therefore, a constitutive model is needed to develop the relationship between current density, carrier concentration, and electrostatic potential. In this work, the constitutive transport model used was the drift-diffusion equation given by

$$\vec{J} = qD\nabla n - qn\mu\nabla\psi - \mu n[k_B T\nabla(\ln n_{ie})] \tag{7}$$

where D is the diffusivity of electrons, μ is the electron mobility, k_B is Boltzmann's constant, T is the temperature, and n_{ie} is the intrinsic carrier concentration of the material. Thus, the steady-state charge continuity equation becomes

$$\nabla \cdot \{qD\nabla n - qn\mu\nabla\psi - \mu n[k_B T\nabla(\ln n_{ie})]\} = 0 \tag{8}$$

in terms of the electron concentration and electro-static potential. The final step is to relate the electron diffusivity D to the other variables in the system, which is described in the ATLAS user manual for both Boltzmann and Fermi-Dirac statistics. With respect to electron transport, the software solves Poisson's equation for electrostatics and the continuity equation given above for the carrier concentration and electro-static potential. The electric field is derived from the basic relation

$$\vec{E} = -\nabla\psi \tag{9}$$

and the current density is calculated from the drift-diffusion equation during the post-processing phase of the model.

2.2.2 Thermal Transport

The temperature distribution in the device is calculated from the classical heat conduction equation

$$\rho c_p \frac{\partial T}{\partial t} = \nabla \cdot (k \nabla T) + H \tag{10}$$

where ρ is the density, c_p is the specific heat per unit mass, k is the thermal conductivity, and H is the local heat generation rate per unit volume. In the steady-state case, the general heat conduction equation reduces to

$$\nabla \cdot (k \nabla T) + H = 0. \tag{11}$$

The thermal conductivity k is kept inside the divergence operator to account for anisotropic and temperature-dependent thermal conductivity relationships. The classical heat conduction equation, which is derived from energy conservation for a continuous medium and which makes use of Fourier's law, is only valid for thermal transport in the diffusive regime. Heat generation in a semiconductor device is primarily due to the three mechanisms of Joule heating, carrier recombination and generation heating and cooling, and thermoelectric (Peltier and Joule-Thomson) effects. Positive and negative values of H in the lattice heat conduction equation correspond to energy being added to or removed from the lattice, respectively. In GaN HEMTs, the dominant heat generation mechanism is Joule heating due to the very high current density and electric fields in the channel. Carrier generation and recombination rates in electronic devices are typically low enough to make the heat generation associated with generation and recombination negligible. Thermoelectric effects are usually neglected in electro-thermal modeling of GaN HEMTs, although this has not been thoroughly investigated [51]. In general, the local heat generation rate per unit volume is given by

$$H = \frac{\left|\vec{J}\right|^2}{q\mu n} + q(G - R)[\phi + TP] - T\left(\vec{J}\nabla P\right) \tag{12}$$

where ϕ is the quasi-Fermi potential and P is the thermoelectric power (or Seebek coefficient). The first, second, and third terms refer to Joule heating, generation and recombination heating and cooling, and thermoelectric heating and cooling, respectively. It is interesting to note that the Joule heat generation term

$$H = \frac{\left|\vec{J}\right|^2}{q\mu n} \tag{13}$$

is always positive because the numerator is the magnitude of the current density vector squared and the variables in the denominator are always positive by definition. This relation comes from the rigorous analysis provided by *Wachutka (1990)* [52] apparently under the assumption of Boltzmann statistics. The fact that the heat generation must always be positive implies that carrier transport is always a dissipative mechanism that transfers energy from the electron system to the lattice.

However, many device models make use of the relation

$$H = \vec{J} \cdot \vec{E} \tag{14}$$

as specified by the ATLAS user manual [50] and given as the leading term of the heat generation rate by *Lindefelt (1994)* [20]. Because current density and electric field are both vector quantities, it is possible to have a situation in which the dot product is negative. This situation corresponds to energy being transferred from the lattice to the electron system and occurs in pn junctions as described in *Lindefelt (1994)*. In this case and others, electron transport is primarily by diffusion in the direction opposite to the electric field because of a large gradient in electron concentration created by the intentional doping profile. For steady-state electro-thermal simulations in ATLAS, the simpler form of the heat equation is used by default. In the case of AlGaN HEMTs, this leads to some locations in which the heat generation rate is negative in the region around the gate. In this work, this was not observed to have a strong effect on the overall temperature distribution.

The general form of the steady-state heat generation rate is invoked by the expression "HEAT.FULL" in ATLAS and the individual terms associated with the various mechanisms can be enabled or disabled by the appropriate expressions. In this work, only Joule heating (always positive) was enabled and the terms associated with generation and recombination and thermoelectric effects were disabled due to convergence issues. There may be an inconsistency in the way that the thermoelectric power is defined in ATLAS with respect to the carrier concentration when Fermi-Dirac statistics are enabled that leads to numerical instabilities in the region around the gate. The ATLAS software essentially solves the equations for electron transport and the heat conduction equation with block iteration or Newton algorithms until the specified convergence criteria are met for the electron concentration, electrostatic potential, and temperature.

2.2.3 Electrical and Thermal Boundary Conditions

Boundary conditions on the equations for electron and thermal transport are an important consideration in device-level electro-thermal modeling. The solutions to the equations are strongly dependent on the boundary conditions used and many of the important results from the simulations (*e.g.*, the drain current) are quantities computed at the boundary.

For the electron concentration, only Dirichlet boundary conditions (prescribed concentration) are used at the ohmic contacts, Schottky contacts, and insulated boundaries (everywhere else). For ohmic contacts such as the drain and source the concentration at the contact boundary is set equal to the local donor concentration, *i.e.* $n_s = N_D^+$ where n_s is the surface electron concentration and

N_D^+ is the donor concentration. The boundary condition on electron concentration is fixed at zero for Schottky contacts. On all of the other boundaries that are insulated or in contact with a dielectric material, the surface electron concentration is set equal to zero.

For the surface potential of ohmic contacts with no contact resistance, the electrostatic potential is fixed with the Dirichlet boundary condition

$$\psi_s = \phi + \frac{k_B T}{q} \ln \frac{n_s}{n_{ie}} \tag{15}$$

in which the quasi-Fermi potential ϕ is set equal to the applied bias voltage of the ohmic contact. When the ohmic contact has a distributed contact resistance, the condition is modified to

$$\frac{1}{R_c}\left[V_{applied} - \left(\psi \pm \frac{k_B T}{q} \ln \frac{N}{n_{ie}}\right)\right] - I = 0 \tag{16}$$

where R_c is the contact resistance at each node (Ω-cm), $V_{applied}$ is the applied voltage at the contact, N is the net doping level, and I is the current through the contact. This equation for the boundary condition relates the voltage drop across the contact resistance to the current through the contact. The surface potential of a Schottky contact is given by

$$\psi_s = \phi_B + \frac{E_g}{2q} + \frac{k_B T}{2q} \ln \frac{N_C}{N_V} + V_{applied} \tag{17}$$

where ϕ_B is the Schottky barrier height, E_g is the bandgap of the material in contact with the Schottky contact, and N_C and N_V are the effective density of states of the conduction and valence bands, respectively. For the remainder of the boundary that is in contact with dielectric materials or on the outer boundary, Neumann conditions are specified on the electrostatic potential that are consistent with zero current flow.

Silvaco ATLAS can implement Dirichlet (prescribed temperature), homogeneous Neumann (zero heat flux), and Robin (mixed or finite conductance) boundary conditions for the heat conduction equation. In this work, the bottom of the substrate was set to a fixed temperature with a Dirichlet condition and the rest of the boundaries were set to be thermally insulating. ATLAS often has convergence problems when implementing Robin type boundary conditions to model convection heat transfer at one of the boundaries or a finite conductance to a temperature sink.

2.3 Test Device Structure

A test device structure was fabricated on a commercially-available GaN-on-SiC wafer by Omair Saadat (Professor Tomas Palacios' research group, Department of Electrical Engineering and Computer Science, MIT) to provide validation for the electro-thermal model. The test device was a 2 x 150 μm HEMT with a gate length of 0.3 μm and other dimensions typical of HEMTs for RF power amplifier applications. This device, in particular, was fabricated with two independent, *in-situ* Schottky diodes parallel to the source contacts that were used to measure the surface

temperature while the device was under test [53]. Measurement data was collected for two different devices on the wafer, 1L3B and 3L3B, as designated by their position on the wafer. Most of the data presented in this thesis was for 1L3B, which showed better performance than 3L3B. A schematic of the device epitaxial structure is shown in Figure 11 along with a top view optical microscope image of the final fabricated device.
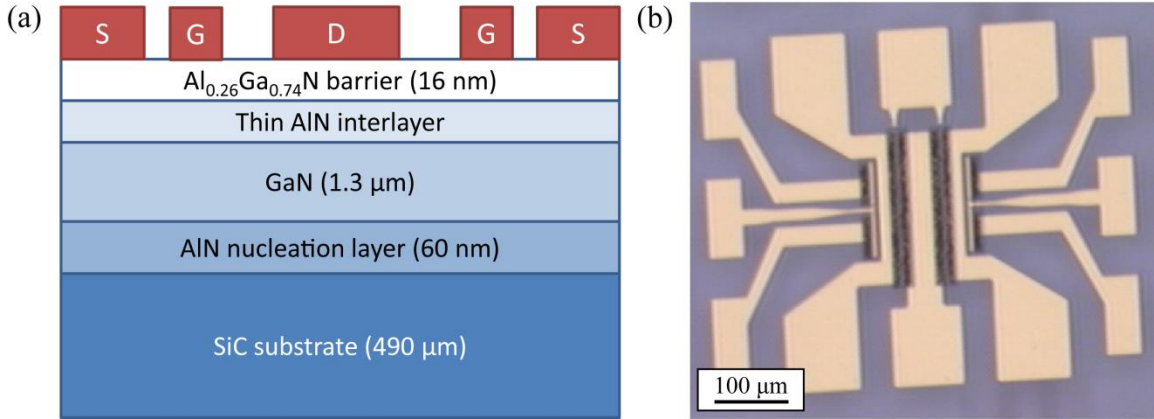
Figure 11: Test device structure for HEMT electro-thermal model
(a) Schematic of epitaxial structure and (b) top side optical microscope image of fabricated device [53]

The epitaxial structure consists of a 490 μm high purity semi-insulating (HPSI) SiC wafer, 60 nm AlN nucleation layer, 1.3 μm GaN device layer, a thin (~1 nm) AlN interlayer, and a 16 nm $Al_{0.26}Ga_{0.74}N$ barrier layer. The AlN interlayer was introduced by the wafer supplier to increase the 2DEG concentration and improve carrier confinement to the 2DEG. The sheet resistance, electron mobility, and carrier concentration of the 2DEG were found to be 298 Ω/sq., 2038 cm^2/V-s and 1.03 x 10^{13} cm^{-2}, respectively, from Hall measurements. The device was probed with four DC probes and two RF probes to simultaneously measure the HEMT characteristics and voltage of the *in-situ* diodes. In the electro-thermal model, a 1000 μm thermal spreading region was added to the side of the device to simulate the large spreading area around the device on the wafer piece which was much larger than the device itself.

2.4 Material Properties and Transport Models

Because GaN-based electronics are a relatively less mature technology than Si- and GaAs-based electronics, the material properties and transport parameters of GaN and its related alloys are less well-characterized. In this thesis, only a summary of the relevant material properties and transport models will be presented as it pertains to the electro-thermal modeling work described herein. *Quay (2008)* [53] serves as a good reference for the mechanical, electrical, optical, thermal, and transport properties of GaN and other III-N semiconductors.

The bandgap of GaN (3.42 eV) and AlN (6.13 eV) were set equal to values obtained from experimental results in the literature [54]-[55]. An alloy bowing parameter of unity was used to compute the bandgap of $Al_xGa_{1-x}N$ with the relation

$$E_g(x) = xE_g(AlN) + (1-x)E_g(GaN) - x(1-x) \qquad (18)$$

where E_g is the bandgap and x is the AlN mole fraction. For the $Al_{0.26}Ga_{0.74}N$ alloy used in the barrier layer in the validation study, the bandgap of AlGaN is 3.93 eV as calculated from the equation above. The alignment of the conduction band at the AlGaN/GaN heterojunction was set using the "ALIGN" parameter rather than the electron affinity rule. The conduction band offset was specified to be 70% of the difference between the bandgap values of AlGaN and GaN, *i.e.,*

$$\Delta E_c = 0.7\big[E_g(Al_xGa_{1-x}N) - E_g(GaN)\big] \tag{19}$$

where ΔE_c is the conduction band offset (calculated to be 0.357 eV for $x = 0.26$) [55]. This is likely more accurate than specifying the electron affinity values for AlGaN and GaN. As a note, the electron affinity for GaN was taken as the ATLAS default (4.31 eV) and the electron affinity of $Al_{0.26}Ga_{0.74}N$ was calculated to be 3.98 eV.

Isotropic values of the dielectric constant of GaN (9.5) and AlN (8.5) were taken from *Ambacher et al. (1999)* [55] and the dielectric constant of $Al_xGa_{1-x}N$ was calculated with the simple linear relationship

$$\epsilon(x) = -0.5x + 9.5 \tag{20}$$

where ϵ is the dielectric constant as a function of Al mole fraction. For the specific alloy $Al_{0.26}Ga_{0.74}N$, the dielectric constant was set equal to 9.24. There is some experimental evidence that the dielectric constant of GaN and AlN is orthotropic with slightly different values in the in-plane and along the c-axis [53], but this was not found to have a significant effect on the electron transport for the model. For the value of the dielectric constant for all other materials, default values were used.

The Schottky barrier height of the gate contact was set through the work function parameter in ATLAS according to the definition

$$\phi_m = \chi + \phi_B \tag{21}$$

where ϕ_m is the work function of the metal and χ is the electron affinity of the semiconductor in contact with the metal [5]. As previously stated, the electron affinity of the $Al_{0.26}Ga_{0.74}N$ barrier layer in contact with the metal was calculated by ATLAS to be 3.98 eV from the conduction band offset rule. The Schottky barrier height was measured experimentally for the 1L3B test device from the forward voltage characteristics of the *in-situ* Schottky diodes adjacent to the HEMT. The temperature was controlled by the probe station chuck. The forward voltage characteristics of the diode are shown in Figure 12 in a linear plot and semilog plot. The properties of the gate contact should be the same as these Schottky diodes because they were formed during the same fabrication step but are simply used for different purposes.
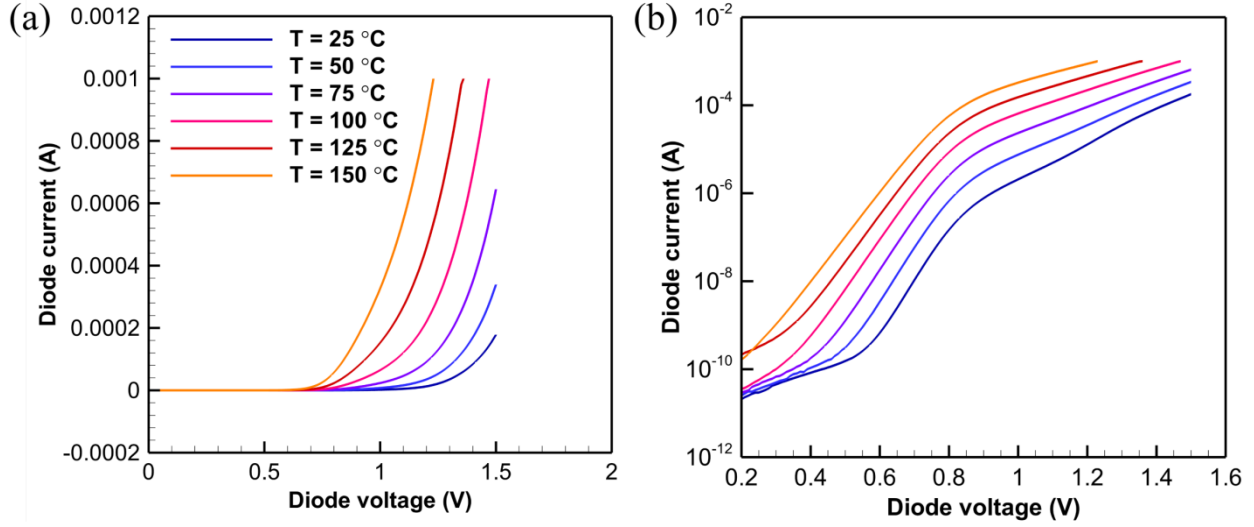
Figure 12: Forward current-voltage characteristics of the in-situ Schottky diodes for 1L3B
(a) Linear and (b) semilog axes

The semilog plot of the forward voltage characteristics of the diodes shows some non-idealities that may be due to poor surface condition at the metal-AlGaN interface. The relatively straight region on the semilog plot of the forward bias characteristics was used to calculate the saturation current. For a Schottky diode governed by thermionic emission, the current is related to the voltage by

$$I = I_s \left[\exp\left(\frac{qV}{nk_BT} \right) - 1 \right] \tag{22}$$

where I is the diode current, I_s is the saturation current, V is the diode voltage, n is the ideality factor, T is the absolute temperature in kelvins [56]. Taking the base ten log of each side yields

$$log_{10}I = \frac{qV}{nk_BT} - log_{10}I_s \tag{23}$$

which is seen in the semilog current-voltage characteristics in the figure above. Thus, the diode saturation current I_s is the intercept of the vertical axis in the semilog figure above. The saturation current is related to the Schottky barrier height of the diode through the equation

$$I_s = A_{eff}A^{**}T^2 \exp\left(\frac{-q\phi_B}{k_B} \right) \tag{24}$$

where A_{eff} is the effective area of the diode and A^{**} is the Richardson constant (8.6 A/cm^2-K^2). One can rearrange to solve for the Schottky barrier height in terms of the other variables

$$\phi_B = -\frac{k_B}{q}\ln\left(\frac{I_s}{A_{eff}A^{**}T^2} \right). \tag{25}$$

The saturation diode was determined from the vertical axis intercept of a linear regression of the semi-log current-voltage measurement data. Then, the Schottky barrier height was calculated from the previous formula taking the effective area of the diode to be the nominal area. The values of the Schottky barrier height for different diode temperatures is shown in the table below.

Table 1: *In-situ* diode Schottky barrier height as a function of temperature

Temperature (°C)	Temperature (K)	Schottky barrier height (eV)
25	298.15	1.05
50	323.15	1.07
75	348.15	1.12
100	373.15	1.12
125	398.15	1.13
150	423.15	1.14

The Schottky barrier height of the diode was found to have a weak dependence on temperature. However, since this parameter is fixed in ATLAS and cannot be changed with device temperature, the value of 1.05 eV at a temperature of 25 °C was chosen for the electro-thermal models. The appropriate value of the gate metal work function is therefore 5.03 eV as given by the previous equation. This value is slightly lower with the work function for typical Ni/Au/Ni Schottky contacts in AlGaN/GaN HEMTs reported in the literature [57].

Polarization charges play a major role in the high electron density present in the 2DEG in GaN-based electronics. The discontinuity in polarization at the AlGaN/GaN heterojunction due to spontaneous and piezoelectric polarization effects leads to a polarization sheet charge given by

$$\sigma = P_{top} - P_{bottom} \tag{26}$$

where σ is the polarization sheet charge and P_{top} and P_{bottom} are the total polarizations of the materials on the top and bottom sides of the interface [55]. The spontaneous polarization of AlN and GaN are material properties that should be true for all samples. For the ternary alloy $Al_xGa_{1-x}N$, the spontaneous polarization can be calculated as a function of the mole fraction from a linear interpolation between the respective values for AlN and GaN, *i.e.,*

$$P_{sp} = -0.052x - 0.029 \tag{27}$$

where P_{sp} is the spontaneous polarization in C/m^2 [55]. In addition to spontaneous polarization, AlN and GaN exhibit an additional polarization when strained due to the piezoelectric effect. The piezoelectric polarization is related to the lattice constants, piezoelectric properties and elastic properties through the expression

$$P_{pe} = 2\frac{a - a_0}{a_0}\left(e_{31} - e_{33}\frac{C_{13}}{C_{33}}\right) \tag{28}$$

where a_0 and c_0 are the equilibrium in-plane and c-axis lattice constants, respectively, a and c are the lattice constants when the material is under strain, e_{31} and e_{33} are piezoelectric properties and C_{13} and C_{33} are elastic constants [55]. When the AlGaN barrier layer (20 to 30 nm thick) is grown epitaxially on the GaN buffer (1 to 2 µm thick), AlGaN is placed under tensile strain and the GaN buffer is in equilibrium. The polarization sheet charge at the AlGaN/GaN interface can be calculated from the equation

$$|\sigma(x)| = |P_{pe}(Al_xGa_{1-x}N) + P_{sp}(Al_xGa_{1-x}N) - P_{sp}(GaN)|. \qquad (29)$$

Using this equation and the material properties given by *Ambacher et al. (1999)* [55], the value of the polarization charge should be $\sigma = 1.44 \times 10^{13}$ C/m^2 for an AlGaN barrier with no strain relaxation. The use of this value, however, in the electro-thermal model led to a threshold voltage much more than negative than the measured value of -2.3 V for device 1L3B. If the experimentally measured Schottky barrier height and the band parameters taken from the literature are kept fixed, the polarization charge is the best parameter to modify in order to obtain a threshold voltage close to the measured value for the device. This may be due to strain relaxation in the AlGaN layer, which decreases the contribution of spontaneous polarization at the AlGaN/GaN interface. If one assumes the strain in the AlGaN barrier is totally relaxed, the polarization sheet charge should be $\sigma = 8.44 \times 10^{12}$ C/m^2. For the electro-thermal model described in this work, a value of $\sigma = 8.5 \times 10^{12}$ C/m^2 was chosen to provide the best agreement in the threshold voltage, 2DEG concentration, and device output characteristics.

In addition to the positive polarization sheet charge that exists at the AlGaN/GaN interface, there should also be a negative sheet charge at the interface between AlGaN and the passivation layer around the gate. This occurs because of the discontinuity in polarization between AlGaN (high polarization) and the passivation material (near zero polarization). Some previous electro-thermal models make use of a negative polarization charge at the AlGaN/passivation interface while others only use a positive sheet charge at the AlGaN/GaN interface [51]. For the alloy Al$_{0.26}$Ga$_{0.74}$N, the negative polarization sheet charge should be $\sigma = -3.25 \times 10^{13}$ C/m^2 (fully strained) or $\sigma = -2.65 \times 10^{13}$ C/m^2 (fully relaxed). However, when these values of negative sheet charge were added to the ATLAS model at the AlGaN/passivation interface, they led to a significant drop in electron concentration in the 2DEG, resulting in a change in resistance and threshold voltage. It is unclear whether this negative sheet charge is neutralized or compensated for in actual AlGaN/GaN HEMTs so that the effective negative sheet charge at the AlGaN/passivation interface is much less in magnitude or near zero. The best fit between the experimental and model output characteristics in this work was achieved with a negative sheet charge of $\sigma = -3.5 \times 10^{12}$ C/m^2 at the AlGaN/passivation interface.

Bulk acceptor type traps were introduced into the GaN layer to model the high density of electron traps frequently present in the GaN layer and to limit the conduction of electrons far away from the AlGaN/GaN heterojunction. The traps were assigned an energy level of 2.20 eV below the conduction band with electron and hole lifetimes of 10^{-12} s and 2×10^{-11} s, respectively. The volumetric density of traps in the GaN layer was set with an exponential dependence on distance away from the AlGaN/GaN heterojunction according to the equation

$$\rho(y) = \rho_0 \left[\exp\left(\frac{y - 0.033}{1.86962} \right) - 1 \right] \tag{30}$$

where ρ is the trap density in cm^{-3}, $\rho_0 = 9 \times 10^{17}$ cm^{-3} is the trap density at the bottom of the GaN layer, and y is the distance below the AlGaN/GaN heterojunction in μm. The value of $\rho_0 = 9 \times 10^{17}$ cm^{-3} was chosen during the process of adjusting the model parameters so as to make the bottom of the GaN layer ($y = 1.325$ μm) highly electrically resistive and the constants in the exponential term were chosen to make the trap density equal to zero at a distance of 0.033 μm below the AlGaN/GaN heterojunction. This causes the acceptor traps to have no effect on conduction in the 2DEG and little effect on conduction in the first ~50 nm of the GaN layer as desired.

The contact resistance of the ohmic contacts (source and drain) is also an important parameter in electro-thermal modeling of AlGaN/GaN HEMTs. This parameter can have a strong effect on the slope of the linear region of the output characteristics, especially when the contact resistivity is relatively high. The contact resistivity can be measured by the transmission line method (TLM) with additional ohmic pads fabricated on the wafer [58]. The contact resistivity of the TLM structures on devices 1L3B and 3L3B were determined to be 3.73×10^{-4} Ω-cm^2 and 3.97×10^{-4} Ω-cm^2, respectively. The best agreement between the modeled and measured output characteristics were achieved when using contact resistivity values of ~1.5×10^{-4} Ω-cm^2. Due to the fact that ATLAS ignores the contact resistivity of metal-semiconductor contacts in computing the rate of heat generation, a thin (1 nm) region with a very low electron mobility (~10^{-3} cm^2/V-s) was added to the AlGaN region around the source and drain contacts. The electron mobility of this resistive region was varied slightly to achieve the best fit between measured and modeled output characteristics.

Selecting the proper mobility model for AlGaN/GaN HEMTs is particularly challenging because there are so many mobility models available in the literature and because some modification of the parameters is often needed to achieve good agreement between measured and modeled device characteristics. In this work, the Farahmand Modified Caughey-Thomas mobility model was used for the low field electron mobility dependence on temperature and doping level for the AlGaN barrier and GaN buffer regions given by

$$\mu_0(N,T) = \mu_{min} \left(\frac{T}{300} \right)^{\beta_1}$$
$$+ (\mu_{max} - \mu_{min}) \left(\frac{T}{300} \right)^{\beta_2} \left\{ 1 + \left[\frac{N}{N_{ref} \left(\frac{T}{300} \right)^{\beta_3}} \right]^{\alpha(T/300)^{\beta_4}} \right\}^{-1} \tag{31}$$

where μ_0 is the low field mobility, N is the doping concentration, T is the absolute temperature, and α, μ_{max}, μ_{min}, N_{ref}, and $\beta_1 \ldots \beta_4$ are parameters obtained from Monte Carlo simulations [59]. The doping level for the AlGaN and GaN device layers was not known exactly and was set

equal to a nominal value of 10^{15} cm^{-3}. The high field mobility with its dependence on electric field was calculated using the Canali model given by

$$\mu(N, T, E) = \mu_0(N, T) \left[\frac{1}{1 + \left(\frac{\mu_0(N,T)E}{v_{sat}}\right)^\beta} \right]^{1/\beta} \tag{32}$$

where μ is the high field electron mobility, E is the electric field magnitude, v_{sat} is the electron saturation velocity, and β is a parameter from the Canali model [50]. The parameter β also varies with temperature according to the expression

$$\beta = \beta_0 \left(\frac{T}{300}\right)^{\beta_{exp}} \tag{33}$$

where β_0 and β_{exp} are fitting parameters of the mobility model with different values for each material. For the electro-thermal simulations in this work, values of $\beta_0 = 1.7$ and $\beta_{exp} = 0$ were used from the reference *Wang et al. (2012)* [60]. The saturation velocities of AlGaN and GaN were set to constant (temperature-independent) values of 7.6 x 10^6 cm/s and 2.0 x 10^7 cm/s, respectively.

The electron mobility of the 2DEG was set using the constant low field mobility model in ATLAS given by

$$\mu_0 = \mu_{300} \left(\frac{T}{300}\right)^\alpha \tag{34}$$

where μ_{300} is the low field mobility at a temperature of 300 K and α is a negative exponent representing the reduction of electron mobility due to increased electron-phonon scattering at high temperatures. The low field mobility at 300 K was taken to be $\mu_{300} = 2038$ cm^2/V-s from Hall effect measurements on the wafer performed at the ambient. The exponent parameter was set equal to the value $\alpha = 1.9$, which is lower than a reported value in the literature of $\alpha = 2.4$ [61], to achieve the best agreement between measured and modeled output characteristics in the saturation region. The high field mobility of the 2DEG was computed with the Canali model with the same parameters as given above for the GaN buffer. The saturation velocity of the 2DEG was set to $v_{sat} = 2.57$ x 10^7 cm/s based on Monte Carlo simulations of electron transport in AlGaN/GaN heterostructures [62].

To achieve the best agreement between measured and modeled output characteristics, the low field mobility of the ohmic contact resistance regions was varied and finally set to a value of $\mu_0 = 3$ x 10^{-3} cm^2/V-s independent of temperature. The electric field dependence of mobility was calculated using the Caughey-Thomas expression for velocity saturation

$$\mu(N,T,E) = \mu_0(N,T)\left[\cfrac{1}{1 + \left(\cfrac{\mu_0(N,T)E}{v_{sat}}\right)^{\beta}}\right]^{1/\beta} \tag{35}$$

where all of the variables are the same as for the Canali model and β is a constant [50]. The value of $\beta = 5$ was chosen to make the electron mobility of this resistive region weakly dependent on the magnitude of the local electric field.

Other important features of the models for carrier transport in ATLAS included the Shockley-Read-Hall model for carrier generation/recombination associated with traps in the bulk GaN buffer and Fermi-Dirac statistics to account for very high carrier concentrations ($\sim 10^{20}$ cm^{-3}) in the 2DEG. For all other electrical and mobility transport parameters not mentioned, the default values were used.

The thermal conductivity of the epitaxial layers and substrate plays a major role in the physics of self-heating in AlGaN/GaN HEMTs. For many crystalline solids, such as wurtzite semiconductors, the thermal conductivity decreases with increasing temperature because of increased phonon-phonon scattering rates according to the relation

$$k(T) = k_0 \left(\frac{T_0}{T}\right)^n \tag{36}$$

where T_0 is an absolute reference temperature (typically 300 K), k_0 is the thermal conductivity at the reference temperature T_0, T is the absolute temperature, and n is a positive number [63]. The theory of phonon transport in semiconductors predicts that $n = 1$ in the range above the ambient temperature (where phonon-phonon scattering dominates) but experiments show $1 \le n \le 1.5$. For the GaN buffer, the following commonly used expression for the thermal conductivity of GaN in the literature was used

$$k_{GaN}(T) = 150 \left(\frac{300}{T}\right)^{1.4} \tag{37}$$

where k_{GaN} is in W/m-K [64]. The thermal conductivity of SiC has been found to be slightly orthotropic, varying in the in-plane and cross-plane directions, via experimental measurements by a major supplier of commercial GaN-on-SiC substrates

$$k_{SiC,a}(T) = 490 \left(\frac{300}{T}\right)^{1.3} \tag{38}$$

$$k_{SiC,c}(T) = 390 \left(\frac{300}{T}\right)^{1.3} \tag{399}$$

where $k_{SiC,a}$ and $k_{SiC,c}$ are the thermal conductivity of SiC in the in-plane (a-axis) and cross-plane (c-axis) directions [65]. The temperature dependence in the expression above was taken from *Sarua et al. (2007)* [64].

The AlN nucleation layer between the GaN buffer and SiC substrate represents the thermal boundary or thermal interfacial resistance (TBR or TIR) between GaN and SiC for the GaN-on-SiC epitaxy. Since the thermal conductivity of the 60 nm thick AlN nucleation layer was not measured directly in this work, an estimated value was chosen from benchmarking measurements of *Manoi et al. (2010)* [33] with the temperature-dependence found experimentally by *Sarua et al. (2007)* [64]. Thus, the thermal conductivity of the AlN nucleation layer was set equal to

$$k_{AlN}(T) = 6\left(\frac{300}{T}\right)^{2.76} \tag{40}$$

based on a typical value of GaN-SiC TBR of 3 x 10^{-8} m^2-K/W measured at 175 °C. The thermal conductivity of the 16 nm $Al_{0.26}Ga_{0.74}N$ barrier was set equal to 65 W/m-K based on experimental measurements by *Liu and Balandin (2005)* [66] for AlGaN films grown by MOCVD on sapphire substrates and measured with the differential 3ω technique. Unfortunately, the samples measured by *Liu and Balandin (2005)* of thicknesses 0.3 to 0.7 µm were much thicker than typical HEMT barrier layers. Therefore, one would expect the thermal conductivity of an AlGaN barrier layer to be much less than the values reported in that paper due to classical size effects [63]. This overestimated value of 65 W/m-K for the AlGaN barrier is not likely to have a strong effect on the device temperature predicted by the model because the dominant heat sinking path is through the substrate rather than through the top of the device. For all other materials not mentioned, such as Al_2O_3 and Au, the thermal conductivity was taken to be the ATLAS default.

2.5 Electrical Characteristics and Device Temperature

As previously noted, some adjustment of the material properties and transport parameters were made until a good agreement was obtained between the modeled and measured HEMT device characteristics. The main parameters which were adjusted were the effective contact resistivity of the ohmic contacts, the magnitude of the polarization sheet charges, the bulk GaN trap density, and the exponent of the temperature-dependent relation for low field mobility of the 2DEG. As many as possible of the parameters were measured independently or taken from the literature without adjustment to maximize the robustness of the model. The modeled and measured electrical output characteristics for the test device 1L3B are shown in Figure 13.
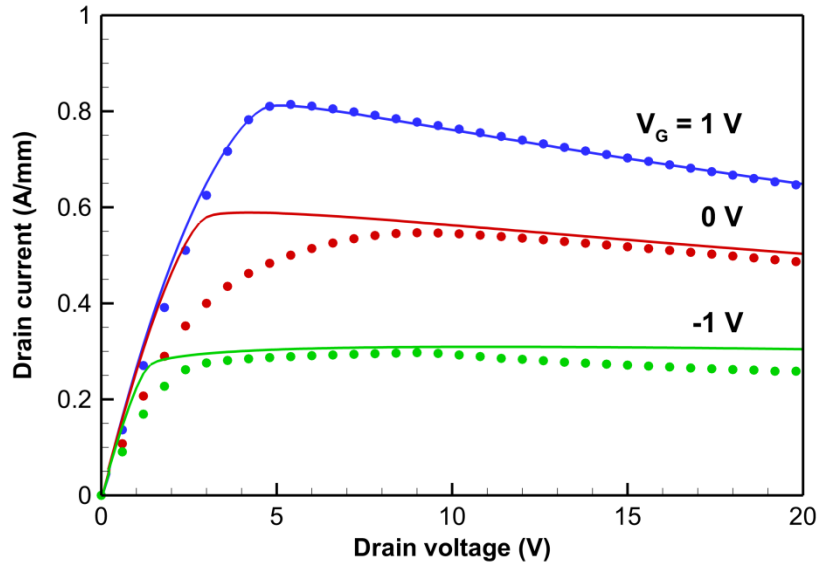
**Figure 13: Measured and modeled test device electrical output characteristics
Circles represent measurements and solid lines represent model curves**

One of the key features in this electro-thermal modeling study was to experimentally measure the device temperature while the device was under test in order to validate the temperature distribution predicted by the model. The electrical parameters were adjusted until a good agreement was obtained for the electrical output characteristics; then, the temperature predicted by the model was compared to the experimental temperature measurements without any further adjustment of the parameters. As the drain, source, and gate pads were probed with standard DC probes, the two in-situ diodes alongside the source contacts were probed with RF probes. The forward current-voltage characteristics has been shown previously in the section describing the Schottky barrier height for the metal-AlGaN contacts. The diodes were biased with a 1 μA current source and the voltage across the diodes was recorded while the HEMT electrical characteristics were measured. The transient voltage readings for the in-situ diodes on device 1L3B are shown in Figure 14 for a current bias of 1 μA at a temperature of 25 °C. The temperature was controlled via the probe station measurement chuck.
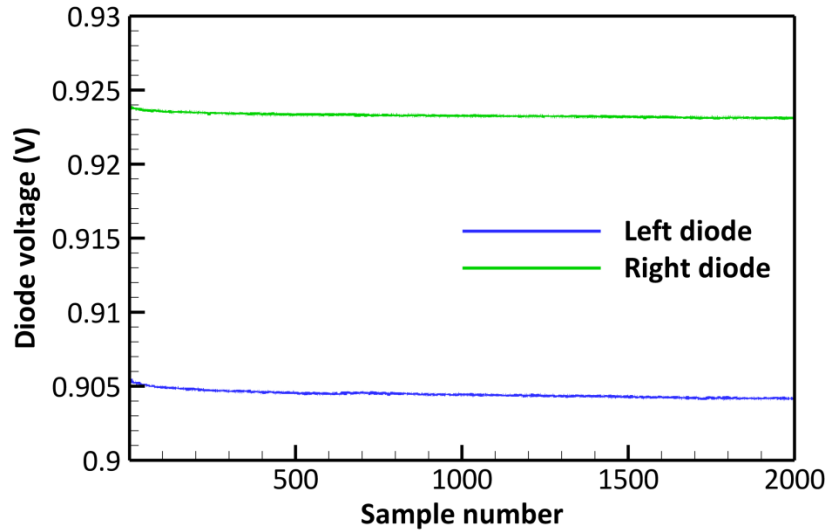
Figure 14: Transient diode voltage readings with 1 µA current bias at 25 °C

As can be seen in Figure 14, the voltage read at the diodes decreased sharply at the beginning of the measurement and then stabilized to a nearly constant value with some observable noise in the signal. In order to determine the relationship between diode voltage and device temperature, the diode voltage was averaged over the last 100 samples for each of the diodes at temperatures of 25 °C to 150 °C in 25 °C increments. It is interesting to note that the readings between the left and right diodes of the same device differed by approximately 20 mV, indicating some variation in properties due to non-idealities in the fabrication process. The voltage-temperature calibration curve was constructed for each of the diodes under test and is shown in Figure 15 for device 1L3B.
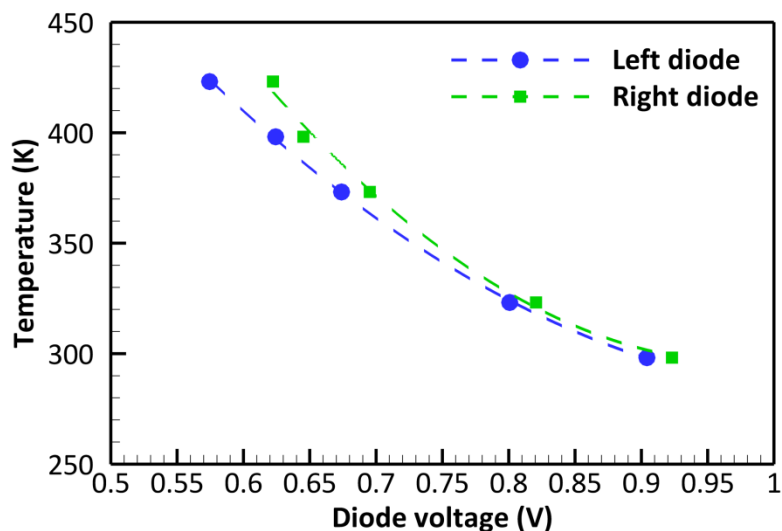
Figure 15: Temperature-voltage calibration curve for 1L3B *in-situ* diodes

The voltage-temperature relationship of a pn junction diode can be determined from the current-voltage relationship based on carrier diffusion

45

$$V(T) = A + \frac{k_B T}{q} \ln \left(\frac{1}{BT^C} \right) \tag{41}$$

where V is the voltage across the diode, T is the absolute temperature, and A, B, and C are constants [5]. Since the term involving the natural logarithm is weakly dependent on temperature, the diode voltage varies approximately linearly with temperature. A similar relationship can be derived for Schottky diodes in which the primary mechanism for carrier transport is due to thermionic emission. In this work, the temperature-voltage diode calibration curves were found to be non-linear and best fit by a quadratic equation. A quadratic regression was performed on the diode calibration curve and the actual temperature measured while the HEMT was under voltage bias was determined from the diode voltage readings and this calibration curve. The measured diode and model temperatures were compared by averaging the temperature predicted by the left and right diodes and sampling the temperature at the center of the anode location in the electro-thermal model. The measured and modeled temperature at the diodes for device 1L3B is shown in Figure 16.
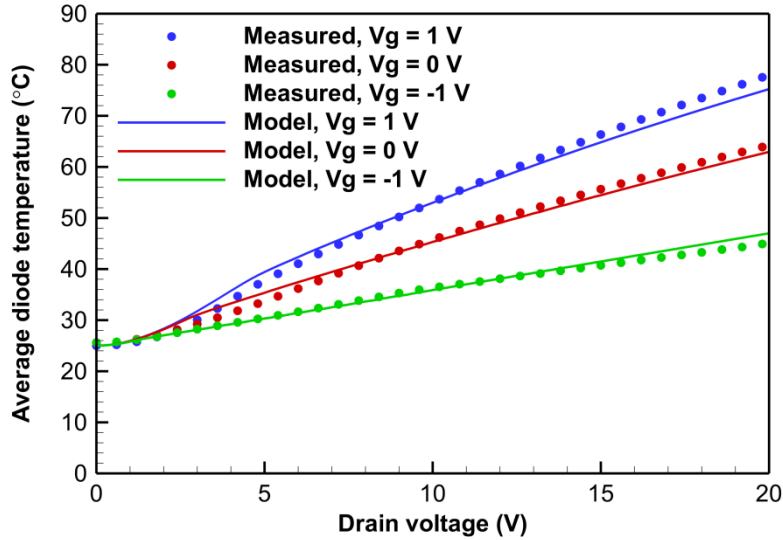
Figure 16: Measured and modeled *in-situ* diode temperature

A reasonably good agreement between the measured and modeled diode temperature at the surface of the device was obtained, indicating that the temperature distribution in the electro-thermal model is accurate. In future work, it would be helpful to improve the reliability of the measurement by improving the fabrication process and the quality of the diodes.

The 2D temperature distribution from the electro-thermal model for the entire 1L3B HEMT at $V_G = 0$ V and $V_D = 20$ V ($P_{loss} = 10.1$ W/mm) is shown in Figure 17. Because the double finger HEMT is left-right symmetric, only half of the device was modeled to reduce the computational expense and a thermally insulated (adiabatic) boundary condition was used to represent the symmetry of the temperature distribution.
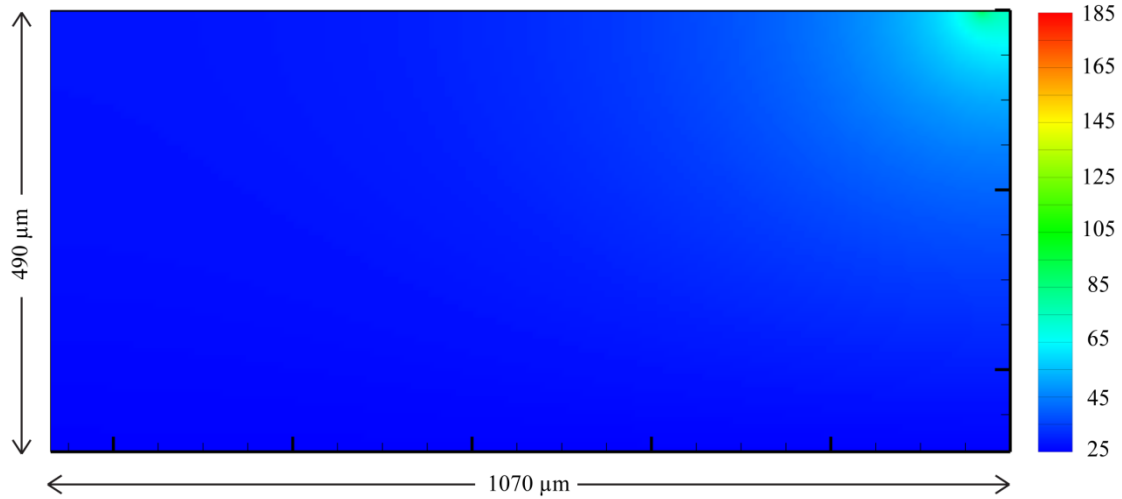
Figure 17: Electro-thermal model 2D temperature (°C) distribution for 1L3B (V_G = 0 V and V_D = 20 V)

As can be seen from the temperature distribution in Figure 17, the peak temperature is approximately 185 °C for a fixed base temperature of 25 °C, the temperature rise is highly localized to the active region. This kind of temperature distribution is indicative of a multi-dimensional conduction situation in which the thermal spreading resistance dominates over the 1D conduction resistance. A closer view of the temperature distribution from the electro-thermal model at V_G = 0 V and V_D = 20 V is shown for the active region in Figure 18.
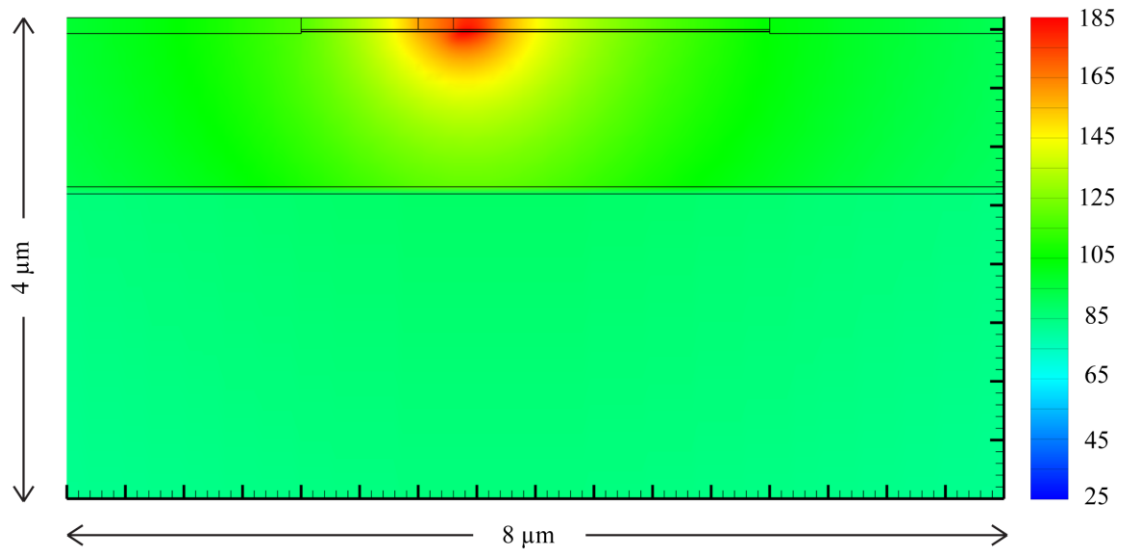
Figure 18: Closer view of the electro-thermal model 2D temperature (°C) distribution for 1L3B near the active region of the HEMT

The peak temperature rise is highly localized at the edge of the gate on the drain side due to the intense heat generation present at that position, as will be discussed in more detail in the following section. The temperature changes by ~70 °C or more within a few microns of the gate because of the small area over which most of the heat is generated and the high thermal spreading resistance in the GaN layer. The low thermal conductivity of the AlN nucleation layer

47

(~6 W/m-K at a temperature of 300 K) further inhibits the flow of heat away from the active region.

2.6 Quantitative Analysis of the Heat Source

There have been many reports in the literature in the past of electro-thermal modeling of AlGaN/GaN HEMTs based on drift-diffusion equation, electrohydrodynamic equations, or Monte Carlo methods. To report the temperature distribution from electro-thermal modeling is not original in itself nor as accurate as some researchers might believe. Most fully-coupled electro-thermal models are restricted to 2D because of the computational expense involved in simultaneously solving the electron and thermal transport equations. Thus, the 2D temperature distribution predicted by electro-thermal models of GaN HEMTs is usually an over-estimate of the real temperature because of the inability of the heat to spread in the direction along the gate width in a 2D model. The actual temperature of the device and the 3D modeled temperature should both be less than that predicted by the 2D model.

However, there is very valuable insight to be gained from electro-thermal models with respect to the heat source distribution associated with Joule heating in GaN HEMTs. Because the thermal spreading resistance and device temperature rise is strongly dependent on the size and shape of the heat source, accurately capturing the heat source is important in analyzing thermal issues in GaN-based electronics. Since the 1970s, researchers in the field of thermal issues in electronics have often assumed that the heat source in high power FETs can be modeled as a uniform heat flux or volumetric heat generation under the gate [67]. This assumption has also often been adopted for GaN HEMTs in analytical and numerical thermal models [16],[23]. Although this may provide a reasonable estimate for moderate gate length devices ($L_G \sim 1$ μm), there is no fundamental, physical reason why the heat source should scale with the gate length. Also, for short gate length devices ($L_G \sim 0.5$ μm), the heat source may not necessarily be much smaller than for a device with $L_G \sim 1$ μm. *Heller and Crespo (2007)* [22] made reference to this point when comparing their electro-thermal modeling results to the analytical model developed by *Darwish et al. (2004)* [23]. *Heller and Crespo (2007)* noted that the temperature rise and thermal resistance had a weak dependence on the gate length as the gate length was varied from 0.1 to 0.8 μm. The non-uniformity it the heat generation (or power dissipation) distribution was reported by *Webb and Russell (1989)* [68] and discussed briefly by *Anholt (1995)* [69] for GaAs MESFETs. In GaN HEMTs, the non-uniformity in the heat generation distribution is particularly important because of the very high power density enabled by the excellent electrical properties of GaN.

The literature is inconsistent with respect to the manner in which researchers choose the width of the heat source in thermal models. As already noted, *Darwish et al. (2004)* [23] and others used the assumption that the heat source is rectangular and equal in length to the gate. In order to demonstrate the effectiveness of few layer graphene heat spreaders in GaN HEMTs, *Yan and Balandin (2012)* [30], modeled a GaN HEMT with $L_G = 3.5$ μm with a heat source length of 4 μm (longer than the gate). In an experimental study involving micro-Raman thermometry, *Rajasingam and Kuball (2004)* [31] proposed that the length of the heat source should be 0.4 ± 0.1 μm for the HEMT they studied with $L_G = 1$ μm (shorter than the gate). Their measurements were also supported by Monte Carlo simulations for electron transport in the

device which showed that the peak electric field extended over a region of approximately the same length. Furthermore, *Hosch et al. (2009)* [70] demonstrated that the temperature rise depends on the gate and drain bias condition with experimental micro-Raman temperature measurements and electro-thermal device simulations.

In general, one should expect the heat source distribution to be non-uniform and dependent on both the geometry of the device (not limited to the gate length) and the drain and source bias conditions. As described in Section 2.2.2 on thermal transport, the local Joule heat generation rate depends on the electric field and current density distributions. The value of electro-thermal modeling over solely thermal modeling is that the electron transport equations solve for the electric field and current density. Thus, a more clear physical understanding of the heat source can be gained and its dependence on a variety of factors can be observed. The current density and electric field distributions in the x-direction are shown in Figure 19 for the 1L3B device for $V_G = 0$ V and $V_D = 10$ V ($P_{loss} = 5.63$ W/mm).
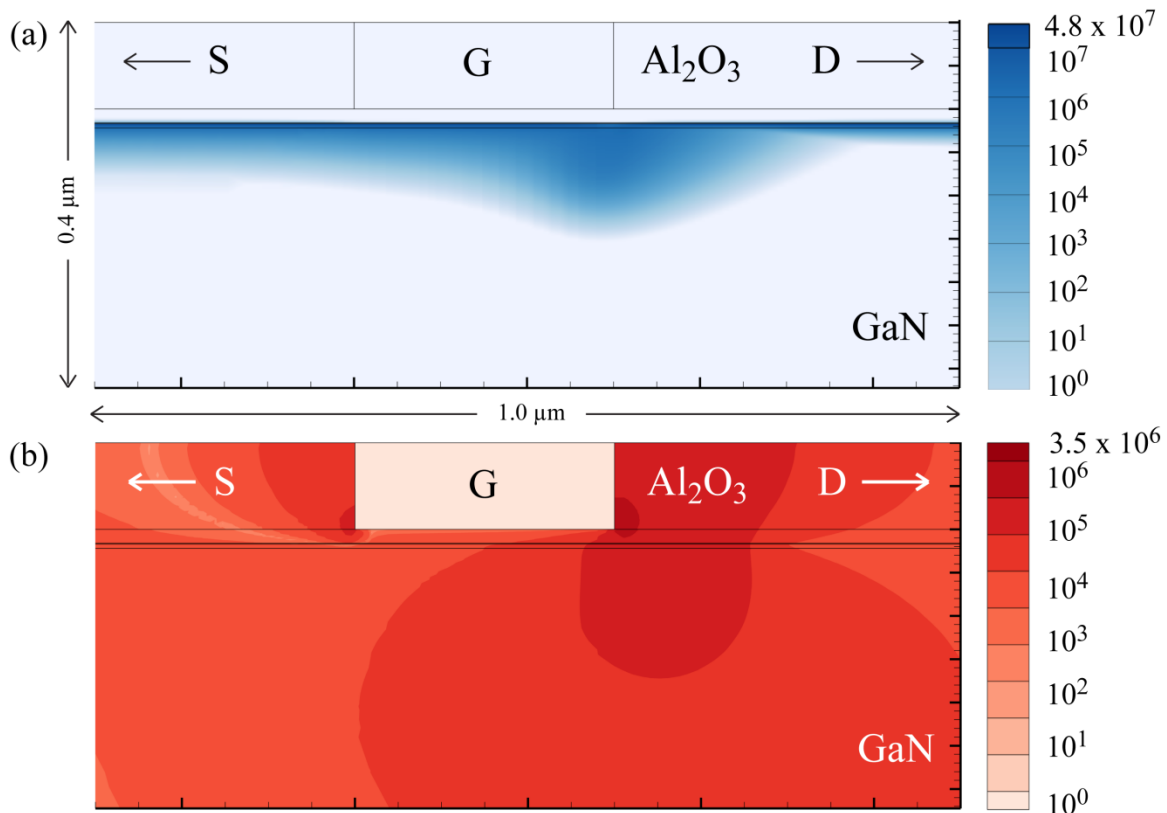
Figure 19: Electro-thermal model contour plots for 1L3B at $V_G = 0$ V and $V_D = 10$ V
(a) Current density in the x-direction (A/cm^2) and (b) Electric field in the x-direction (V/cm)

The current density distribution in the figure above shows that the current density is highest in the 2DEG near the AlGaN/GaN junction as expected and that the GaN buffer carries some of the current at the edge of the depletion region. The lateral electric field peaks at the edge of the gate contact on the drain side because of the high lateral electric field between the gate and the drain and the crowding of the electric field at the edges of the gate contact. The heat generation profile

obtained from the electro-thermal model for 1L3B demonstrates the expected region of intense heat generation at the edge of the gate on the drain side as seen in Figure 20.
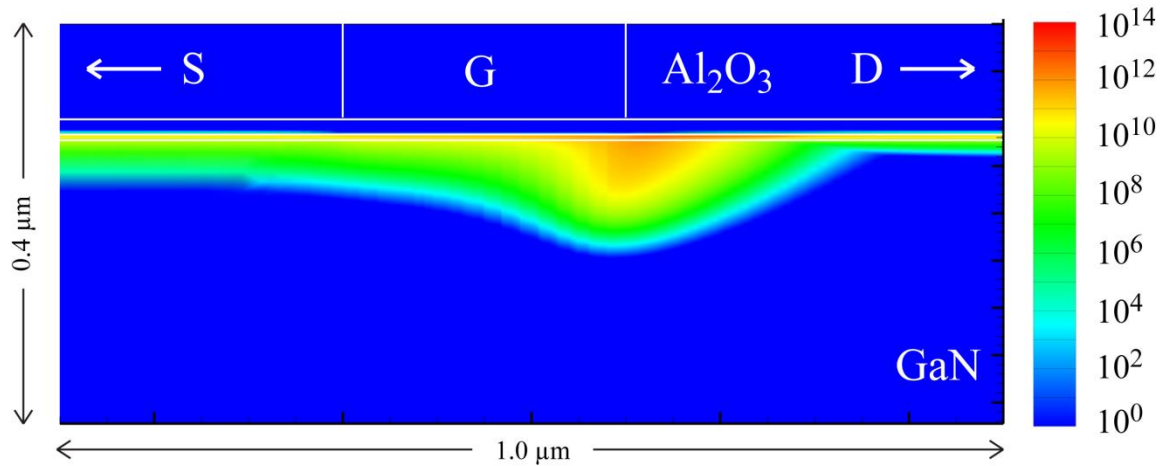
Figure 20: Heat generation (cm^3) distribution for 1L3B at $V_G = 0$ V and $V_D = 10$ V

The heat generation distribution peaks at the edge of the gate on the drain side because of the simultaneous high electric field and current density present there. However, a level of moderate heat generation ($\sim 10^8$ to 10^{12} W/cm^3) also extends along the channel and into the GaN buffer because of moderate electric field and current density magnitudes. This indicates that the heat source distribution extends across a region much longer than the intense region at the edge of the gate. In order to quantitatively capture the heat source distribution with its non-negligible contribution in the rest of the channel and extension into the GaN buffer, the 2D heat generation profile was integrated in the vertical direction to obtain the 1D heat generation profile along the channel. The 1D integrated heat generation distribution along the channel is shown in Figure 21 for 1L3B at the bias point $V_G = 0$ V and $V_D = 10$ V.
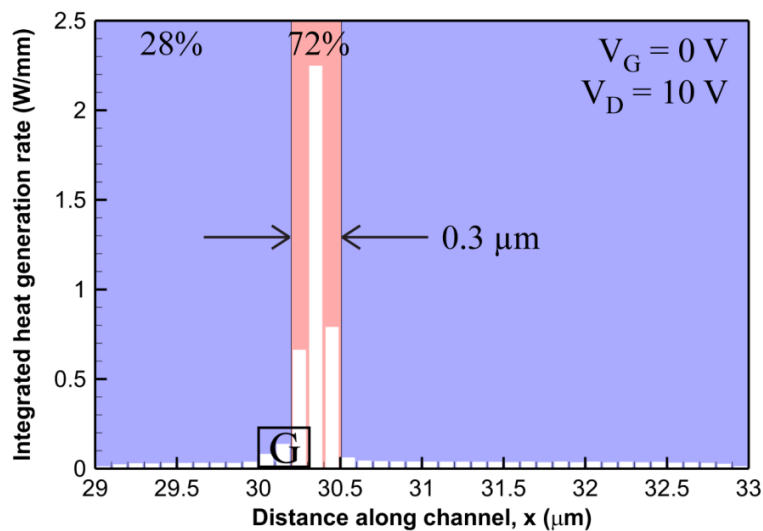
Figure 21: 1D integrated heat generation along the channel at $V_G = 0$ V and $V_D = 10$ V

50

The integrated heat generation along the channel shown in Figure 21 is indeed highly non-uniform with a clearly defined intense region approximately 0.3 μm length and a dilute region that extends throughout the rest of the channel. The intense and dilute regions contribute 72% and 28%, respectively, to the total heat generation of 5.15 W/mm in the channel. This a surprising result because it has been assumed by many researchers in the field that the heat source distribution should be limited to a much shorter region equal to the gate length (0.3 μm for this device). The much wider spread in the heat generation distribution with a significant fraction of the total heat generation in the dilute region indicates that inaccurate assumptions about the heat source length may result in incorrect device temperatures predicted by thermal models.

Gaining a quantitative understanding of the dependence of the heat source distribution on the gate and drain bias is one of the primary purposes of this electro-thermal modeling study. As the drain bias increases, the total amount of heat generated in the channel, *i.e.,* $P_{loss} = I_D V_D$, increases. In addition, the heat source distribution changes because the extent of the high electric field region between the gate and the drain increases. The 1D integrated heat source distribution for $V_G = 0$ V and $V_D = 10$ V and 20 V is shown in Figure 22.
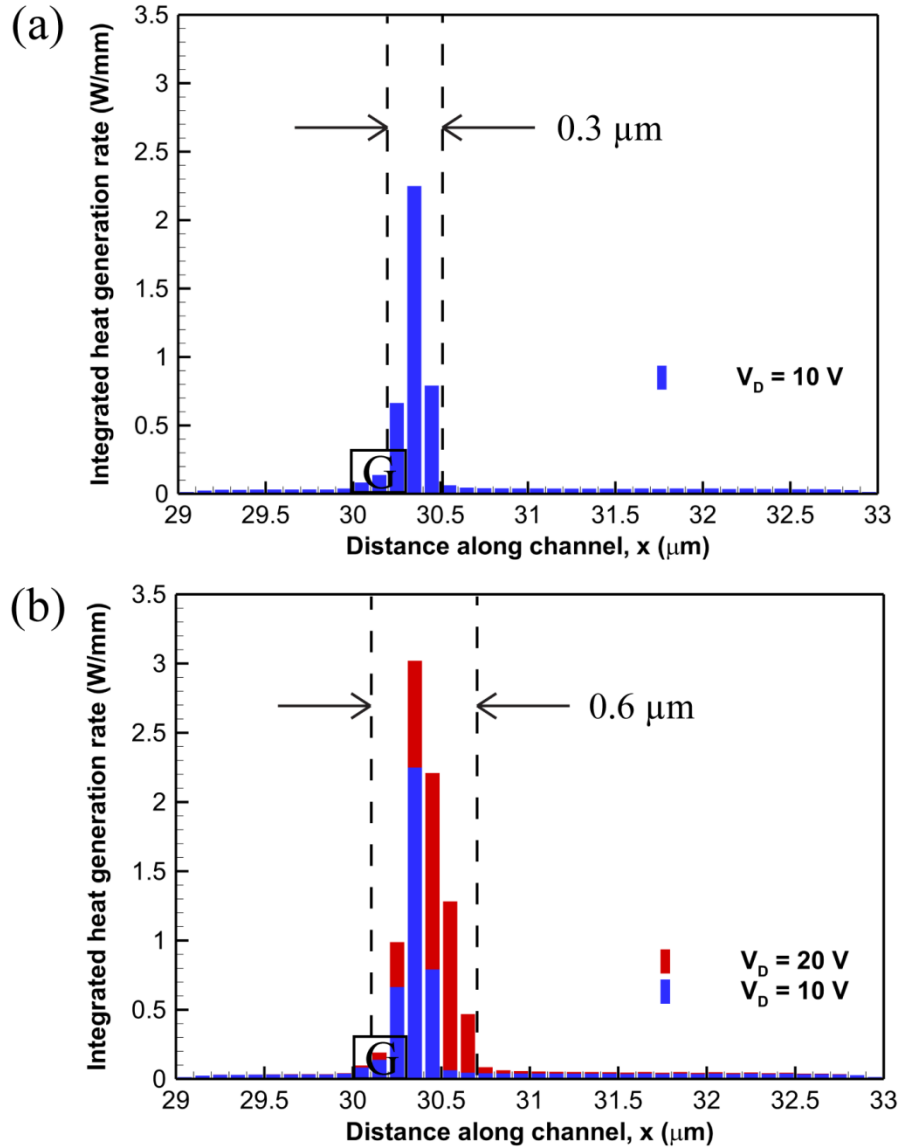
Figure 22: 1D integrated heat generation along the channel dependence on drain bias
(a) Profile for $V_G = 0$ V and $V_D = 10$ V and (b) Profile for $V_G = 0$ V and $V_D = 20$ V with the profile for $V_G = 0$ V and $V_D = 10$ V overlaid for comparison

One can verify that the total heat generated in the channel, shown as the sum of the vertical bars in Figure 22, increases as the drain voltage is increased from 10 V to 20 V while the gate voltage is kept constant. As previously noted, approximately 70% of the heat generated in the channel for the condition $V_G = 0$ V and $V_D = 10$ V occurs in an intense region of 0.3 µm length. As the drain bias is increased to 20 V, the total heat generated in the device is 9.64 W/mm with 8.16 W/mm (~85%) generated in a region of ~0.6 µm length. Thus, the heat source becomes more highly concentrated in the intense region adjacent to the gate (less fraction of the heat generated in the rest of the channel) and the width of the intense region increases. This trend plays an important role in describing the heat source at the quiescent DC bias point for an RF or high-frequency amplifier.

52

The dependence on the shape of the heat source and the temperature distribution in GaN HEMTs on the gate bias is an interesting research topic because it demonstrates the strong dependence of thermal issues on device electrical operation. In a previous experimental study by *Hosch et al. (2009)* [70], the device temperature was found to change by an appreciable amount (up to ~30 °C) when the gate and drain bias were changed so as to keep the total power dissipated constant. Due to the spatial averaging inherent in micro-Raman spectroscopy, the device needed to be biased at power dissipation levels higher than 5 W/mm to observe a noticeable temperature difference. A plot of the measured GaN temperature by micro-Raman thermometry as a function of the dissipated power is shown in Figure 23 [70].
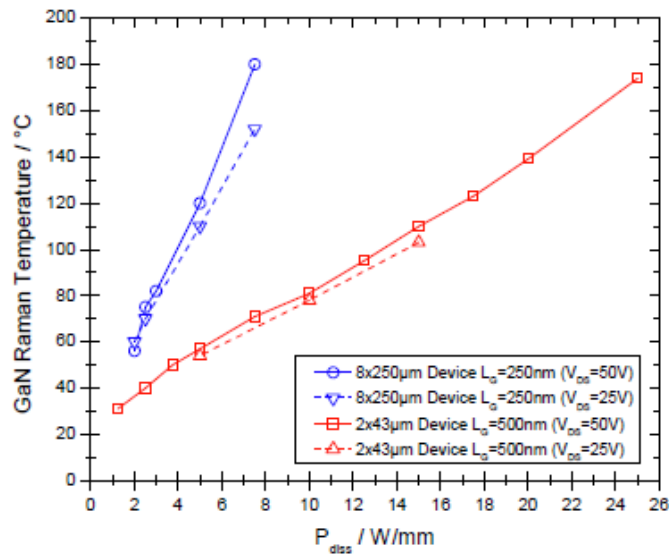
Figure 23: GaN temperature measured by micro-Raman spectroscopy for various bias points [70]

In order to capture the effect of the varying gate bias on the heat source and temperature distribution, the drain bias was changed so as to keep the total heat generated constant. Then, the change in temperature between two different bias points must be due to the change in the heat source shape, related to thermal spreading resistance, rather than the total amount of power generated. This can be visualized by moving along a line of constant power or a hyperbola in the drain current-drain voltage plane or electrical output characteristics as shown in Figure 24.
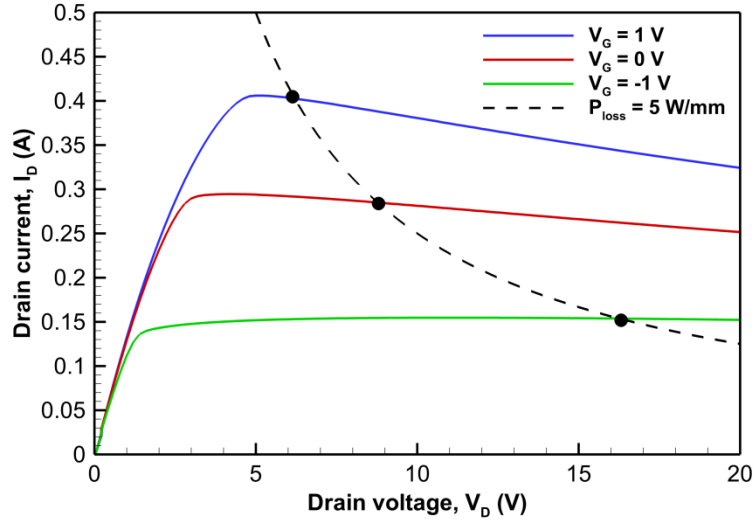
Figure 24: 1L3B modeled output characteristics with 5 W/mm power dissipation line

A decrease in gate voltage (more negative) requires that the drain voltage be increased to keep the total dissipated power constant. Therefore, the trends previously described involving the impact of the drain bias on the heat source will also be present when changing the gate bias. As the gate bias becomes more negative, the heat source is expected to become more concentrated at the edge of the gate in the intense region because of the decreasing current in the rest of the channel. There should be a very high power dissipation at the edge of the gate because of the high electric field associated with the reverse biased Schottky contact and the increased lateral electric field due to the high drain bias. The 1D integrated heat source distribution for the bias conditions $V_G = 1$ V and $V_D = 6.21$ V and $V_G = 0$ V and $V_D = 8.77$ V corresponding to a power dissipation of 5 W/mm are shown in Figure 25.
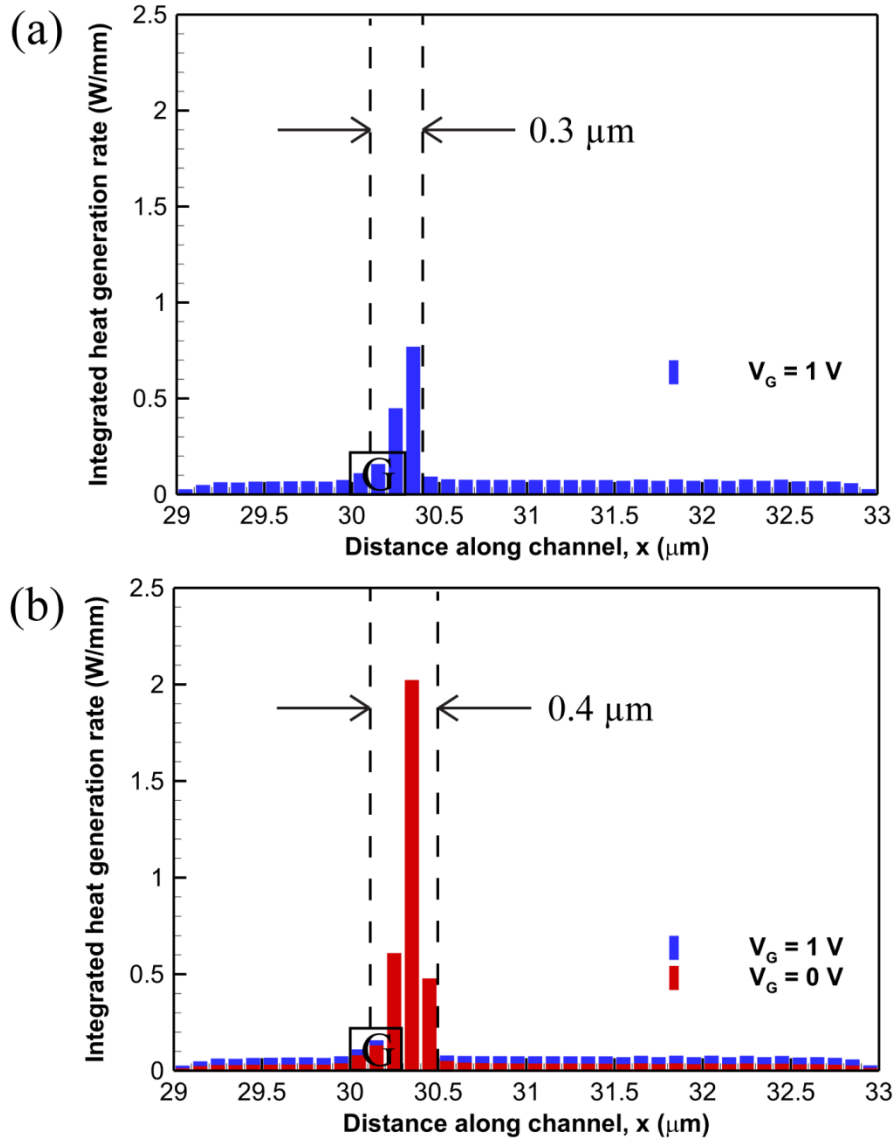
Figure 25: 1D integrated heat generation along the channel for 5 W/mm power dissipation (a) Profile for $V_G = 1$ V and $V_D = 6.21$ V and (b) Profile for $V_G = 0$ V and $V_D = 8.77$ V with the profile for $V_G = 1$ V and $V_D = 6.21$ V overlaid for comparison

The electro-thermal modeling results confirm that as the gate bias becomes more negative, less power is dissipated over the dilute region in the channel and the heat generation become increasingly concentrated at the edge of the gate on the drain side. From $V_G = 1$ V to $V_G = 0$ V, there is a slight increase in the length of the intense region from 0.3 μm to 0.4 μm. However, this effect is most likely due to the increasing lateral electric field because the drain bias correspondingly increased from $V_D = 6.21$ V to $V_D = 8.77$ V. The change in the heat source has a noticeable impact on the device temperature in the near-junction region as shown in Figure 26, increasing the peak temperature from 78.4 °C to 92.0 °C as V_G changes from 1 V to 0 V.
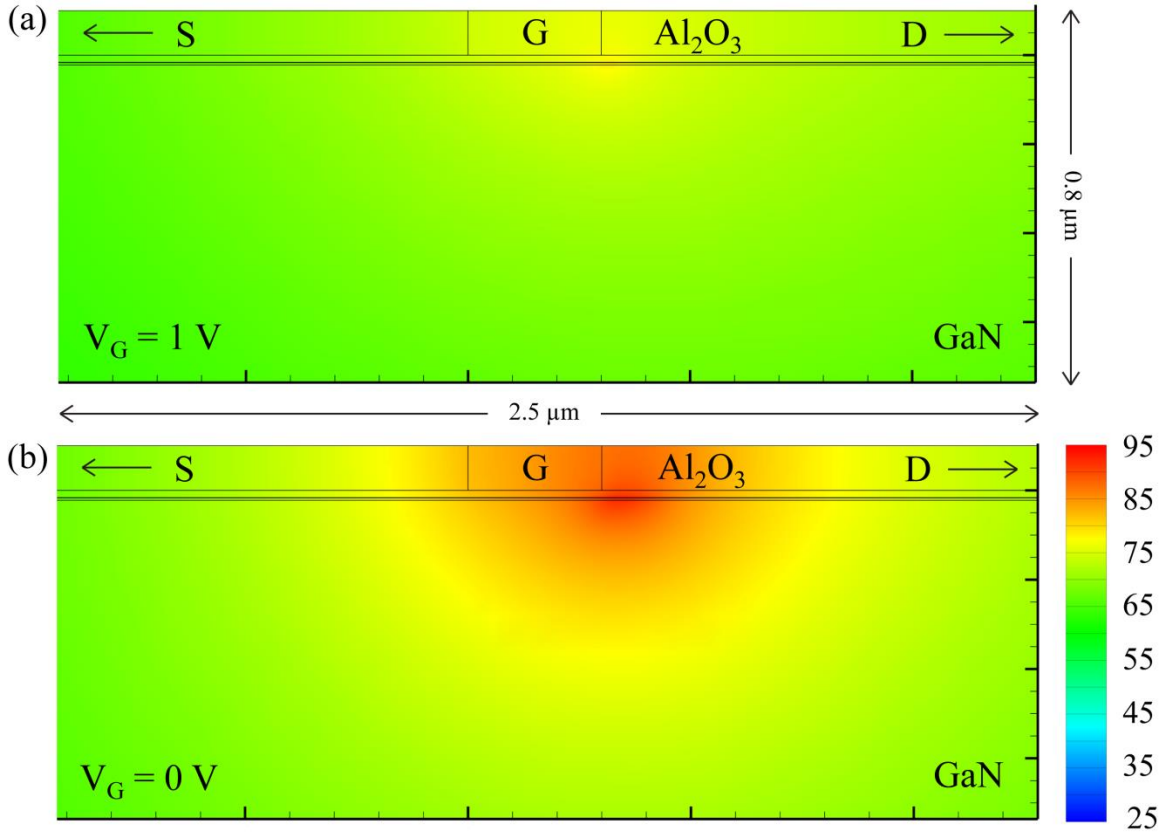
Figure 26: Electro-thermal device temperature profile (°C) at 5 W/mm power dissipation
(a) $V_G = 1$ V and $V_D = 6.21$ V and (b) $V_G = 0$ V and $V_D = 8.77$ V

This finding highlights the importance of accurately capturing the heat source distribution appropriately and considering the dependence on gate and drain bias. As the bias condition changes, the heat source may changes shape, leading to varying values of the thermal spreading resistance and temperature rise. Although the difference in temperature may be difficult to detect with even the highest spatial resolution temperature measurement techniques, such as micro-Raman spectroscopy, it is important to be aware of when comparing peak temperature values. In the case of RF and high frequency GaN amplifiers, it is common to bias the amplifier at a gate voltage near the threshold voltage (V_T) and a large drain voltage (20 V to 30 V or higher). Thus, with respect to thermal issues in GaN HEMTs and MMICs for defense and communications applications, it is important to gain a clearer description of the heat source for a typical quiescent DC bias condition. A schematic of HEMT electrical output characteristics with the typical quiescent bias point is shown in Figure 27.
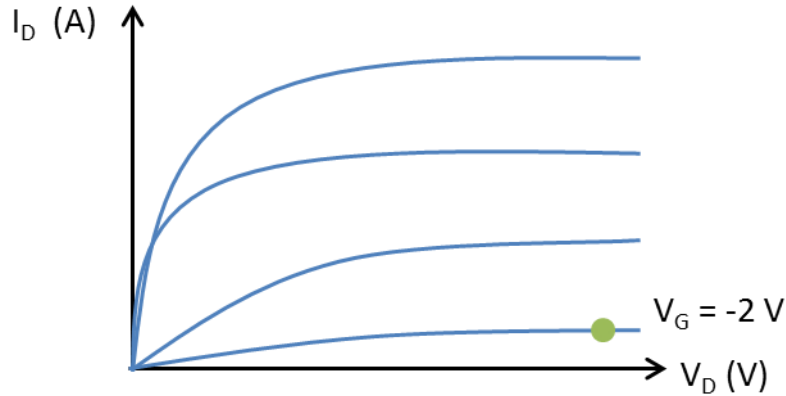
Figure 27: Schematic of HEMT output characteristics with quiescent bias point

Although measurement data was not collected for the test devices at such high drain voltages, the bias point V_G = -2 V and V_D = 50 V was simulated for the purposes of studying the heat source distribution in a typical quiescent DC bias condition. The 1D integrated heat source distribution for the simulated bias point V_G = -2 V and V_D = 50 V is shown in Figure 28.
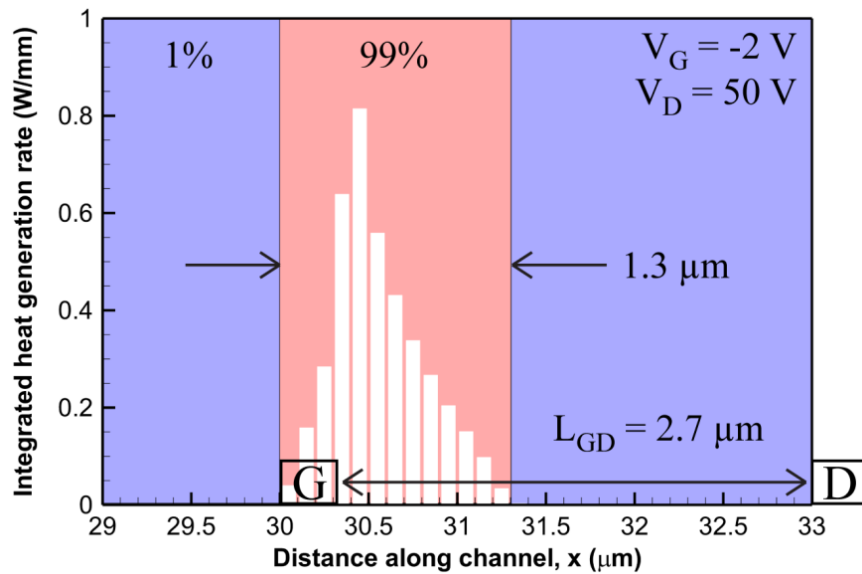
Figure 28: 1D integrated heat source distribution for typical RF quiescent bias point simulated at V_G = -2 V and V_D = 50 V (P_{loss} ~4 W/mm)

As can be seen in Figure 28, the heat source distribution along the channel is highly non-uniform and extends throughout a region approximately 1.3 μm in length, much larger than the gate length of only 0.3 μm. About 99% of the heat generated in the channel is generated in the intense region, indicating that the heat source can be modeled as a single source with a certain distribution. The heat source distribution is asymmetric and peaks near the edge of the gate on the drain side as previously noted. The length over which the heat source extends, however, may be more dependent on the gate to drain spacing (L_{GD}) rather than the gate length itself. This observation was previously noted by *Anholdt (1995)* [69] but not described in detail. The use of a

single heat source equal to gate length may significantly overestimate the peak temperature in the device, as shown by comparison of the channel temperature profiles in Figure 29 for the electro-thermal model and a single heat source of length 0.3 μm with the same total heat generation.
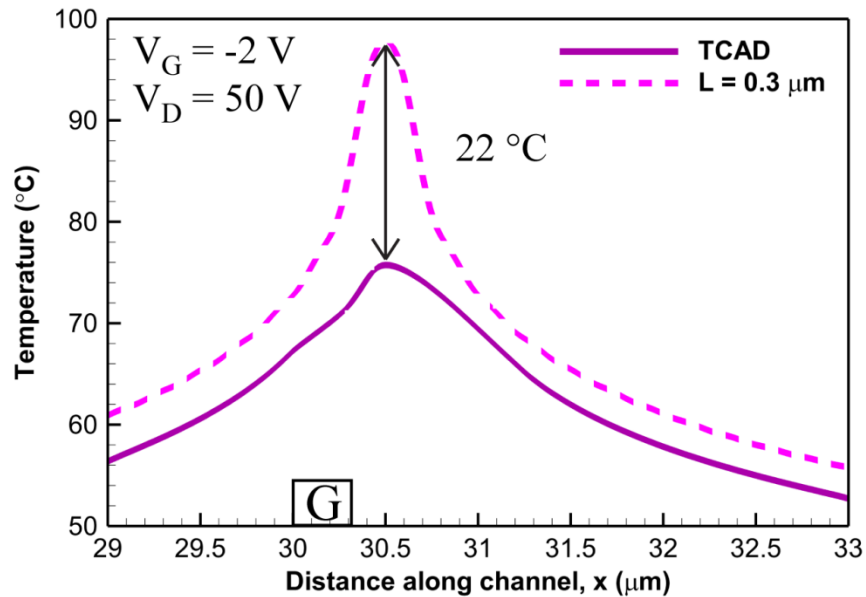
Figure 29: Channel temperature profiles from electro-thermal (TCAD) simulation and a 0.3 μm heat source with the same power dissipation for RF quiescent point bias (V_G = -2 V and V_D = 50 V)

In the case of the device tested, the use of a single heat source equal to the gate length overestimated the peak device temperature by 22 °C (~29% of the peak electro-thermal modeled temperature). To the best knowledge of the author, this effect has not been adequately appreciated nor described in detail in the past. The observations of *Hosch et al. (2009)* [70] could only provide limited experimental evidence of the effect of bias condition and heat source distribution on the device temperature because of the spatial resolution of micro-Raman spectroscopy.

2.7 Conclusions

A significant portion of this thesis has been devoted to efforts in electro-thermal modeling of AlGaN/GaN HEMTs because of the critical importance of understanding the heat source in these devices. Many researchers have not thoroughly investigated this topic because it is labor-intensive and requires knowledge in the fields of semiconductor device physics and heat transfer. While this study provided new quantitative insight into this problem, there are a number of remaining questions to be answered. In particular, it would be interesting to investigate the dependence of the heat source on the gate length and gate-to-drain spacing for several different HEMTs though modeling and experimental temperature measurements. It would also be worthwhile to investigate how the heat source distribution changes for new device structures, such as InGaN/GaN HEMTs, field plates, or recessed etched gate structures. In the future, more modeling studies and experimental temperature measurements are needed.

3. Thermal Modeling

As demonstrated in the previous chapter of this thesis, fully coupled electro-thermal modeling of high power semiconductor devices can be both challenging to implement properly and computationally expensive. Therefore, it is most common for thermal management researchers and thermal design engineers in industry to only model the thermal aspects of semiconductor devices. Fully coupled electro-thermal modeling is very helpful for gaining physical insight into the mechanism of heat generation and the effect of self-heating on device performance but is often restricted to 2D models. Three-dimensional thermal analysis of a device is more realistic because it accounts for thermal spreading in all three directions, which is an important factor in thermal issues in GaN-based electronics. The disadvantage, however, of neglecting the electron transport equations is that the heat source due to electrical power dissipation must be prescribed as boundary condition or volumetric heat source in the thermal model. The ability to provide more quantitative and accurate insight about the size and shape of the heat source for thermal modeling is one of the central aims of this thesis in the previous chapter.

Thermal modeling by FEA techniques is a standard approach in both research fields and commercial design of GaN-based electronics. The numerical thermal modeling discussed in the first section of this chapter is not new in itself; the initial description is intended to provide a context for the parametric studies and comparison to analytical heat transfer models in the following sections. The section on analytical modeling, however, presents a new result obtained in collaboration with Prof. Yuri Muzychka of Memorial University of Newfoundland, St. Johns, Canada for temperature rise in compound, orthotropic systems with interfacial resistance [71]. This model for steady-state temperature rise provides an accurate method for computing the temperature distribution in GaN HEMTs without some of the key simplifications made in previous analytical models. A number of additional important considerations in the field of device-level thermal analysis of GaN HEMTs are given in the following section with an emphasis on physical insight and best modeling practices. Finally, temperature rise in GaN HEMTs in power switching applications is discussed, in which the transient temperature rise during switching transitions plays a major role.

3.1 Finite Element Analysis (FEA) Modeling

Finite element analysis (FEA) modeling in COMSOL Multiphysics 4.2a/4.3a was implemented to solve for the temperature distribution in GaN HEMTs subjected to the relevant boundary conditions on the temperature and heat flux associated with self-heating and thermal management. This section is intended as introductory description of the FEA models used in this chapter while the primary results from the models are presented in the following sections. As an example, a standard device layout for a multi-finger GaN HEMT for RF applications is shown in Figure 30 based upon the device described by *Garven and Calame (2009)* [15].
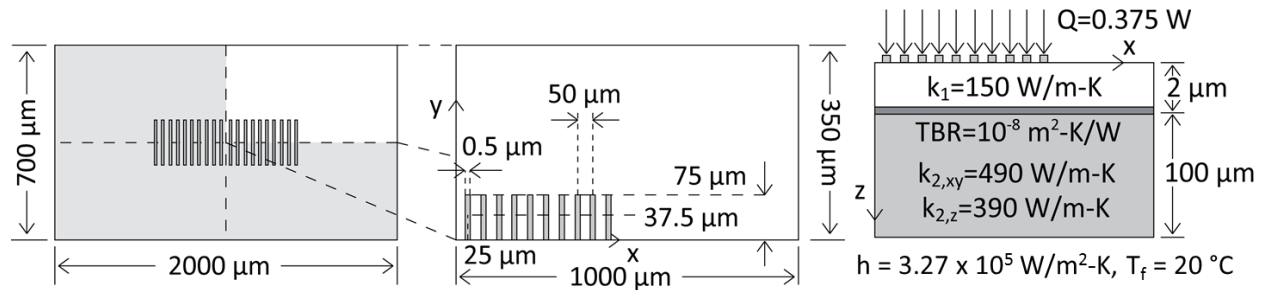
Figure 30: Example device layout for multi-finger GaN HEMT model

For the device shown in Figure 30, symmetry in the expected temperature distribution about the x- and y-axes are used to reduce the full model to a quarter model. Insulated (adiabatic) boundary conditions on the heat flux are implemented to capture the symmetry in the temperature distribution. The heat source is implemented as a heat flux boundary condition of 10^{10} W/m^2 distributed uniformly over the typical area of a gate (0.5 µm x 150 µm) to be consistent with 5 W/mm power dissipation. In this case, the heat transfer coefficient at the base was set equal to 3.27 x 10^5 W/m^2-K with an ambient temperature of 20 °C to model spreading through a SnAg die attach layer to a high performance micro-channel cooler [15]. The geometry of the quarter model in COMSOL Multiphysics 4.3a is shown in Figure 31 along with a typical tetrahedral mesh used to discretize the domain.
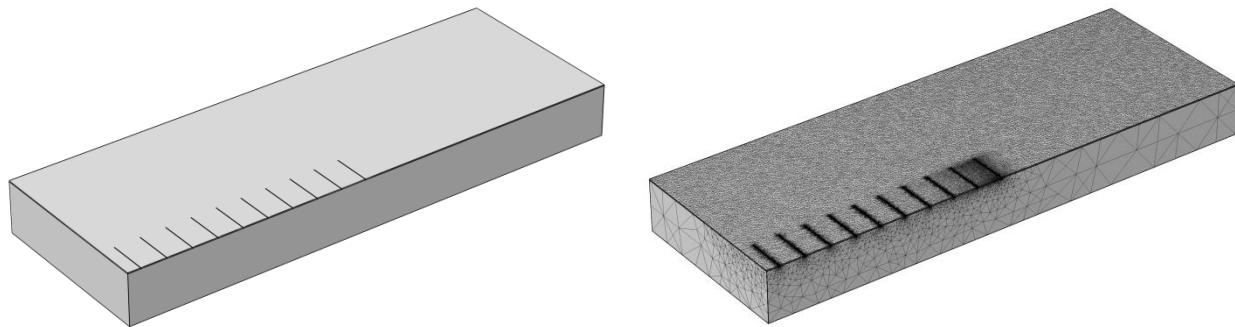
Figure 31: Quarter-model of multi-finger GaN HEMT in COMOSOL Multiphysics 4.3a
(a) Device geometry and (b) device geometry overlaid with tetrahedral mesh

As can be seen from the figure above, a high density of elements are required in the regions around the heat source where the local heat flux is very high (~10^9 to 10^{10} W/m^2) and the temperature changes rapidly over short distances (~1 to 2 µm). As a note, the areas over which the prescribed heat flux was applied (0.5 µm) is three orders of magnitude smaller than the other major length scales in the device geometry (100 to 1000 µm). In most of the thermal models for multi-finger GaN HEMTs, 3D tetrahedral elements with quadratic or cubic basis functions were used as the meshing scheme. Hexahedral (triangular prism) elements were also tested for computational efficient and accuracy; tetrahedral elements were found to be more robust and stable in COMSOL. The model with the mesh pictured in the figure above had 1.26 million tetrahedral elements with an average element quality of 0.426. Average computational times for

a linear solution (constant thermal conductivity) for the steady-state temperature rise require 5 to 10 minutes on a RedHat Linux workstation with one CPU (quad core) and 12 GB RAM.

The temperature distribution shown for the example case in Figure 32 provides some initial insight into the key factors in temperature-rise in GaN HEMTs. The first observation is that the temperature rise is non-uniform with sharp peaks in the region around the heat sources. This kind of temperature distribution is evident of large thermal spreading resistance due to the spreading out of a small heat source to a much larger area. Therefore, the temperature drop across the 2 µm thick layer can be comparable to the temperature drop in the substrate.
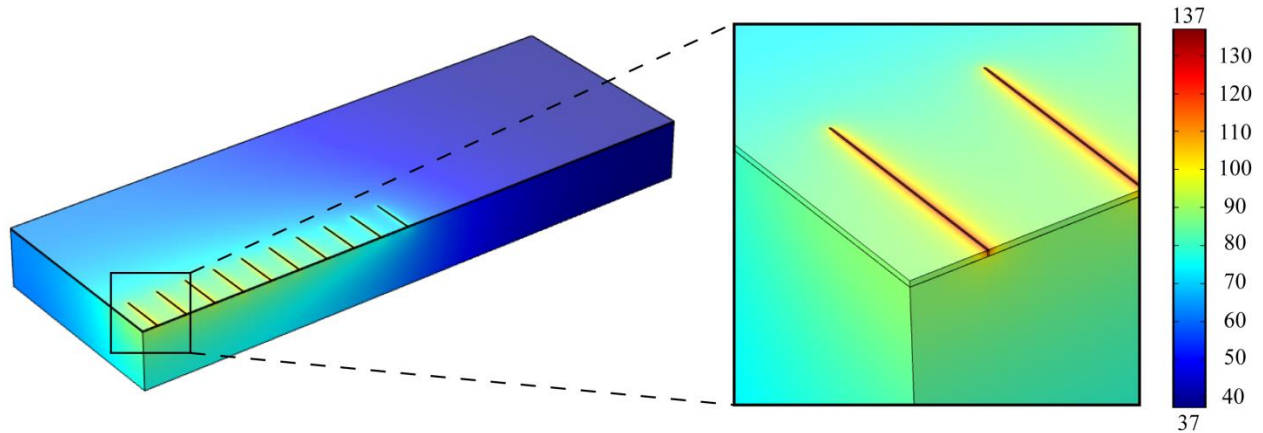
Figure 32: Temperature distribution (°C) in example multi-finger GaN HEMT

3.2 Analytical Thermal Modeling

Over the past few decades, researchers have used analytical techniques based on infinite series solutions [72], spreading angle models [73], and finite element or finite difference models [15] to predict device temperatures. While the use of infinite series solutions to the heat equation is common in package-level analysis, device-level simulations are more often approached with finite element analysis (FEA) techniques. However, GaN devices often have a very small heat source area ($\sim$0.5 x 150 µm^2) corresponding to the length and width of each gate finger [15]. In high-frequency power amplifier applications, there is a significant near-junction thermal resistance associated with the spreading of heat from the small heat sources to a much larger substrate base ($\sim$2 x 2 mm^2). The wide difference in thicknesses of the epitaxial layers of GaN (1-2 µm), aluminum nitride (AlN) nucleation layer (20-100 nm), and silicon carbide (SiC) substrate (100 µm), requires a very fine mesh near the heat sources to achieve accurate results from FEA models. Therefore, infinite series solutions offer an attractive alternative to FEA for accurate junction temperature analysis, provided they can represent the device structure and thermal properties.

A number of effects, including GaN-SiC thermal boundary resistance (TBR) and orthotropic and temperature-dependent thermal conductivity, have been identified as important in thermal analysis of GaN-based devices [32], [23]. GaN-SiC TBR originates from the finite conductance at the GaN-AlN and AlN-SiC interfaces of the nucleation layer, the conduction resistance of the AlN nucleation layer with poor crystal quality, and the conduction resistance of the poor quality GaN region adjacent the GaN-AlN interface. The effective GaN-SiC TBR (or inversely, the

61

effective interfacial conductance) has been measured experimentally and benchmarked for a variety of academic and industrial GaN-on-SiC epitaxy producers [33]. This finite interfacial conductance has been shown to result in a significant additional temperature rise, which should be accounted for by infinite series solutions. Furthermore, semi-insulating SiC substrates have been shown to have significant anisotropy in thermal conductivity [65], [75]. Previous closed-form analytical models based on prolate spheroidal and ellipsoid coordinate systems [23] and multi-finger thermal resistance expressions [51] cannot adequately account for all of these effects. In the case of the expression developed by *Darwish et al. (2004)* [23], the GaN-SiC TBR is not accounted for and the base of the SiC substrate is held at a fixed temperature. *Freeman and Mueller (2008)* [51] considered the effect of TBR and temperature-dependent thermal conductivity but could only solve for the average channel temperature, which may be much lower than the peak device temperature in a GaN HEMT. The analytical model presented in this paper accounts for GaN-substrate TBR or finite interfacial conductance, orthotropic and temperature-dependent thermal conductivity, and convection at the substrate base. The results of the analytical model are compared with results obtained using the commercial COMSOL Multiphysics™ 4.2a/4.3a FEA software package. This flexible analytical model based on infinite series techniques can capture these important effects and accurately predict the average and peak temperatures in GaN devices and is shown to be much more computationally efficient than FEA models.

3.2.1 Problem Formulation

The analytical model developed by *Muzychka et al. (2013)* [71] solves for the steady-state temperature at the surface or source plane of an eccentric heat source located on a rectangular, compound domain with interfacial resistance between the two layers. Although this layout may be applied to a variety of thermal spreading problems in electronics packaging, it is particularly suitable for solving for the device-level temperature distribution in GaN HEMTs. Through a convenient transformation, orthotropic thermal conductivity of both layers can be accounted for by the model. The layout of the model for a single heat source is shown in Figure 33 from the top (xy-plane) and side (xz-plane) views.
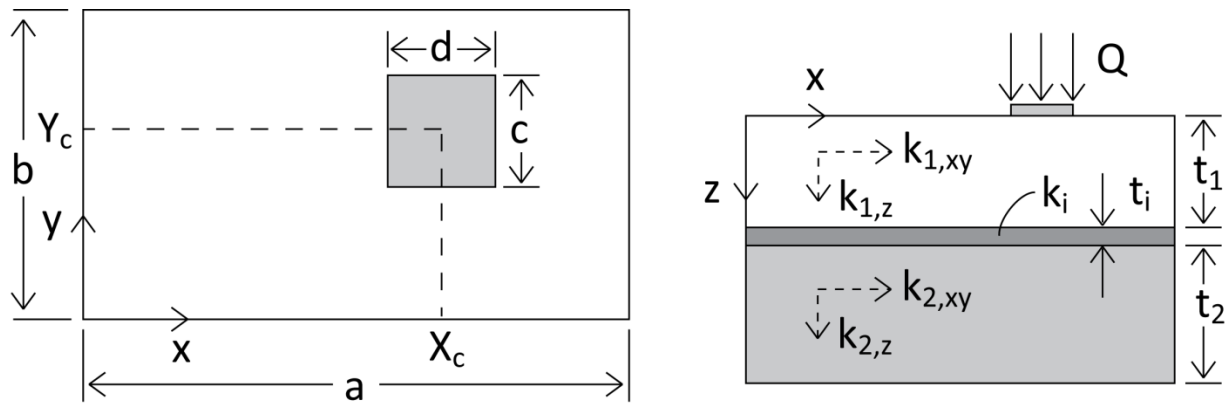
Figure 33: Model layout for temperature rise in rectangular, compound domain

The equations and derivation presented in this section is from the manuscript *Muzychka et al. (2013)* [71] with some abbreviation and additional comments provided by the author for completeness. Mathematic modeling of the problem and solution for the Fourier coefficients was performed by Prof. Yuri S. Muzychka as part of an on-going collaboration for analytical modeling of the problem. The aim is to solve for the temperature distribution in the domain by solving the linear heat conduction equation with orthotropic thermal conductivity in the xy-plane and z-axis

$$k_{xy}\left(\frac{\partial^2\theta}{\partial x^2} + \frac{\partial^2\theta}{\partial y^2}\right) + k_{zz}\frac{\partial^2\theta}{\partial z^2} = 0 \tag{42}$$

where $\theta(x,y,z) = T(x,y,z) - T_\infty$ is the temperature excess above the ambient, $k_{xy} = k_{xx} = k_{yy}$ is the thermal conductivity along the x- and y-axes and k_z is the thermal conductivity along the z-axis for each domain. For mathematical convenience, separate z-coordinates are defined for each layer. The heat source is modeled as a prescribed flux over a rectangle of dimensions c by d centered about the point (X_c, Y_c). Then, the boundary conditions in the source plane are

$$-k_{1,z}\frac{\partial\theta_1}{\partial z} = q \tag{43}$$

over the regions covered by the heat sources and

$$\frac{\partial\theta_1}{\partial z} = 0 \tag{44}$$

elsewhere to represent an adiabatic surface in the plane $z_1 = 0$. At the interface between the first and second layer, the continuity condition requires

$$k_{1,z}\frac{\partial\theta_1}{\partial z}\bigg|_{z_1=t_1} = k_{2,z}\frac{\partial\theta_2}{\partial z}\bigg|_{z_2=0} \tag{45}$$

for perfect interface contact and

$$-k_{1,z}\frac{\partial\theta_1}{\partial z}\bigg|_{z_1=t_1} = h_c[\theta_1(x,y,t_1) - \theta_2(x,y,t_1)] \tag{46}$$

for a finite interfacial conductance h_c between the two layers. At the sink plane, the boundary condition may be a fixed temperature excess

$$\theta_2 = \theta_{base} \tag{47}$$

where θ_{base} is the prescribed base temperature or

$$-k_{2,z}\frac{\partial\theta_2}{\partial z}\bigg|_{z_2=t_2} = h_s[\theta_2(x,y,t_2)] \tag{48}$$

for convection at the sink with a heat transfer coefficient h_s. All other surfaces are assumed to be adiabatic. This is a good assumption in general for GaN HEMTs in which the side and top surfaces are cooled by natural convection with a heat transfer coefficient on the order of 10 to 100 W/m^2-K. Thus, the dominant heat sinking path is through the substrate, which is attached to a temperature controlled stage in the laboratory or a package with heat sink in deployed devices. Orthotropic thermal conductivity of the layers can be accounted for through the transformation

$$\xi = \frac{z}{\sqrt{k_{zz}/k_{xy}}} \tag{49}$$

such that Laplace's equation simplifies to

$$\frac{\partial^2 \theta}{\partial x^2} + \frac{\partial^2 \theta}{\partial y^2} + \frac{\partial^2 \theta}{\partial \xi^2} = 0 \tag{50}$$

which can be solved in closed-form with Fourier series techniques. In applying the boundary and continuity conditions, the effective isotropic layer properties are defined as

$$\bar{k} = \sqrt{k_{xy}k_{zz}} \tag{51}$$

and

$$\bar{t} = \frac{t}{\sqrt{k_{zz}/k_{xy}}} \tag{52}$$

where $\bar{k}$ and $\bar{t}$ are the effective isotropic thermal conductivity and thickness of each of the layers. The details of the transformation of the boundary conditions can be found in *Muzychka et al. (2013)* [71].

3.2.2 General Solution, Fourier Coefficients, and Spreading Functions

The general solution for the excess temperature rise is then

$$
\begin{aligned}
\theta(x,y,z) = A_0 + B_0 z &+ \sum_{m=1}^{\infty} \cos(\lambda_m x) \left[A_1 \cosh(\lambda_m z) + B_1 \sinh(\lambda_m z) \right] \\
&+ \sum_{n=1}^{\infty} \cos(\delta_n y) \left[A_2 \cosh(\delta_n z) + B_2 \sinh(\delta_n z) \right] \\
&+ \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \cos(\lambda_m x) \cos(\delta_n y) \left[A_3 \cosh(\beta_{mn} z) + B_3 \sinh(\beta_{mn} z) \right]
\end{aligned} \tag{53}
$$

where $\lambda_m = m\pi/a$, $\delta_n = n\pi/b$, and $\beta_{mn} = \sqrt{\lambda_m^2 + \delta_n^2}$ are the eigenvalues in the problem. The Fourier coefficients needed to evaluate the temperature field as a function of depth z into the

model are quite complicated and are not given explicitly in *Muzychka et al. (2013)* [71]. Since the temperature rise in the source plane is the most important, the temperature excess formula is much simpler when $z = 0$. This is because $\cosh x \to 1$ and $\sinh x \to 0$ when $x = 0$. The temperature rise in the source plane is then given by

$$
\begin{aligned}
\theta(x, y, 0) = A_0 + \sum_{m=1}^{\infty} A_m \cos(\lambda_m x) + \sum_{n=1}^{\infty} A_n \cos(\delta_n y) \\
+ \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} A_{mn} \cos(\lambda_m x) \cos(\delta_n y)
\end{aligned}
\tag{54}
$$

where the Fourier coefficients are found by matching the boundary condition in the source plane. The Fourier coefficients in simplified form are given by

$$
A_0 = \frac{Q}{ab}\left(\frac{t_1}{k_1} + \frac{1}{h_c} + \frac{t_2}{k_2} + \frac{1}{h_s}\right)
\tag{55}
$$

$$
A_m = \frac{4Q \cos(\lambda_m X_c) \sin\left(\frac{1}{2}\lambda_m c\right)}{abck_1\lambda_m^2 \phi(\lambda_m)}
\tag{56}
$$

$$
A_n = \frac{4Q \cos(\delta_n Y_c) \sin\left(\frac{1}{2}\delta_n d\right)}{abdk_1\delta_n^2 \phi(\delta_n)}
\tag{57}
$$

$$
A_{mn} = \frac{16Q \cos(\lambda_m X_c) \sin\left(\frac{1}{2}\lambda_m c\right) \cos(\delta_n Y_c) \sin\left(\frac{1}{2}\delta_n d\right)}{abcdk_1\lambda_m\delta_n \phi(\beta_{mn})}
\tag{58}
$$

and the spreading function ϕ depends on the features of the model, such as number of layers and interfacial and sink conductances. The mean temperature rise of a heat source with dimensions c by d centered about the point (X_c, Y_c) can be obtained by integrating the expression for the source plane temperature excess over the heat source. The mean temperature of the heat source is simply

$$
\begin{aligned}
\bar{\theta} = A_0 + 2\sum_{m=1}^{\infty} A_m \frac{\cos(\lambda_m X_c) \sin\left(\frac{1}{2}\lambda_m c\right)}{\lambda_m c} + 2\sum_{n=1}^{\infty} A_n \frac{\cos(\delta_n Y_c) \sin\left(\frac{1}{2}\delta_n d\right)}{\delta_n d} \\
+ 4\sum_{m=1}^{\infty} \sum_{n=1}^{\infty} A_{mn} \frac{\cos(\lambda_m X_c) \sin\left(\frac{1}{2}\lambda_m c\right) \cos(\delta_n Y_c) \sin\left(\frac{1}{2}\delta_n d\right)}{\lambda_m c \delta_n d}.
\end{aligned}
\tag{59}
$$

When the model consists of a compound (two-layer) substrate with interfacial resistance and convection at the sink, the spreading function is given by

$$\phi(\gamma) = \frac{C_1 + C_2 \tanh(\gamma t_1)}{C_1 \tanh(\gamma t_1) + C_2} \tag{60}$$

$$C_1 = \gamma \tanh(\gamma t_2) + h_s/k_2 \tag{60a}$$

$$C_2 = \frac{k_1}{k_2}\left[\gamma\left(1 + \frac{h_s}{h_c}\right) + \left(\frac{h_s}{k_2} + \frac{\gamma^2 k_2}{h_c}\right)\tanh(\gamma t_2)\right] \tag{60b}$$

where γ is equal to λ_m, δ_n, or β_{mn} depending upon the Fourier coefficient to be evaluated. The spreading function and zeroth Fourier coefficient change if there is no interfacial resistance and/or sink convection or if a single layer substrate is used.

For a compound (two-layer) substrate with a uniform sink temperature (in the limit $h_s \to \infty$)

$$\phi(\gamma) = \frac{D_1}{D_2} \tag{61}$$

$$D_1 = 1 + \frac{\gamma k_1}{h_c}\tanh(\gamma t_1) + \frac{k_1}{k_2}\tanh(\gamma t_1)\tanh(\gamma t_2) \tag{61a}$$

$$D_2 = \frac{\gamma k_1}{h_c} + \frac{k_1}{k_2}\tanh(\gamma t_2) + \tanh(\gamma t_1) \tag{61b}$$

and

$$A_0 = \frac{Q}{ab}\left(\frac{t_1}{k_1} + \frac{1}{h_c} + \frac{t_2}{k_2}\right). \tag{62}$$

For a compound (two-layer) substrate with perfect interfacial contact (in the limit $h_c \to \infty$) and convection at the sink

$$\phi(\gamma) = \frac{E_1}{E_2} \tag{63}$$

$$E_1 = \gamma \tanh(\gamma t_2) + \frac{h_s}{k_2} + \frac{k_1}{k_2}\tanh(\gamma t_1)\left[\gamma + \frac{h_s}{k_2}\tanh(\gamma t_2)\right] \tag{63a}$$

$$E_2 = \tanh(\gamma t_1)\left[\gamma \tanh(\gamma t_2) + \frac{h_s}{k_2}\right] + \frac{k_1}{k_2}\left[\gamma + \frac{h_s}{k_2}\tanh(\gamma t_2)\right] \tag{63b}$$

and

$$A_0 = \frac{Q}{ab}\left(\frac{t_1}{k_1} + \frac{t_2}{k_2} + \frac{1}{h_s}\right). \tag{64}$$

For a compound (two-layer) substrate with perfect interfacial contact (in the limit $h_c \to \infty$) and uniform sink temperature (in the limit $h_s \to \infty$)

$$\phi(\gamma) = \frac{1 + \frac{k_1}{k_2}\tanh(\gamma t_1)\tanh(\gamma t_2)}{\tanh(\gamma t_1) + \frac{k_1}{k_2}\tanh(\gamma t_2)} \tag{65}$$

and

$$A_0 = \frac{Q}{ab}\left(\frac{t_1}{k_1} + \frac{t_2}{k_2}\right). \tag{66}$$

For a single layer substrate with convection at the sink

$$\phi(\gamma) = \frac{\gamma\tanh(\gamma t_1) + \frac{h_s}{k_1}}{\gamma + \frac{h_s}{k_1}\tanh(\gamma t_1)} \tag{67}$$

and

$$A_0 = \frac{Q}{ab}\left(\frac{t_1}{k_1} + \frac{1}{h_s}\right). \tag{68}$$

Finally, for a single layer substrate with uniform sink temperature (in the limit $h_s \to \infty$)

$$\phi(\gamma) = \coth(\gamma t_1) \tag{69}$$

and

$$A_0 = \frac{Q}{ab}\left(\frac{t_1}{k_1}\right). \tag{70}$$

For reference, in an isotropic, three-layer substrate with uniform sink temperature, *Kokkas (1974)* [72] reported the spreading function and zeroth term Fourier coefficient to be

$$\phi(\gamma) = \left\{\tanh(\gamma t_1) + \frac{1 - [\tanh(\gamma t_1)]^2}{\tanh(\gamma t_1) + \left(\frac{k_2}{k_1}\right)\left[\frac{1 + (k_2/k_3)\tanh(\gamma t_2)\tanh(\gamma t_3)}{\tanh(\gamma t_2) + (k_2/k_3)\tanh(\gamma t_3)}\right]}\right\}^{-1} \tag{71}$$

and

67

$$A_0 = \frac{Q}{ab}\left(\frac{t_1}{k_1} + \frac{t_2}{k_2} + \frac{t_3}{k_3}\right). \tag{72}$$

The above cases represent a variety of situations that occur in thermal conduction problems semiconductor devices and electronics packages. In the case of GaN HEMTs, the most rigorous model should include orthotropic properties of the substrate (second layer), interfacial resistance between the two layers, and convection at the sink to account for the resistance between the substrate base and the ambient. The analytical model presented in *Muzychka et al. (2013)* [71] and referenced here accounts for these important factors that have been neglected in previous analytical thermal models [23]. The disadvantage of using these Fourier series based solutions is that a large number of terms may be required for the solution to converge.

The solution for the temperature distribution in the source plane for a problem with multiple heat sources can be determined using the influence coefficient method described in *Muzychka et al. (2013)* [71]. There is a computational advantage in the influence coefficient because of the symmetry of the matrix of influence coefficients; for N discrete heat sources, only $N(N+1)/2$ influence coefficients need be computed rather than N^2.

3.2.3 Analytical Model Validation

In order to demonstrate the accuracy and computational efficiency of the analytical model developed in this paper, several validation studies were conducted with comparison to 3D FEA models. Although FEA models for thermal conduction are flexible and reliable, a large number of elements may be needed to ensure convergence of the solution, particularly when the relevant length scales vary significantly in the domain of interest. Similarly, the summation of the infinite series in the analytical model may require a large amount of terms in order to converge. In the validation studies described in this section, the convergence of both the FEA and analytical models was first verified by increasing the number of elements and summation terms, respectively. Because of the growing interest in wide bandgap semiconductor devices for high power, high frequency applications, thermal analysis and management of GaN electronics has become an important research topic. GaN HEMTs were used as the test structure for the analytical model validation, and it is shown that the analytical model can calculate the device temperature much more quickly than the FEA model without reducing the accuracy of the solution.

Single Heat Source

As in many semiconductor devices, heat generation in GaN HEMTs is primarily due to Joule heating; electrons transfer their energy gained from the external electric field to the lattice via electron-phonon scattering while they traverse the channel. The most intense region of heat generation or power dissipation is adjacent to the gate toward the drain side. Although the exact length of the heat source depends on the drain and gate bias voltages, many reports in the literature, such as *Garven and Calame (2009)* [15], approximate the heat generation region in the channel as a heat flux uniformly under the gate. In all of the models compared in this section, the thicknesses of the GaN and SiC substrate layers were chosen to be 2 μm and 100 μm,

respectively, to represent standard values for high power RF monolithic microwave integrated circuits (MMICs). For both the FEA and analytical methods, a 1000 µm by 350 µm quarter model with the appropriate symmetric boundary conditions was used to represent the entire device in order to reduce the computational expense.

In the first validation study, a device with single heat source of 0.5 µm by 150 µm was modeled to represent a double finger GaN HEMT as shown in Figure 34 with the basic dimensions of the model described by Garven et al. [15]. The magnitude of the heat source was chosen to be 0.375 W per heat source corresponding to a power dissipation of 5 W/mm (normalized to the gate length) or a heat flux of 10^{10} W/m^2. The interface contact between the GaN and SiC layers was assumed to be perfect (no interfacial resistance), and the base temperature was set to a fixed value of 25 °C. These boundary conditions are highly idealized but are commonly used in device-level thermal models to determine the near-junction thermal resistance of GaN HEMTs [23].
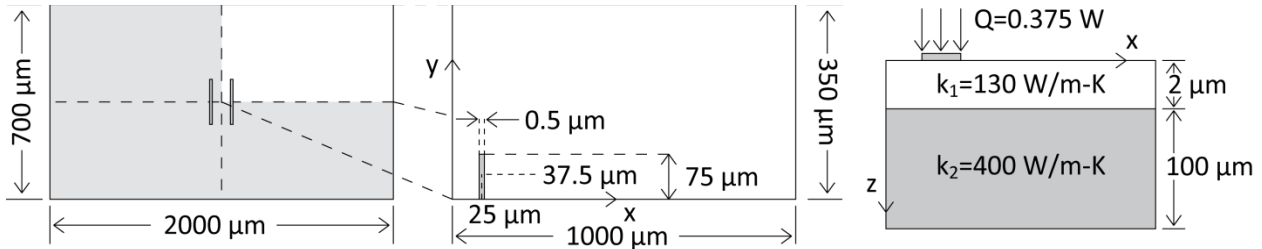
Figure 34: Device layout for single source model validation study

The temperature values at specified points from the analytical model were obtained by evaluating Eq. (54) with the appropriate form of the spreading function ϕ as given by Eq. (65) for a fixed base temperature with MATLAB R2012a with $M = N = 10^4$ terms. For summation of infinite series over a finite number of terms, the number of terms required to obtain a sufficiently accurate solution is an important concern. The centroidal and average temperature rise for the single heat source analytical model was computed as a function of the number of terms M and N (with $M = N$) in the summation. The centroidal and mean temperatures of the heat source in the single heat source validation study as well as the percent error compared to the solution with 10^5 terms are shown in Figure 35.
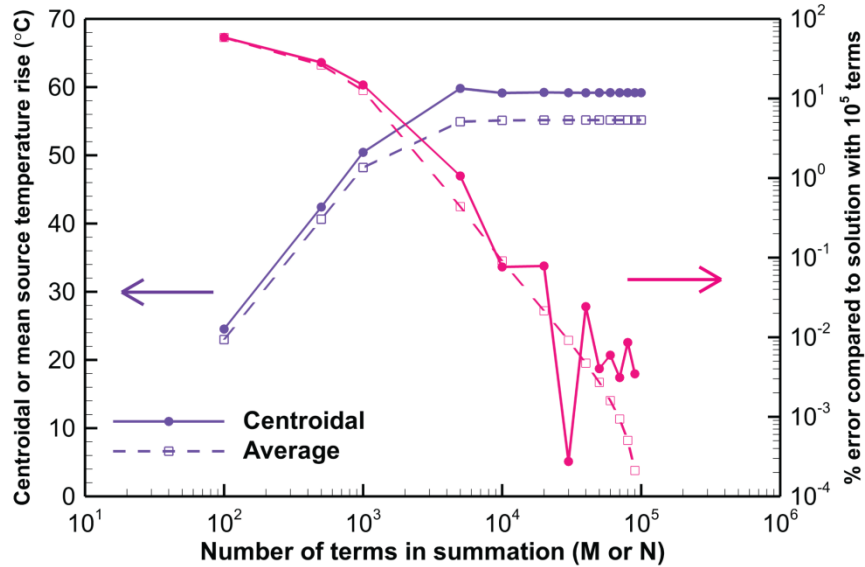
Figure 35: Analytical model Fourier series convergence study
Centroidal and mean source temperatures of the single heat source computed by the analytical model as a function of the number of terms in the summation. The left axis shows the centroidal and mean source temperatures and the right axis shows the percent error compared to the most accurate solution is 10^5 terms.

The most appropriate number of terms to use in the summations in Eqs. (54) and (65) is a trade-off between the desired accuracy of the solution and the required time for computation. As the number of terms in the series increases, the accuracy increases but the computation time is longer. The value of $M = N = 10^4$ was found to be a good compromise between accuracy and computational efficiency as it resulted in accuracy to within 0.1% of the solution with $M = N = 10^5$ terms but only required ~21.7 seconds compared to ~2160 seconds (1%). The truncation or round-off error associated with the precision of the double floating point data type in MATLAB and computation of the summation was also investigated and found to be negligible.

The data from the FEA model was calculated by solving for the 3-D temperature distribution in COMSOL Multiphysics™ 4.3a for a model with approximately ~1.7 million cubic tetrahedral elements for high accuracy and extracting a line cross-section. The convergence of the FEA solution was checked by progressively refining the mesh; the maximum error in the temperature along the cross-section between the model with ~1.7 million elements and a model with ~6.0 million elements was 0.52%. The source plane temperature profile along the line x-axis for the entire domain is shown in Figure 36 with a closer view of the region around the heat source at $x = 25$ μm shown in the inset.
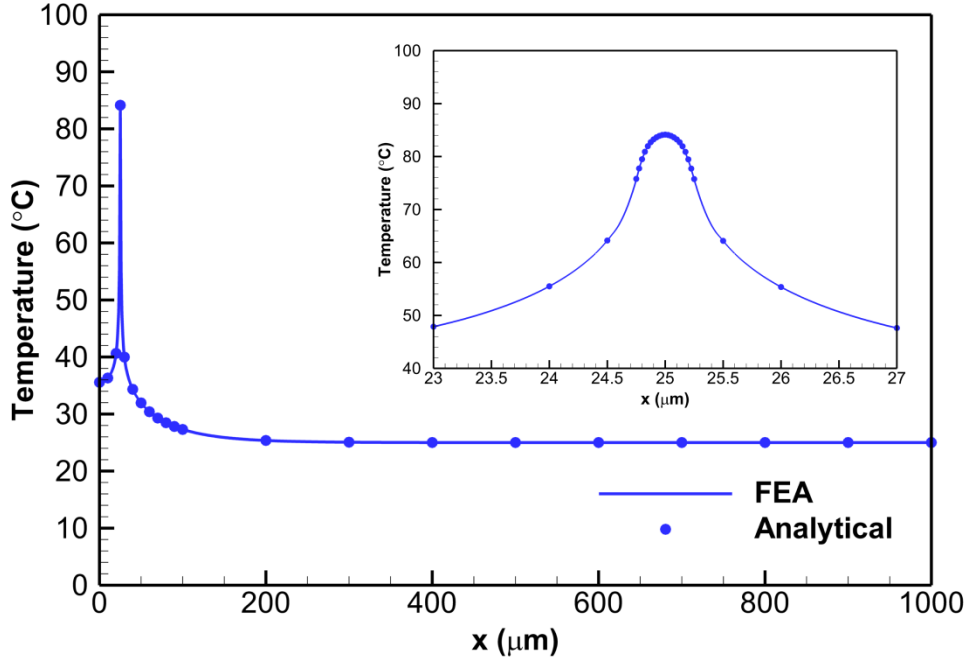
Figure 36: Single heat source temperature profile comparison for perfect interface contact
Model implemented isotropic thermal conductivity (k_1 = 130 W/m-K, k_2 = 400 W/m-K), perfect interface contact and T_{base} = 25 °C. Inset shows a closer view of the temperature profile near the heat source.

The agreement between the analytical and FEA model temperature profiles is very good, with less than 0.5% difference between the temperature rises above the sink temperature. It can be seen that the temperature profile peaks sharply within a few micrometers around the heat source, indicating that heat flow from the 0.5 μm source is greatly constricted when spreading to the fixed temperature sink at the base.

In the second model validation study, the effect of finite interfacial conductance was added to the basic model presented in the previous section. Thermal interfacial resistance between the GaN layer and SiC substrate due to lattice mismatch and crystal disorder has been shown experimentally to have a significant impact on device temperatures. The interfacial conductance between the GaN layer and SiC substrate was set to a characteristic value of 10^8 W/m^2-K while the base temperature remained fixed 25 °C. In this case for a finite interfacial conductance and a fixed base temperature, Eq. (61) was used to compute the spreading function ϕ with $M = N = 10^4$ terms in order to calculate the temperature at various points along the source plane. The temperature distribution was computed with the same 3D FEA model as in the previous study (~1.7 million elements) with the addition of a thin, thermally resistive layer between the GaN and SiC layers. Convergence of the FEA model was also verified by refining the mesh and ensuring that the maximum error in the temperature was ~0.5%. The source plane temperature profile along the x-axis is shown in Figure 37 for both the FEA and analytical models with a closer view of the temperature profile shown in the inset.
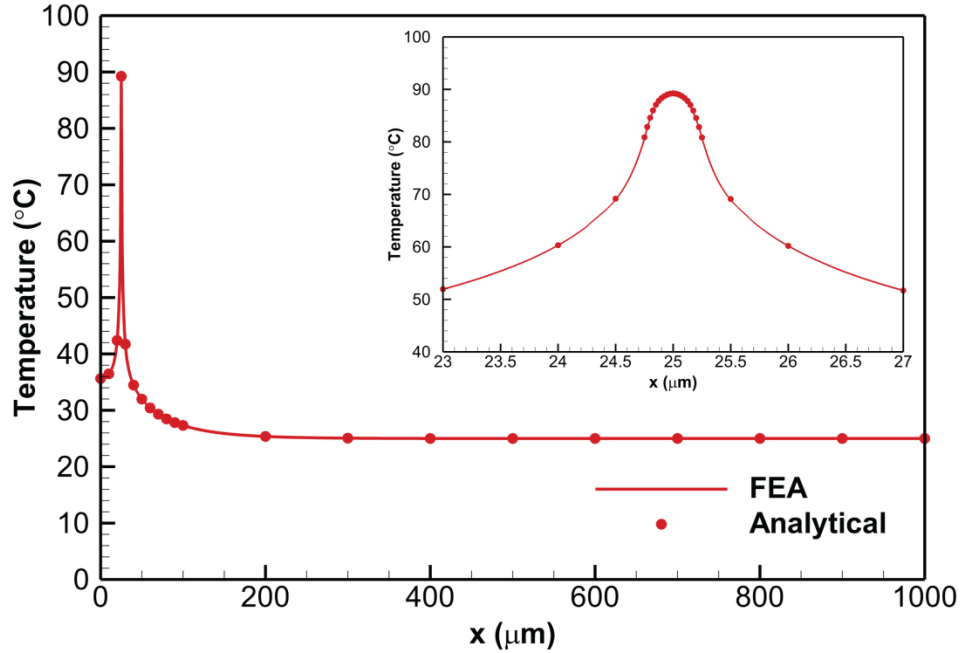
Figure 37: Single heat source temperature profile comparison for finite interfacial conductance
Model implemented isotropic thermal conductivity (k_1 = 130 W/m-K, k_2 = 400 W/m-K), h_c = 10^8 W/m²-K and T_{base} = 25 °C. Inset shows a closer view of the temperature profile near the heat source.

As in the previous case, an excellent agreement is obtained between the FEA and analytical methods, with less than 0.4% difference in the temperature rises above the base predicted by the two methods. For the case with an interface conductance of 10^8 W/m²-K, the peak device temperature increases from 84.1 °C to 89.2 °C (~8.7% increase in the temperature rise above the base) compared to the case with perfect interface contact. Although the increase in 1-D conduction resistance is quite small, this moderate increase in peak temperature occurs because of the increased spreading resistance created by the finite interfacial conductance.

The relationship between finite interfacial conductance and device temperature was further investigated through a parametric study that included the effects of orthotropic thermal conductivity of the layers. Because the peak device temperature depends strongly on the thermal conductivity of the SiC substrate, it is desirable to capture the effect of different values for thermal conductivity in the xy-plane ($k_{2,xy}$= 490 W/m-K) and z-direction ($k_{2,z}$ = 390 W/m-K) [65]. Although the thermal conductivity of GaN is isotropic, a comparison of the FEA and analytical models was also performed with $k_{1,xy}$ = 75 W/m-K and $k_{1,z}$ = 150 W/m-K as a validation of the analytical method. The average and maximum temperatures of the 0.5 µm by 150 µm single heat source representing 5 W/mm uniform power dissipation under the gate were calculated for interfacial conductance values ranging from 6.7 x 10^6 to 10^9 W/m²-K to represent the range of experimentally observed thermal interfacial resistances of 10^{-9} to 1.5 x 10^{-7} m²-K/W [33]. A comparison of the temperatures predicted by the FEA and analytical models for (i) isotropic thermal conductivity of both layers, (ii) orthotropic SiC thermal conductivity, (iii) orthotropic thermal conductivity of both layers are shown in Figure 38 with the exact values given in Table 2. The convergence of the FEA solution for the model with an interfacial

conductance of 6.7 x 10^6 W/m²-K with orthotropic thermal conductivity of both layers, which represents the highest temperature model, was also verified.
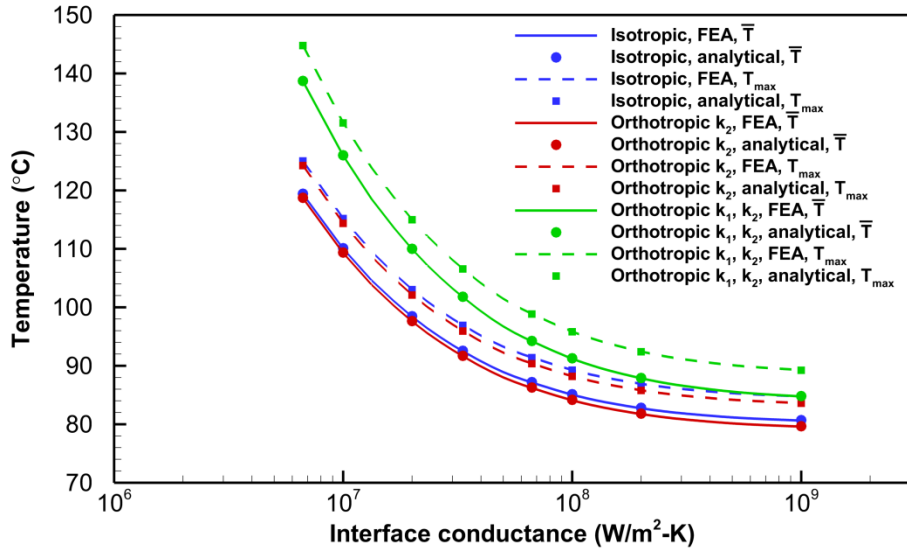
Figure 38: Average and maximum temperatures (°C) of the heat source as a function of interface conductance (W/m²-K)
Data shown for (i) isotropic k_1 and k_2, (ii) orthotropic k_2, and (iii) orthotropic k_1 and k_2.

Table 2: Average and centroidal temperature for the single heat source validation study with variable interfacial conductance and fixed base temperature of 25 °C

(i) Isotropic thermal conductivity of both layers (k_1 = 130 W/m-K, k_2 = 400 W/m-K)

h_c (W/m²-K)	Analytical		FEA	
	$\bar{T}$ (°C)	$\hat{T}$ (°C)	$\bar{T}$ (°C)	$\hat{T}$ (°C)
1.00E+09	80.6982	84.7218	80.6303	84.7667
2.00E+08	82.8053	86.8737	82.7376	86.9188
1.00E+08	85.1258	89.2491	85.0581	89.2943
6.67E+07	87.2095	91.3870	87.1419	91.4323
3.33E+07	92.5674	96.9057	92.4999	96.9511
2.00E+07	98.4659	103.0168	98.3985	103.0622
1.00E+07	110.103	115.1824	110.0357	115.2280
6.67E+06	119.4132	125.0214	119.3460	125.0671

(ii) Isotropic GaN and orthotropic SiC thermal conductivity ($k_1 = 130$ W/m-K, $k_{2,xy} = 490$ W/m-K, $k_{2,z} = 390$ W/m-K)

h_c (W/m^2-K)	Analytical		FEA	
	$\bar{T}$ (°C)	$\hat{T}$ (°C)	$\bar{T}$ (°C)	$\hat{T}$ (°C)
1.00E+09	79.6642	83.5822	79.5965	83.6274
2.00E+08	81.8103	85.7732	81.7426	85.8184
1.00E+08	84.1672	88.1852	84.0996	88.2304
6.67E+07	86.2797	90.3520	86.2122	90.3973
3.33E+07	91.6995	95.9327	91.6321	95.9782
2.00E+07	97.6528	102.0987	97.5855	102.1443
1.00E+07	109.3730	114.3478	109.3057	114.3935
6.67E+06	118.7343	124.2382	118.6671	124.2840

(iii) Orthotropic thermal conductivity of both layers ($k_{1,xy} = 75$ W/m-K, $k_{1,z} = 150$ W/m-K, $k_{2,xy} = 490$ W/m-K, $k_{2,z} = 390$ W/m-K)

h_c (W/m^2-K)	Analytical		FEA	
	$\bar{T}$ (°C)	$\hat{T}$ (°C)	$\bar{T}$ (°C)	$\hat{T}$ (°C)
1.00E+09	84.8056	89.2295	84.7557	89.2808
2.00E+08	87.9156	92.3903	87.8658	92.4417
1.00E+08	91.2734	95.8077	91.2237	95.8593
6.67E+07	94.2497	98.8417	94.2000	98.8933
3.33E+07	101.7971	106.5565	101.7475	106.6082
2.00E+07	109.9996	114.9769	109.9501	115.0287
1.00E+07	126.0090	131.5225	125.9596	131.5744
6.67E+06	138.7136	144.7592	138.6642	144.8113

The average and maximum temperatures predicted by the FEA and analytical models agree to within 0.13% for all of the values of interface conductance evaluated. The difference between the cases for isotropic thermal conductivity of both layers with $k_2 = 420$ W/m-K and orthotropic SiC thermal conductivity with $k_{2,xy} = 490$ W/m-K and $k_{2,z} = 390$ W/m-K is minor because the isotropic value used in the calculations is close to the geometric mean of the orthotropic values as given by Eq. (51). The shallow slope of the temperature vs. interface conductance curve shown in Figure 37 suggests that reduction of the interfacial resistance below a value of 10^{-8} m^2-K/W will not result in a significant decrease in the peak device temperature. However, increase of the thermal resistance beyond ~2 x 10^{-8} m^2-K/W will result in significantly higher device temperatures due to the rapid increase in spreading resistance.

Multiple Heat Sources

Because many GaN HEMTs use multi-finger structures to increase the power density of RF amplifier MMICs, evaluating the temperature rise and thermal resistance due to multiple heat sources is critical. When the heat sources are located close to one another, the temperature rise

due to thermal crosstalk between the sources may be much greater than each heat source alone. In the two following multiple heat source validation studies, ten heat sources of 0.5 μm length by 150 μm width, 50 μm source-to-source spacing, and 0.375 W magnitude were modeled on the device structure shown in Figure 39 with the same general dimensions as the previous single heat source studies. First, a device with perfect interface conductance between the GaN layer and SiC substrate and a fixed base temperature of 25 °C was studied.
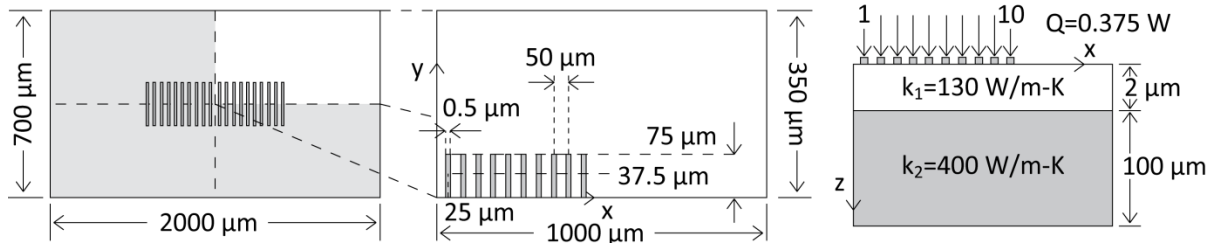
Figure 39: Device layout for multiple (ten) heat source model validation study

In this case, the 3-D FEA model in COMSOL Multiphysics™ required ~5.5 million cubic tetrahedral elements to resolve the ten heat sources accurately. Proper convergence of the FEA model was validated by progressively refining the mesh with negligible change in the temperature. The influence coefficient method was implemented in the analytical model to capture the thermal crosstalk between multiple heat sources. The source plane temperature profile along the x-axis is shown for both the FEA and analytical methods in Figure 40 with a closer view of the innermost and highest temperature heat source shown in the inset.
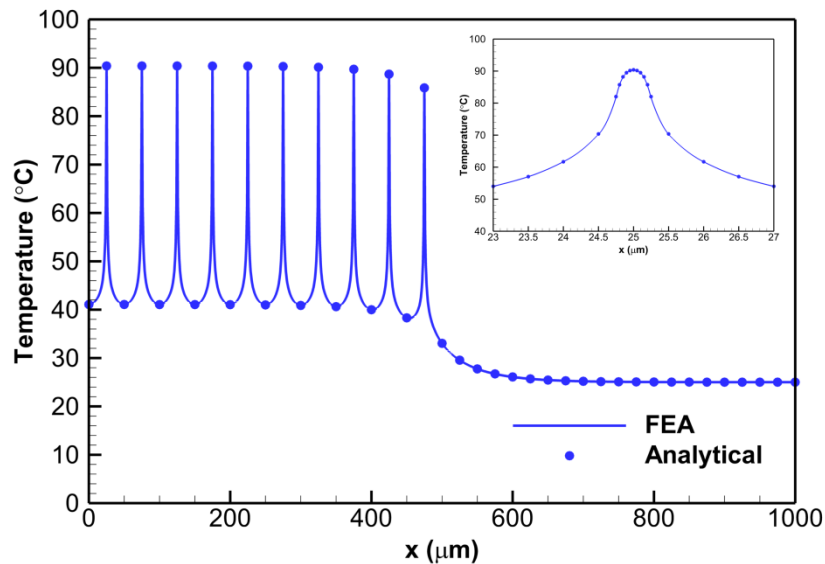
Figure 40: Temperature profile comparison for ten heat sources with perfect interface contact
Model implemented isotropic thermal conductivity ($k_1 = 130$ W/m-K, $k_2 = 400$ W/m-K) and uniform base temperature $T_{base} = 25$ °C. Inset shows a closer view of the innermost (highest temperature) heat source.

The agreement between the FEA and analytical models is again very good, even in the sharply peaked region around the innermost (highest temperature) heat source. For reference, the average

and centroidal temperatures of the ten heat sources are given in Table 3. The average and centroidal temperatures of the heat sources calculated by the FEA and analytical models differ by less than 0.2% and 0.1%, respectively.

Table 3: Average and centroidal temperature for first multiple heat source validation study
The model features are isotropic thermal conductivity (k_1 = 130 W/m-K, k_2 = 400 W/m-K), perfect interface contact and fixed base temperature of 25 °C. Heat sources #1 and #10 represent the innermost and outermost heat sources, respectively.

Source no.	Analytical		FEA	
	$\bar{T}$ (°C)	$\hat{T}$ (°C)	$\bar{T}$ (°C)	$\hat{T}$ (°C)
1	85.7429	90.3903	85.6503	90.4505
2	85.7410	90.3883	85.6489	90.4482
3	85.7359	90.3831	85.6434	90.4421
4	85.7238	90.3705	85.6312	90.4287
5	85.6957	90.3410	85.6023	90.4002
6	85.6298	90.2715	85.5355	90.3303
7	85.4743	90.1057	85.3802	90.1645
8	85.1015	89.7021	85.0082	89.7612
9	84.1800	88.6818	84.0864	88.7418
10	81.6877	85.8465	81.5951	85.9047

In a second multiple heat source study, the effect of orthotropic SiC thermal conductivity ($k_{2,xy}$ = 490 W/m-K and $k_{2,z}$ = 390 W/m-K), finite interfacial conductance, and convection at the substrate base are demonstrated on a device structure with ten discrete heat sources. The dimensions of the device are the same as the model described in the previous study and shown in Figure 39. However, the thermal conductivity of the GaN layer was changed to k_1 = 150 W/m-K for comparison to a following validation study that accounts for thermal conductivity variation with temperature. In addition, a finite interfacial conductance of 9.28 x 10^7 W/m^2-K corresponding to experimental measurements by *Sarua et al. (2007)* [32] and base convection coefficient of 3.27 x 10^5 W/m^2-K were specified to account for the liquid micro-channel convective cooling (h_s = 6.5 x 10^4 W/m^2-K and T_∞ = 20 °C) and the 1D conduction resistance of a 50 µm thick SnAg solder layer (k = 33 W/m-K) described by *Garven and Calame (2009)* [15]. The temperature distribution in the source plane was computed by the same COMSOL Multiphysics™ model with ~5.5 million elements and appropriate boundary conditions and the analytical model via the spreading function ϕ given by Eq. (60) and the influence coefficient method. The convergence of the FEA solution for this second validation study was also independently confirmed. The source plane temperature profile along the x-axis is shown in Figure 41 with a closer view of the innermost heat source shown in the inset and the exact values of the average and centroidal temperatures are given in Table 4.
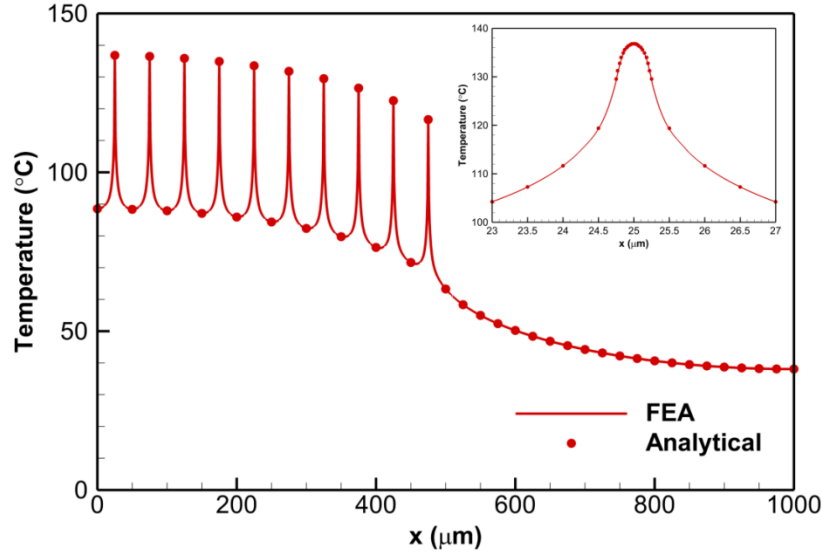
Figure 41: Temperature profile comparison for ten heat sources with finite interfacial conductance Models implemented isotropic k_1 = 150 W/m-K, orthotropic k_2 ($k_{2,xy}$ = 490 W/m-K, $k_{2,z}$ = 390 W/m-K), h_c = 9.28 x 10^7 W/m²-K, h_s = 3.27 x 10^5 W/m²-K and T_∞ = 20 °C.

Table 4: Average and centroidal temperature for multiple heat source validation study with finite interfacial conductance and convection at the sink.
Details of the model are given in the paragraph and Figure 41 above.

Source no.	Analytical		FEA	
	$\bar{T}$ (°C)	$\hat{T}$ (°C)	$\bar{T}$ (°C)	$\hat{T}$ (°C)
1	132.0707	136.8180	131.9875	136.8685
2	131.7606	136.5063	131.6777	136.5563
3	131.1312	135.8735	131.0480	135.9225
4	130.1635	134.8995	130.0802	134.9477
5	128.8255	133.5506	128.7414	133.5995
6	127.0670	131.7733	126.9820	131.8219
7	124.8082	129.4810	124.7234	129.5298
8	121.9120	126.5204	121.8281	126.5697
9	118.0943	122.5617	118.0102	122.6119
10	112.5145	116.6032	112.4312	116.6515

The average and centroidal temperatures of the heat sources calculated by the two methods differ by less than 0.1% for this second multiple source validation study. The use of these boundary conditions is more realistic than perfect interface contact and a fixed base temperature for calculating the temperature rise GaN HEMTs. This validation study demonstrates the flexibility of the analytical model in accounting for a thermal resistance at the base of the device due to additional package layers or thermal management devices.

3.2.4 Computational Efficiency of the Analytical Model

Although the finite element method is an accurate and flexible technique for solving conduction heat transfer problems, it is often much less computationally efficient than closed-form analytical solutions. In particular, solving for the temperature field in a domain for which the heat source is much smaller than the other characteristic lengths in the model requires a high density of elements near the heat source. As the number of discrete heat sources increases, the number of finite elements and computation time required to achieve an accurate solution also increases. This requirement of very high mesh density makes FEA models less feasible and analytical models much more attractive for solving heat conduction problems involving many discrete heat sources.

In the case of high-power solid state electronic devices, such as GaN HEMTs, multi-finger structures are often used to increase the power density of the device and minimize the area of substrate required. Since power is dissipated near the gate contact in each unit cell, each gate finger represents a discrete heat source with a characteristic dimension (~1 µm) that is two orders of magnitude smaller than other characteristic lengths in the device structure (~100 µm). Furthermore, concurrent electrical and thermal design of GaN RF MMICs and power electronics has motivated researchers to pursue closed-form analytical expressions that can be evaluated much more quickly than FEA models [23]. In this work, the computation time required to calculate the mean temperature of the heat sources was compared between the FEA and analytical models for multi-finger GaN HEMTs described in second multiple heat source validation study with isotropic k_1, orthotropic k_2, finite interface conductance, and convection at the base. As shown in Figure 42, the computation time for the analytical model remained fairly constant at ~22 seconds while the FEA model required ~4 minutes to ~8 hours for two to ten heat sources. The rapid increase in the computation time required for the FEA model was due to the increase in number of finite elements required to achieve an accurate solution with each additional heat source. For very large models, such as the multiple heat source model with ten heat sources and ~5.5 million elements, the computation time can be prohibitively long on desktop workstations.
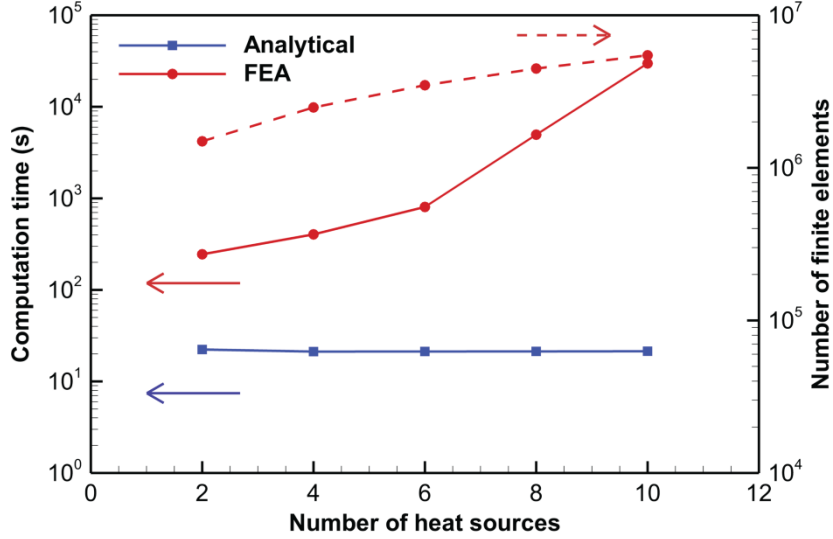
Figure 42: Comparison of the time required to compute the mean temperature rise of the heat sources for models with multiple heat sources
The calculation was carried out for the same type of model as the second multiple heat source study (isotropic k_1, orthotropic k_2, finite interfacial conductance, and convection at the sink. The left axis shows the computation time (s) for the FEA and analytical models (solid lines). The right axis shows the number of finite elements required for each FEA model (dashed line).

In the past, researchers have sought other methods other than infinite series solutions for solving multi-dimensional heat conduction problems because of the large number of terms needed for the summations to converge. Evaluation of the double summation in Eq. (53), for instance, may require a significant amount of computation time when each index is summed over 10^4 terms. Furthermore, calculation of the temperature rise due to N_s discrete heat sources requires $O(N_s^2)$ operations to account for the mutual temperature rise due to adjacent heat sources. In the case of the analytical model for temperature rise in GaN HEMTs, a number of factors reduce the computation time required to solve the 3D thermal conduction problem. First, the dimensions of each of the heat sources (nominally taken as the gate area in this work) are typically the same for a multi-finger HEMT. Thus, the single summation in the y-direction over the index n in Eq. (54) is identical for all heat source pairs and only needs to be computed once. The inner summation in the y-direction over the index n in the double summation in Eq. (54) is also identical for each of the heat source pairs and only needs to be computed once for each value of the index m. For each heat source pair i, j using the influence coefficient method for the mean source temperature rise given by Eq. (59), only two single summations in the x-direction over the index m need to be computed. This effectively means that the computation time of the analytical model does not increase significantly with the number of heat sources because these two single summations require relatively little time to compute. In addition, the symmetry of the influence coefficient matrix $\boldsymbol{F}$ reduces the number of pairs computed to approximately half that of the full matrix. These properties of the analytical model make it highly attractive for calculating temperature rise in multi-finger GaN HEMTs because it is up to two orders of magnitude faster to compute than an equivalent FEA model.

79

3.3 Important Factors in Thermal Modeling

One of the most valuable aspects of thermal modeling of temperature rise in GaN HEMTs is understanding how the temperature distribution depends on various factors of the device geometry and properties. There have been many parametric studies in the past in the field of thermal issues in GaN HEMTs but many do not provide helpful physical insight needed to improve thermal management technologies. The following section of this thesis attempts to give an explanation of how several key factors affect the temperature distribution in GaN HEMTs with an emphasis on thermal spreading resistance. The number of factors discussed in this section is not exhaustive but covers a few of the most important ones, including substrate material, GaN-substrate thermal boundary resistance, size of the heat source, and thermal boundary conditions.

3.3.1 Near-junction Thermal Spreading Resistance

The importance of near-junction thermal spreading resistance in GaN HEMTs, particularly in RF and high frequency applications, cannot be overemphasized. Unfortunately, the description of temperature rise in GaN HEMTs and the discussion of its dependency on a number of factors is rarely discussed in the context of thermal spreading resistance. Understanding thermal spreading resistance and its relationship to the layout of GaN HEMTs and thermal properties of the epitaxial layers help to clearly see how many factors affect temperature rise.

In the heat transfer literature, thermal spreading resistance is a concept developed to describe the effect of heat spreading from a source with a small area to a sink with a much larger area. The total thermal resistance between the heat source and heat sink is

$$R_{total} = \frac{\bar{T}_{source} - \bar{T}_{sink}}{Q} \tag{73}$$

where $\bar{T}_{source}$ and $\bar{T}_{sink}$ are the average temperature over the source and sink areas, respectively, and Q is the rate of heat transfer at the source [74]. The total resistance can be conceptualized as the sum of two resistances, the 1D conduction resistance between the contact plane and the sink plane and the spreading resistance, *i.e.,*

$$R_{total} = R_{sp} + R_{1D} \tag{74}$$

where R_{1D} and R_{sp} are the thermal 1D and spreading resistances, respectively. When the 1D conduction resistance is defined as

$$R_{1D} = \frac{\bar{T}_{cp} - \bar{T}_{sink}}{Q} \tag{75}$$

where $\bar{T}_{cp}$ is the average temperature of the contact or source plane, the thermal spreading resistance is defined as

$$R_{1D} = \frac{\bar{T}_{source} - \bar{T}_{cp}}{Q}. \tag{76}$$

In GaN HEMTs, the heat source with a characteristic length of ~1 µm is much smaller than the other characteristic dimensions of the device of ~100 µm. Thus, thermal models and experimental temperature measurements often show sharply peaked temperature distributions in the channel with the center of the peak at the edge of the gate on the drain side. As an example, the FEA model for the first single heat source validation study gives $\bar{T}_{source}$ = 80.09 °C, $\bar{T}_{cp}$ = 25.28 °C, and $\bar{T}_{sink}$ = 25 °C. From the definitions given above, R_{1D} = 0.758 °C/W and R_{sp} = 146 °C/W. Thus, the near-junction (within 100 µm of the heat source) thermal spreading resistance is two orders of magnitude larger than the 1D conduction resistance.

Although the specific ratio between the 1D conduction and spreading resistances changes with the epitaxial structure, materials properties, and device layout, the thermal spreading resistance in the near-junction region of GaN HEMTs tends to be the dominant thermal resistance. Advanced and next generation thermal management technologies must be primarily aimed at reducing the thermal spreading resistance in the near-junction region.

3.3.2 Substrate Thermal Conductivity

Introduction of high thermal conductivity substrates is one of the most commonly used and effective strategies to reduce the thermal spreading resistance in GaN HEMTs because solid materials are relatively simple to implement. A high thermal conductivity solid improves thermal spreading in the region within ~1 to 2 µm of where the heat is generated near the AlGaN/GaN interface. Various complications, however, also arise from the epitaxial growth of GaN on foreign substrates or from substrate transfer processes. High thermal interfacial resistances may mitigate the benefit of a high thermal conductivity because the interface is so close to the heat source.

In the past, GaN HEMTs were fabricated on SiC or sapphire substrates due to the similar (wurtzite) crystal structure with GaN and relatively small (~3%) different in lattice constants. A substantial amount of research effort has been devoted to optimizing the epitaxial growth process and improving the quality of GaN device layers. GaN epitaxial growth has also been introduced on Si substrates because of the demand for lower cost GaN-based electronics; devices can be fabricated on less expensive and larger area Si wafers than either SiC or sapphire. Polycrystalline and single crystal diamond have also been introduced for GaN-based electronics because of the very high thermal conductivity of diamond (~2000 W/m-K) for the highest performance applications in which cost is not the driving factor. Ultimately, the application space, budget, and desired properties of the GaN device determine the most suitable substrate choice. The analysis presented in this thesis is intended to give an illustration of how the substrate thermal conductivity changes the temperature distribution. In Figure 43, the thermal spreading and 1D conduction resistances are shown as a function of substrate thermal conductivity for a single finger HEMT with the same layout as the first example in the previous section.
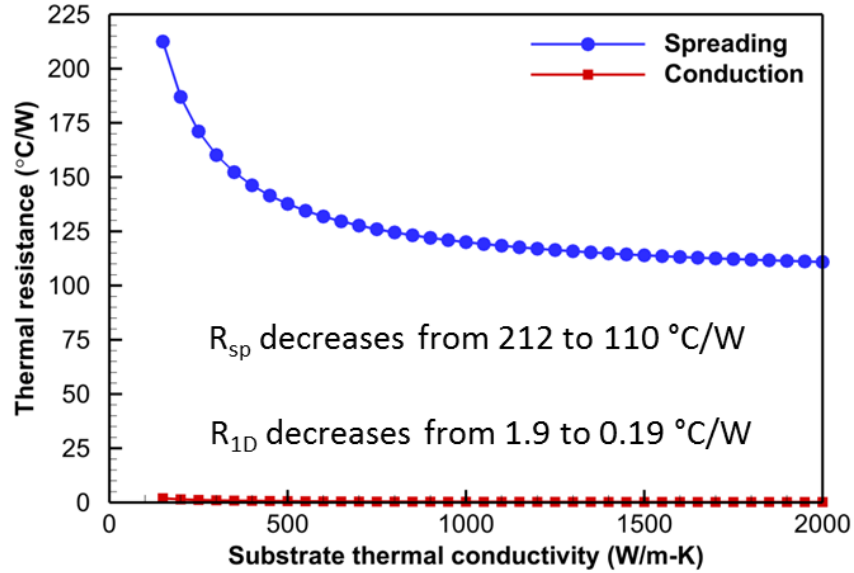
Figure 43: Thermal resistance as a function of substrate thermal conductivity for single finger HEMT computed with analytical model

As described in the discussion of the importance of thermal spreading resistance, the thermal spreading resistance is approximately two orders of magnitude larger than the 1D conduction resistance for thermal conductivity values in the range of 150 to 2000 W/m-K. When the substrate thermal conductivity is increased from 150 to 2000 W/m-K, the 1D conduction resistance decreases significantly from 1.9 to 0.19 °C/W as expected. However, the decrease in thermal spreading resistance from 212 to 110 °C/W far outweighs the change in 1D conduction resistance and is ultimately responsible for the decrease in total thermal resistance. The main effect of increased substrate thermal conductivity is to reduce the thermal spreading resistance, which is the dominant component of the total thermal resistance.

It is also interesting to note that the thermal spreading resistance declines most rapidly as the substrate thermal conductivity is increased from 150 to 500 W/m-K. As the thermal conductivity continues to increase, the thermal spreading resistance reaches a plateau; further increase in thermal conductivity results in a relatively minor decrease in spreading resistance. In this case, the spreading resistance does not change significantly for substrate thermal conductivity values above 1000 W/m-K. This effect should also occur in large, multi-finger HEMTs but very high thermal conductivity substrates (~2000 W/m-K) should provide more of a benefit than demonstrated here because of the higher total power dissipation and need for efficient spreading.

3.3.3 GaN-substrate Thermal Boundary Resistance

The growth of GaN on foreign substrates results in a thermal interfacial or boundary resistance between GaN and its substrate. This thermal boundary resistance (TBR) is the result of thermal resistances at the interface between dissimilar materials (*e.g.,* GaN and AlN) due to different phonon dynamics and poor crystalline quality near the interfaces. The flow of heat from one material to another experiences a resistance because of the non-ideal transmission of phonons across the interface; this can be described partly by the acoustic mismatch model (AMM) or

diffuse mismatch model (DMM) [63]. There should also be a conduction resistance associated with the buffer or nucleation layers for GaN-based epitaxial structures that adds to the GaN-substrate thermal resistance. Often, the thermal conductivity of these layers is much lower than that of the bulk material because of a high density of defects caused by the crystal nucleation and growth process. The effective thermal boundary resistance between GaN and its substrate is essentially the sum of these resistances: interfacial resistance between dissimilar materials and conduction resistance of the layers themselves. A schematic showing the components of the GaN-substrate TBR with TEM images of the interface from the literature is shown in the figure below.
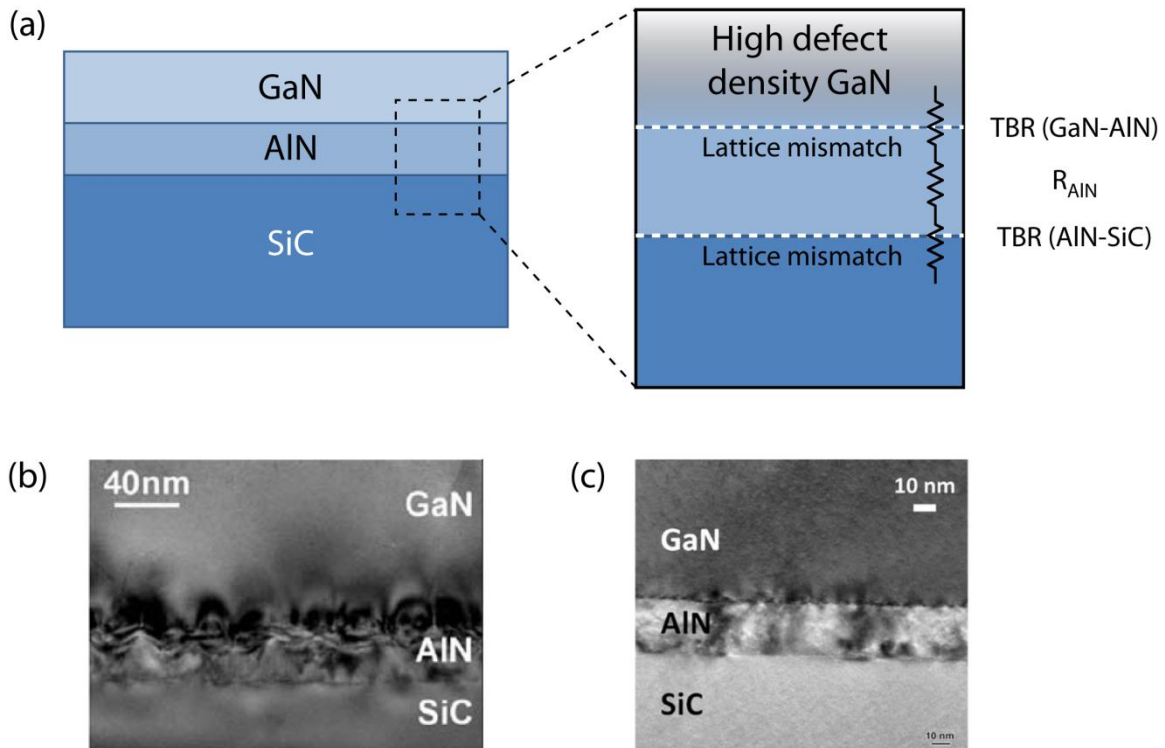
Figure 44: Schematic of GaN-substrate thermal boundary resistances with TEM images
(a) Schematic of GaN-on-SiC epitaxy, (b) and (c) TEM images of GaN-substrate interface [75]-[76]

The limited experimental measurements of GaN-substrate TBR indicate that the value is much higher than that predicted by either the AMM or DMM and that highly defective regions must play a role in the high thermal resistance. Most of the data reported in the literature was obtained from micro-Raman thermometry measurements [32]-[33] for the GaN-on-SiC epitaxy since it is the most common for high power RF devices. Typical values range from 10^{-8} to 3 x 10^{-8} m^2-K/W at room temperature and a strong temperature dependency, $TBR \sim T^{2.76}$, has been reported by *Sarua et al. (2007)* [32]. The single finger analytical model for a GaN-on-SiC HEMT with 100 μm substrate was used to study the impact of GaN-substrate TBR on the spreading and 1D conduction resistances as shown in Figure 45. The thermal spreading resistance is calculated using the average temperature of the heat source from the analytical thermal model.

83
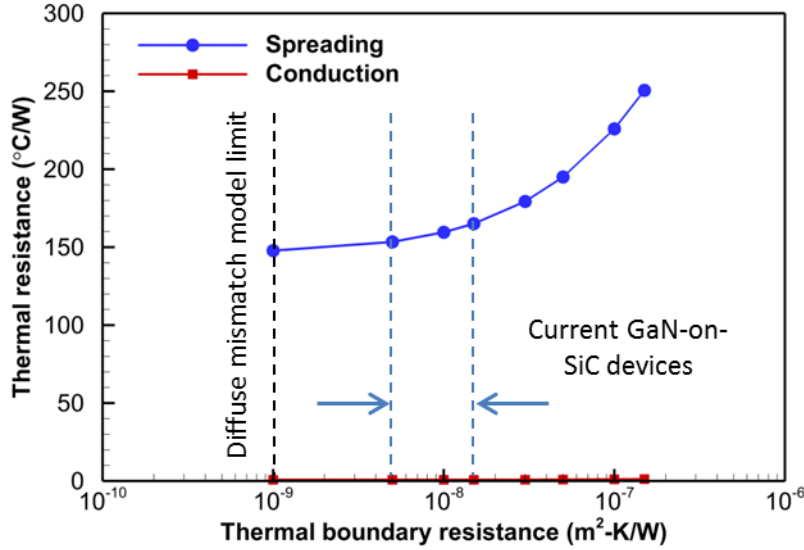
Figure 45: Thermal resistance as a function of GaN-substrate TBR for single finger GaN-on-SiC HEMT

As in the previous discussion of the effect of substrate thermal conductivity, the thermal spreading resistance is approximately two orders of magnitude greater than the 1D conduction resistance. The GaN-substrate TBR can greatly increase the spreading resistance because it is within 1 to 2 μm of the heat source at the AlGaN/GaN interface. As the TBR is decreased from 2×10^{-7} to 10^{-9} m^2-K/W, the thermal spreading resistance decreases from 250.6 to 147.8 °C/W (41%). Prior benchmarking studies of GaN-SiC TBR measured indirectly by micro-Raman spectroscopy indicated that the TBR of wafers from academic and commercial suppliers may exceed 2×10^{-8} m^2-K/W at room temperature [33]. This has been cited to lead to additional temperature rises of ~20 to 30% over the case with zero TBR. However, more recent experimental measurement data obtained from time domain thermal reflectance (TDTR) reported TBR values as low as ~4 to 5×10^{-9} m^2-K/W at room temperature for GaN-on-SiC epitaxy grown by Raytheon [75]. Although surprising, these measurements demonstrate that the GaN-on-SiC has been optimized by some commercial suppliers and may not be as much of a concern as in the past. It is unlikely that the GaN-SiC TBR can be further reduced since the DMM model predicts a boundary resistance of ~10^{-9} m^2-K/W for the GaN-AlN and AlN-SiC interfaces; the conduction resistance of the AlN nucleation layer must also be taken into account. For the data shown in the figure above, reduction of the GaN-SiC TBR from 5 to 1×10^{-9} m^2-K/W would only reduce the spreading resistance from 153.4 to 147.8 °C/W (~3.7%). Thermal boundary resistance may also play a significant role in optimizing the effectiveness of GaN-on-diamond structures, but this has not been thoroughly investigated to date [19].

3.3.4 Size of the Heat Source

The importance of the heat source in GaN-based electronics has already been thoroughly discussed in Chapter 2 of this thesis with respect to electro-thermal modeling. This discussion is intended to provide additional physical insight in light of the development of an analytical expression for temperature rise and thermal spreading resistance in GaN HEMTs. Usually, the heat source in GaN HEMTs is modeled as a uniform surface flux at the top of the GaN layer or a

region of uniform heat generation. As a demonstration of the impact of the heat source length on the thermal spreading and 1D conduction resistances, the thermal resistance is plotted against the heat source length in Figure 46. The layout used for the thermal model is the same 3D single heat source as in the previous cases with the heat source length swept from 0.1 to 2 μm (no GaN-SiC interfacial resistance and substrate base fixed at 25 °C).
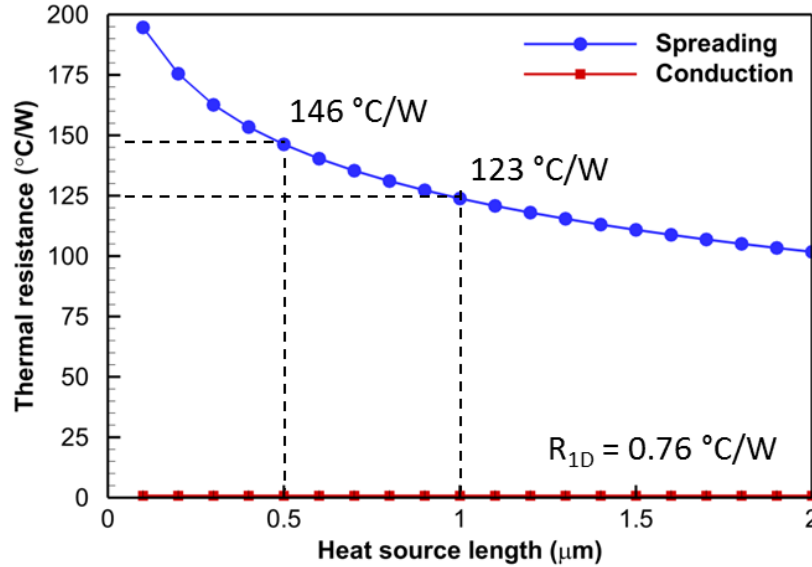
Figure 46: Thermal resistance as a function of heat source length for single finger GaN-on-SiC HEMT

For the baseline model with a heat source length of 0.5 μm (typically assumed to be equal to the gate length for a RF GaN HEMT), the thermal spreading resistance of 146 °C/W is two orders of magnitude greater than the 1D conduction resistance of 0.76 °C/W. As the heat source length increases, the thermal spreading resistance decreases because the heat is essential more well spread from the source. The change in the heat source length has no effect on the 1D conduction resistance, which depends only on the thermal properties of the structure. The thermal spreading resistance depends to the leading order on the inverse square of the heat source length, *i.e.,*

$$R_{sp} \sim \frac{1}{c^2} \tag{77}$$

where c is the heat source length. The relationship given above is not exact for 3D conduction problems because spreading the in the y-direction also plays a significant role. However, it can be seen that the thermal spreading resistance and device temperature increases more dramatically as the heat source length is less than 0.5 μm. This motivates the importance of correctly estimating the heat source length for high frequency HEMTs with short gate lengths in order to produce an accurate thermal model. As discussed in the chapter on electro-thermal modeling, both this thesis and a previous electro-thermal modeling investigation [22] found that the heat source is spread over an area much longer than the gate.

85

3.3.5 Substrate Base Thermal Boundary Condition

In previous electro-thermal and thermal modeling research, the importance of the thermal boundary condition at the base of the substrate has been significantly underappreciated. The top and sides of a packaged GaN HEMT are usually surrounded by air and cooled by natural convection; thus, the thermal resistance to the ambient through the top and sides is much higher than the thermal resistance to the substrate base. The substrate base, which is bonded to the package and cooled by a heat sink, is the dominant heat sinking path in GaN HEMTs. The thermal boundary condition (BC) specified at the substrate base can play a major role in shaping the temperature distribution and determining the peak temperature in the device.

GaN HEMTs are often evaluated for electrical performance on a probe station with a large, temperature-controlled chuck. In the research setting, these HEMTs may also be smaller devices with a few gate fingers for demonstration of advances in electrical characteristics. Thus, it is reasonable to assume that the base of the substrate remains at a fixed, uniform temperature equal to the ambient. For packaged and deployed multi-finger HEMTs, the base of the substrate is often attached to a spreader with AuSn solder; the spreader is bonded with silver epoxy to the package. The, the package is attached a heat sink with thermal grease and the heat sink rejects heat to the ambient by natural or forced convection. In this second situation, it is most appropriate to specific a heat transfer coefficient or finite conductance from the base to the ambient. These two situations, *i.e.,* a uniform prescribed temperature and heat transfer coefficient, are depicted in Figure 47.
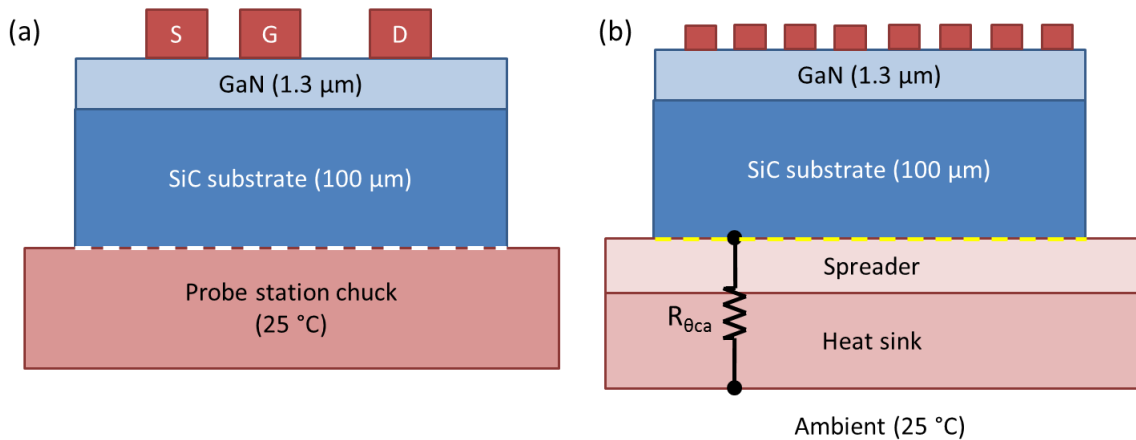
Figure 47: Schematic of GaN HEMT under different measurement conditions
(a) On probe station chuck in which uniform temperature is appropriate and (b) in package in which heat transfer coefficient is appropriate

A thermal model was developed in COMSOL Multiphysics 4.2a for a 20 finger GaN HEMT similar to those in the multi-finger FEA and analytical model comparison given previously in this chapter, with constant values of k_{GaN} = 150 W/m-K and k_{SiC} = 420 W/m-K. A quarter-model reduction was used to reduce the computation expense because of the symmetry in the device layout. The surface heat flux over each of the 0.5 μm x 75 μm heat sources was set to 10^{10} W/m^2 in accordance with a power dissipation of 5 W/mm normalized to the gate width. The thermal resistance between the substrate base and the ambient was taken to be a typical value of 5 °C/W, representing the sum of the thermal resistances of the package, thermal interface

86

materials, and heat sink. In the case of the uniform base temperature, the base temperature was prescribed to be 100 °C to represent this thermal resistance with respect to an ambient temperature of 25 °C. For the case of a heat transfer coefficient at the base, the heat transfer coefficient was set equal to 1.43 x 10^5 W/m^2-K as calculated from the thermal resistance of 5 °C/W and the area of the substrate base. The temperature distribution in the quarter-model device is shown in Figure 48 for the two types of boundary conditions discussed.
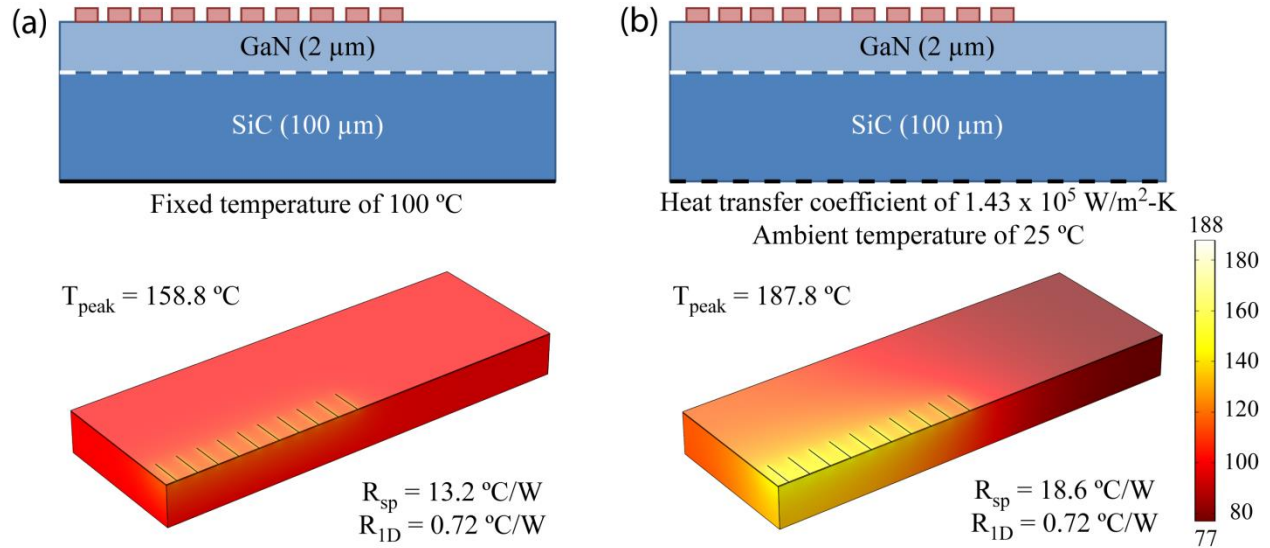
Figure 48: Temperature distribution (°C) for different substrate boundary conditions (a) Prescribed base temperature of 100 °C and (b) heat transfer coefficient of 1.43 x 10^5 W/m^2-K for a thermal resistance of 5 °C/W and ambient temperature of 25 °C

As can be seen in Figure 48, the temperature distribution is quite different for the two thermal boundary conditions used. When the substrate base temperature is prescribed to a uniform value of 100 °C, the peak temperature is 158.8 °C and the rise in temperature is highly localized to the region around the heat sources. Most of the device is at a temperature very close to the temperature prescribed at the base. In contrast, the peak temperature in the heat transfer coefficient case is 187.8 °C (29 °C higher) and the temperature rise above the base is more spread throughout the device. This behavior is observed because the finite heat transfer coefficient at the base creates an additional spreading resistance (18.6 °C/W versus 13.2 °C/W for the uniform base temperature case). The uniform base temperature represents the best case in which the spreading and heat sink operate ideally. A finite base heat transfer coefficient, however, is more realistic for large, multi-finger packaged devices. The 1D temperature distribution along the channel in the center and highest temperature part of the device is shown in Figure 48.
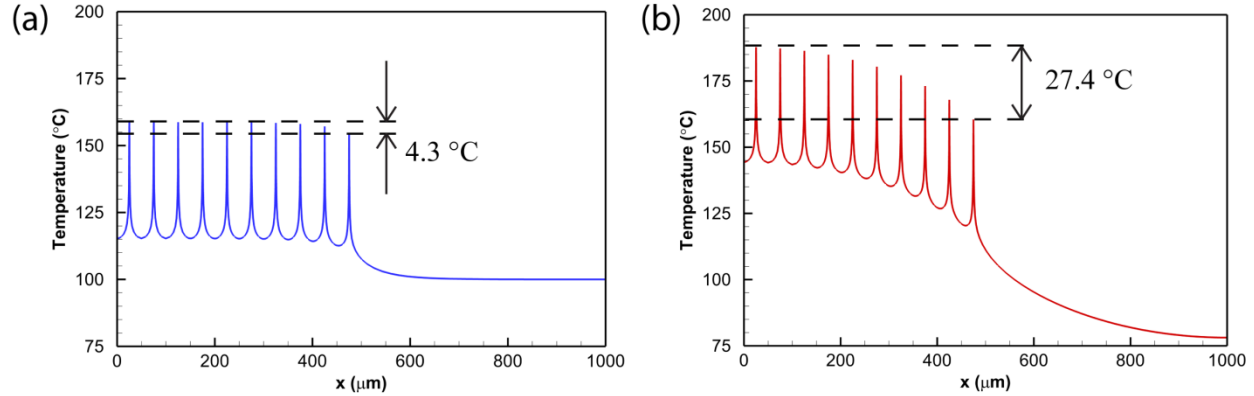
**Figure 49: 1D temperature distribution for different substrate boundary conditions
(a) Prescribed base temperature of 100 °C and (b) heat transfer coefficient of 1.43 x 10^5 W/m^2-K for a
thermal resistance of 5 °C/W and ambient temperature of 25 °C**

The 1D temperature distribution along the channel highlights an important behavior with respect to the temperature variation between the fingers. When the base temperature is prescribed to be a uniform value, the difference in temperature between the inner and outer heat sources is only 4.3 °C and the temperature of most of the gate fingers is almost the same. This is the behavior observed for multi-finger HEMTs by *Darwish et al. (2004)* [23] who found that only 3 to 5 gate fingers are required to apply the unit cell model with adiabatic planes. However, the difference in temperature between the inner and outer heat sources is 27.4 °C/W in the heat transfer coefficient case and there is significant variation in temperature between the fingers. This behavior was observed by *Garven and Calame (2009)* [15] but not explained in terms of the thermal spreading resistance. In this work, the substrate thermal boundary condition was found to be very significant and it is recommended to use the prescribed temperature boundary condition with caution. It is not merely sufficient to set the temperature of the base to an elevated value to account for the base to ambient thermal resistance in HEMT thermal models in which the thermal spreading resistance dominates. A finite heat transfer coefficient must be explicitly implemented at the substrate base in order to model the temperature distribution in the device.

3.3.6 Heat Flux in the Near-junction Region

Despite the large amount of research and large number of publications on thermal issues in GaN electronics, relatively little has been reported on the heat flux values in the near-junction region. To the best knowledge of the author, there are not any publications that discuss this topic in depth nor that provide helpful metrics on the heat flux for developing advanced thermal management technologies. The results reported in this section are a brief introduction to understanding the heat flux in the near-junction region of multi-finger GaN HEMTs and illustrate the challenges associated with thermal management within 100 μm of the heat source.

Heat flux data taken from the multi-finger HEMT FEA thermal model in the previous section with the heat transfer coefficient boundary condition were used to generate the heat flux plots shown below. The heat flux magnitude in the z-direction (downward in the page) is shown in Figure 50 for a 2D cross-section at the center of the quarter-model on a log scale.
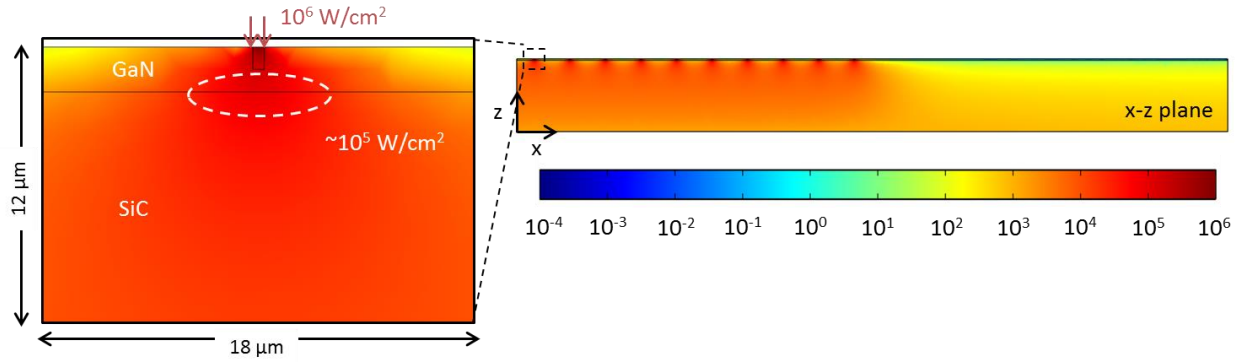
Figure 50: Heat flux magnitude in the z-direction (W/cm^2) for multi-finger GaN HEMT

As noted by *Garven and Calame (2009)* [15], the heat flux dissipated by the source is ~10^6 W/cm^2 for a typical RF GaN HEMT with 5 W/mm power dissipation. Due to aggressive spreading in the 1 to 2 μm thick GaN layer, the heat flux is reduced to ~10^5 W/cm^2 at the GaN/SiC interface. Because of the high thermal conductivity of SiC, this heat flux is further reduced to ~10^3 W/cm^2 at the base of the substrate. This is further demonstrated in the 2D cross-section plots in the xy-plane of the model for the GaN/SiC interface, a depth of 10 μm, and at the substrate base in Figure 51.
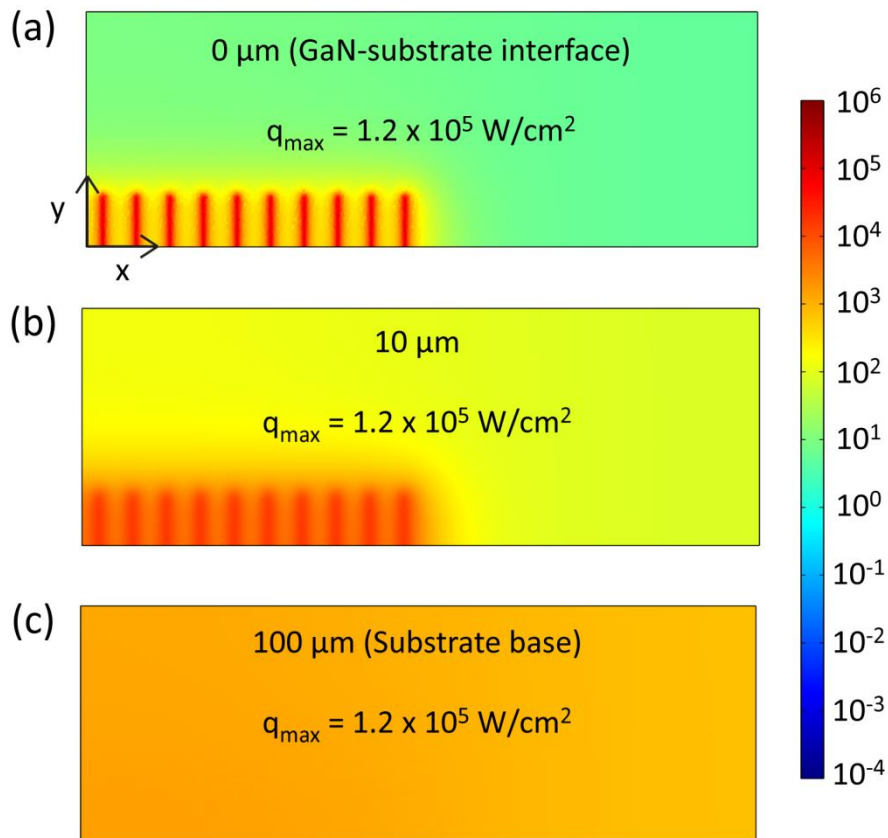
Figure 51: Heat flux magnitude in the z-direction (W/cm^2) viewed in the xy-plane
(a) GaN/SiC interface, (b) a depth of 10 μm into the substrate, and (c) at the substrate base

89

Although this is only an introductory discussion of the heat flux magnitude in the near-junction region, it provides order of magnitude values for the heat flux at the bottom of the GaN layer and the base of the substrate. Liquid-vapor phase change based solutions do not seem feasible at the bottom of the GaN layer because such solutions must manage heat fluxes on the order of 10^5 W/cm^2 (an order of magnitude greater than the kinetic limit for evaporation). Micro-channel cooler based solutions within the SiC substrate must manage fluxes ~10^3 to ~10^4 W/cm^2, which may also be prohibitively challenging. Even direct attachment of thermal management devices without sufficient spreading in a solid material at the base of the substrate requires innovative design of liquid passages or phase change devices [16]. More investigation of this topic will be performed in the future to understand the regime of thermal management solutions that are capable of managing these very high fluxes in the near-junction region.

3.4 Transition Temperature Rise in Power-switching Applications

GaN-based electronic devices have attracted increasing attention in power switching applications due to their high breakdown voltages and low on-resistances that exceed the capabilities of Si- and SiC-based devices. In high-voltage hard-switching transitions, the transient peak temperature may be higher than expected due to the simultaneous high voltage and high current. Although transient temperature rises have been actively investigated for Si- and SiC-based IGBTs and MOSFETs [78]-[80], relatively fewer efforts have focused on this topic for GaN HEMTs. Self-heating effects in GaN-on-SiC high frequency devices have been characterized for DC bias [28] and pulse mode [34] conditions with spatial and temporal resolutions of ~0.5 μm and ~10 ns, respectively. A number of studies have also investigated transient thermal behavior in GaN HEMTs with time-dependent power dissipation pulses [81]-[83]. However, these previous investigations mostly reported temperature rises on the order of microseconds to seconds rather than for the time period of ~20 to 50 ns over which switching transitions occur. Consideration of the large transient temperature rise during switching transitions in GaN power HEMTs needs to be investigated in more detail to ensure device reliability and lifetime.

In this thesis, transient thermal behavior during high-voltage turn-off transitions in GaN HEMTs is reported using a 3D thermal finite element model (FEM). The relative contributions of the epitaxy, substrate, and package thermal properties are described and devices on Si, sapphire, and SiC substrates are compared. The results highlight the importance of considering transient device operation to appropriately capture the thermal behavior.

3.4.1 Device Model

To obtain realistic predictions of GaN HEMT junction temperatures, we modeled a GaN-on-Si power switching device developed by *Saito et al. (2012)* [84] that is representative of high voltage (V_{BD} = 630 V) devices. The device structure, as shown in Figure 52(a), consists of a thin Al$_{0.25}$Ga$_{0.75}$N barrier, 1.5 μm GaN layer, and 2.3 μm transition/nucleation layers [84] on a 300 μm Si substrate [85] with transition current and voltage waveforms shown in Figure 52(b). The gate length and width are 1 μm and 36 mm, respectively. In addition, the on-resistance and threshold voltage reported for this device are 0.32 Ω and -6.5 V, respectively.
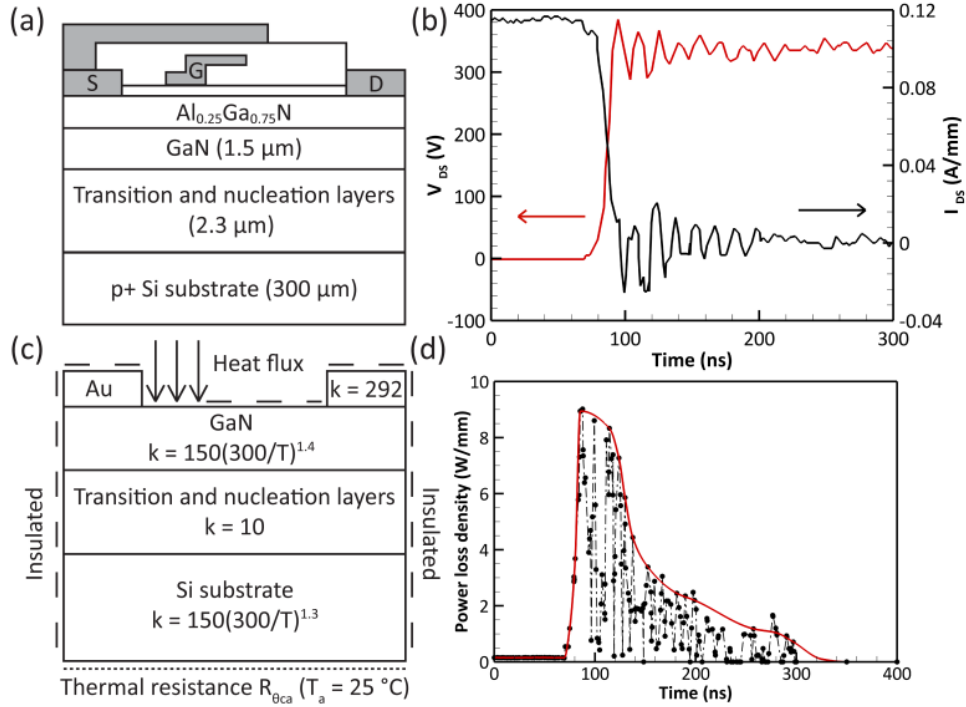
Figure 52: Device model details and power dissipation waveform for GaN power HEMT (a) Device structure, (b) current and voltage waveforms during switching transition, (c) thermal model structure with thermal conductivity in W/m-K, and (d) transient power dissipation profile derived from [84]

The HEMT epitaxial structure was simplified to include the GaN layer, transition/nucleation layers, source and drain contacts, and substrate as shown in Figure 52(c) and the power dissipation was modeled as a surface heat flux under the gate [15]. The current and voltage waveforms shown in Figure 52(b) [84] were multiplied and smoothed with piecewise cubic polynomials to obtain the transient heat flux shown in Figure 52(d). The boundary condition for the substrate base is defined with a thermal resistance of $R_{\theta ca}$ = 20 °C/W to an ambient temperature of 25 °C while the top is insulated (adiabatic) due to negligible heat loss via natural convection. The sides are also insulated because there is minimal lateral heat spreading due to the presence of many adjacent gate fingers [23]. Temperature-dependent thermal properties of the materials were taken from literature [32]. The effective thermal conductivity of the combined 2.3 μm thick transition and nucleation layers on Si substrates was estimated to be 10 W/m-K to represent an AlN/GaN superlattice (SL) and high defect density AlGaN/AlN nucleation layers [86]. The thermal conductivity of the SL and AlGaN/AlN nucleation layers are not known precisely and were set to be temperature-independent because of minimal sensitivity of thermal conductivity of SLs with changes in temperature above 25 °C.

In addition to devices on Si substrates, we studied the effect of sapphire and SiC substrates by replacing the 2.3 μm nucleation/transition layers and Si substrate with a 30 nm AlN nucleation layer and both sapphire and SiC substrates of varying thicknesses. The thermal conductivity of the AlN nucleation layer was set to 1.1 (sapphire) and 2.8 (SiC) W/m-K at 300 K with temperature dependence of $\sim T^{-2.76}$ to match experimental measurements of GaN-sapphire and

91

GaN-SiC effective thermal boundary resistances (TBR) [32].

3.4.2 Transition Temperature Rise

The surface heat flux corresponding to the on-state power dissipation in the channel results in a highly-localized hot spot in the GaN layer. The transient peak temperature profile for the GaN HEMT on a 300 µm Si substrate is shown in Figure 53 with the 3D temperature distribution in the on-state and at its peak temperature during the switching transition shown in the inset. Due to the thermal resistance of 20 °C/W between the base of the substrate and the ambient, the base temperature rises to 134 °C. The thermal resistance of the GaN layer, transition/nucleation layers, and substrate contribute an additional on-state temperature rise, ΔT_1, above the substrate base. During the turn-off transition of ~30 ns, a short burst of power raises the peak device temperature ΔT_2 (transition temperature rise) above the on-state peak temperature. In the off-state, the peak device temperature decreases rapidly due to heat spreading through the substrate and minimal ($\leq 10^{-3}$ W/mm) power dissipation in the off-state. Although the on-state temperature rise above the substrate base is only 20.1 °C, the transition temperature rise is approximately 68.3 °C.
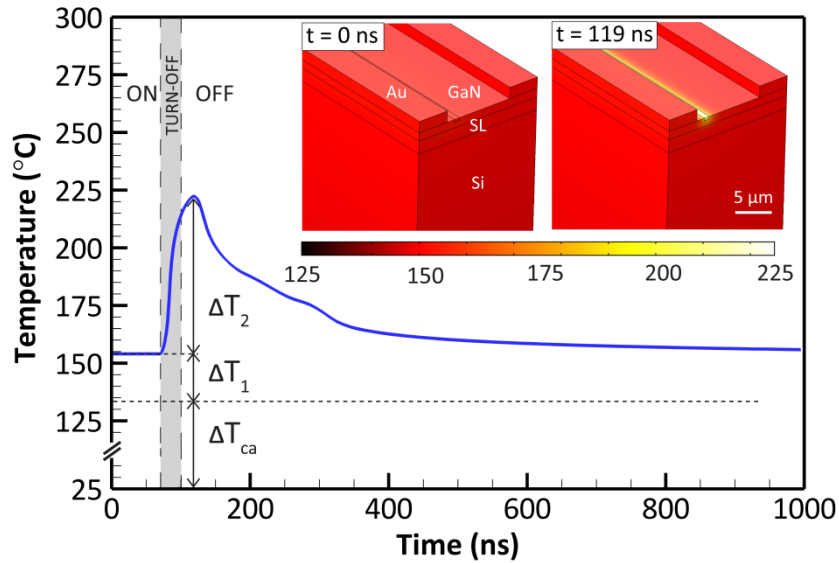
Figure 53: Transient peak temperature predicted at the center of the gate region for a GaN HEMT on a 300 µm Si substrate
Inset shows the 3D temperature (°C) distribution in the on-state (steady-state) and at its peak during the switching transition (transient)

3.4.3 Dependence on Substrate and Package Resistances

Transient temperature rise due to a heat flux applied over very short time scales results in localized heating of the area near the incident heat flux. The transition temperature rise associated with the intense burst of power during GaN HEMT switching is expected to be limited to the GaN layer since the characteristic thermal diffusion time, $\tau = L^2/\alpha$, where L and α are the thickness and thermal diffusivity of the layer (~45 ns for 1.5 µm thick GaN at 25 °C), is comparable to the ~30 ns turn-off transition time. The magnitude of the transition temperature rise depends primarily on the thermal properties of the GaN layer, which was found to be

dependent on the substrate and package thermal resistances. This result is due to the strong dependence of thermal diffusivity on temperature, which is determined at the onset of the switching transition by the substrate and package thermal resistances. Although silicon substrates are the most economical, sapphire substrates have been actively investigated owing to their superior electrical insulating properties despite the lower thermal conductivity. SiC has been primarily used for high-frequency devices, in which the steady-state power dissipation is much larger, due to its high thermal conductivity. The total peak-to-case, on-state, and transition temperature rises are shown in Figure 54 for GaN HEMTs on Si, sapphire, and SiC substrates of varying thicknesses. As expected, the peak device temperature increases with substrate thickness due to the increased thermal resistance of the substrate. We found the transition temperature rise to be much less dependent on the substrate thickness than the on-state temperature rise, as indicated by the slopes of the curves in Figure 54(b).
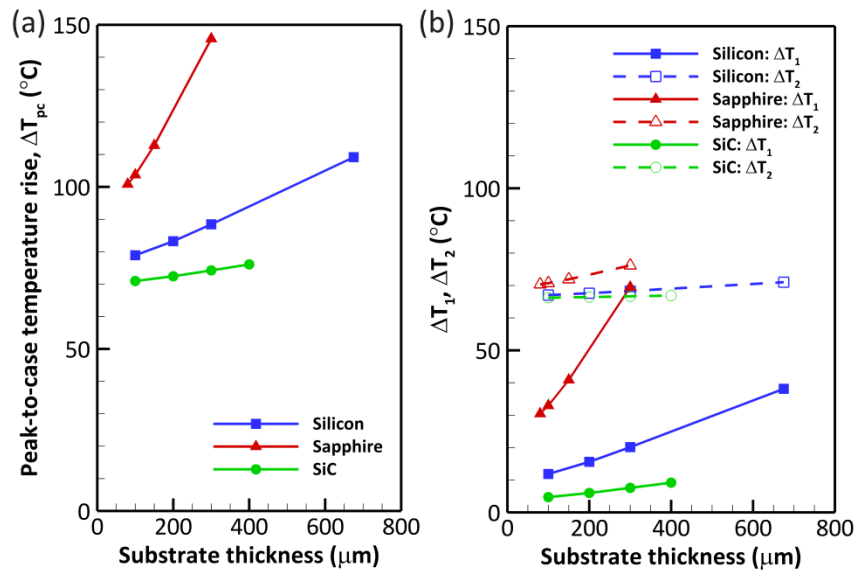
Figure 54: Dependence of temperature rise on substrate material and thickness
(a) Total peak-to-case, ΔT_{pc}, and (b) on-state, ΔT_1, and transition, ΔT_2, temperature rises as a function of substrate thickness for Si, sapphire, and SiC substrates. Substrate thermal conductivities values were $k(T) = 35(300/T)$ and $k(T) = 420(300/T)^{1.3}$ in W/m-K for sapphire and SiC, respectively.

Since the package thermal resistance (20 °C/W in the previous cases) is higher than the substrate thermal resistance, we predicted the case-to-ambient thermal resistance to have a significant effect on the transition temperature rise. Figure 55(a) and (b) show the transition temperature rise and associated average GaN thermal diffusivity values for Si and sapphire substrates with a fixed case-to-ambient resistance of 20 °C/W while Figure 55(c) and (d) show ΔT_2 and α_{GaN} for fixed substrate thicknesses and varying package resistance.
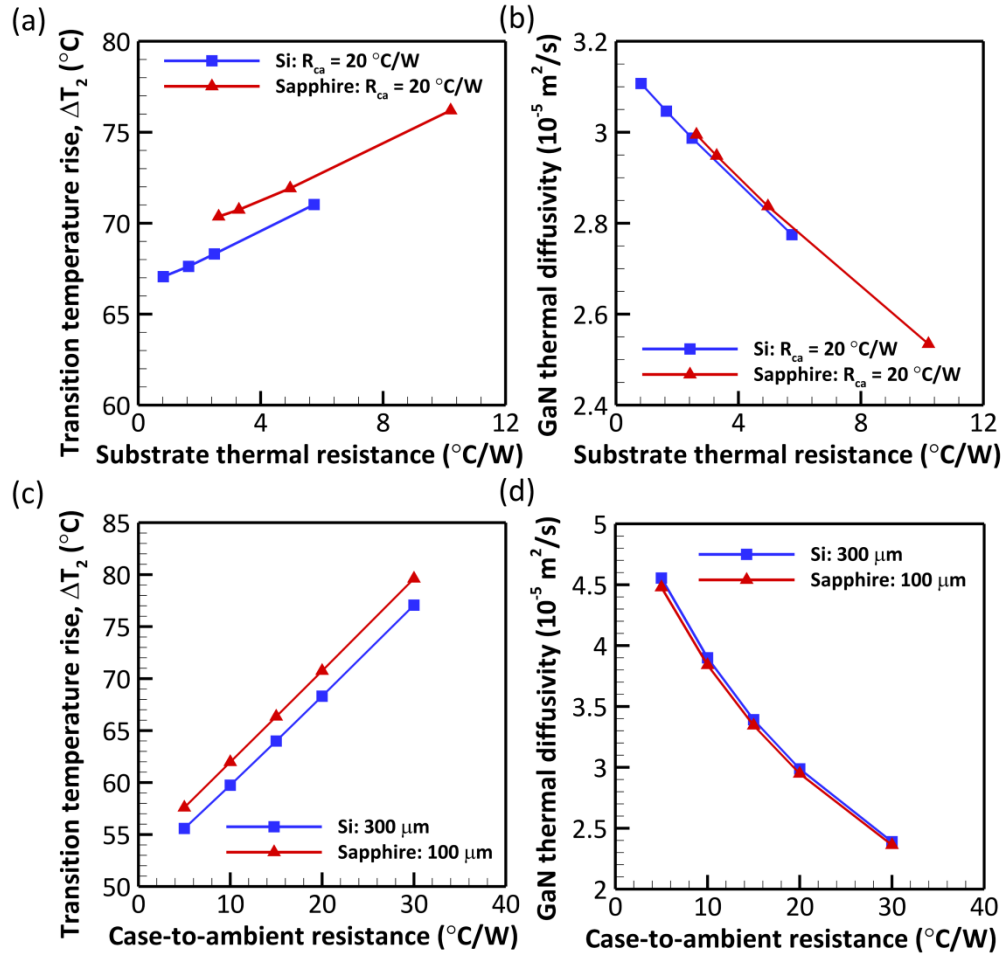
Figure 55: Transition temperature rise dependence on substrate and package thermal resistances
(a) Transition temperature rise and (b) average GaN thermal diffusivity as a function of substrate thermal resistance and (c)-(d) case-to-ambient package resistance for Si (■) and sapphire (▲) substrates

For instance, the transition temperature rise increases from 67.1 °C to 71.0 °C (4%) when the Si substrate thickness is increased from 100 μm to 675 μm ($R_{\theta ca}$= 20 °C/W) while ΔT_2 increases from 55.6 °C to 77.1 °C (39%) when the package resistance increases from 5 °C/W to 30 °C/W for a 300 μm Si substrate. As the GaN thermal diffusivity decreases, the rate at which heat propagates away from the active area also decreases, leading to a higher transition temperature rise. The transition temperature rise was found to be more dependent on the package resistance than on the substrate resistance for the range of substrates investigated. Therefore, it is important to minimize the package resistance in order to minimize both the on-state (steady-state) and transition temperature rises.

These results also confirm that sufficiently thin sapphire substrates may be used to replace Si substrates despite their well-known low thermal conductivity. The comparison for standard 300 μm Si and 100 μm sapphire substrates, which represents the thinnest available, shown in Figure 55(c) and (d) demonstrates their similar thermal performance. The offset of ~2.5 °C in the transition temperature rise curves for the two substrates in Figure 55(c) is primarily due to the

difference in thermal resistance of the GaN-substrate transition and nucleation layers. GaN-on-SiC HEMTs were omitted from the study on package thermal resistance because they are rarely used in power switching applications due to their high cost.

Using a 3D thermal finite element model for a nearly commercial device structure on a 300 μm Si substrate and time-dependent power dissipation profile, the on-state temperature and transition temperature rise were predicted to be 154 °C and 68.3 °C, respectively. Transient temperature rises during the turn-off transition are localized to the GaN layer but depend strongly on the substrate and package thermal resistances because of the temperature dependence of GaN thermal diffusivity. These transition temperature rises should be taken into account in order to design devices for long term performance and reliability.

4. Kirchhoff Transform for Temperature-dependent Thermal Conductivity

The Kirchhoff transform provides a convenient method for solving steady-state and transient heat conduction problems when the thermal conductivity of the medium under consideration varies with temperature. In the steady-state case, the non-linear heat conduction equation

$$\nabla \cdot (k\nabla T) = 0 \qquad (78)$$

in which $k = k(T)$ is difficult to solve in closed form. *Kirchhoff (1894)* [87] introduced an integral transform in terms of a function U

$$U = K\{T\} = \int^{T} k(\tau)d\tau \qquad (79)$$

where the lower bound of the integral can be any value. Kirchhoff demonstrated that

$$\frac{dU}{dT} = k \qquad (80)$$

by the Fundamental Theorem of Calculus and

$$\nabla U = k\nabla T \qquad (81)$$

by the chain rule of differentiation so that the non-linear conduction equation is transformed into the linear Laplace's equation

$$\nabla^2 U = 0. \qquad (82)$$

If the boundary conditions on T (the actual problem to be solved) can be transformed into boundary conditions on U (the linearized problem), one can solve the linear equation for U with standard analytical techniques. The actual temperature can be determined by the inverse Kirchhoff transform

$$T = K^{-1}\{U\} \qquad (83)$$

once the functional form of the thermal conductivity relationship $k = k(T)$ is given. In his lectures on the theory of heat transfer and thermodynamics, Kirchhoff derived this transform for the one-dimensional case only and did not provide an example of how to use the transform [87].

Since its introduction by *Kirchhoff (1894)* [87], the transform has been presented in a number of different forms. Perhaps the most widely cited is that by *Carslaw and Jaeger (1959)* [88]

$$\theta = \frac{1}{k_0} \int_0^{T} k(\tau)d\tau \qquad (84)$$

where θ (later termed the apparent temperature) is related to the Kirchhoff function $U = k_0\theta$ and k_0 is the thermal conductivity of the medium evaluated at $T = 0$. The units of θ are those of temperature and θ is governed by Laplace's equation

$$\nabla^2\theta = 0. \tag{85}$$

Carslaw and Jaeger provide several important details about the transform that were not mentioned in Kirchhoff's introduction. The authors state (without proof) that the transform may be applied with boundary conditions of the first kind (Dirichlet or prescribed temperature) or second kind (Neumann or prescribed heat flux) but not boundary conditions of the third kind (Robin or mixed) that represent convection at the boundary. Carslaw and Jaeger point out that one can solve for the apparent temperature θ using solutions derived for heat conduction in solids with constant thermal conductivity.

Various other references in the literature [89]-[90] define the Kirchhoff transform as

$$\theta = \frac{1}{k_0}\int_{T_0}^{T} k(\tau)d\tau \tag{86}$$

where the constant T_0 "denotes a convenient reference temperature" with $k_0 \equiv k(T_0)$. This definition is essentially the same as that presented by Carslaw and Jaeger in which $T_0 = 0$. *Joyce (1974)* [91] suggested that the apparent temperature be defined as

$$\theta = T_0 + \frac{1}{k_0}\int_{T_0}^{T} k(\tau)d\tau \tag{87}$$

where he defined T_0 as the heat-sink boundary temperature for thermal spreading problems in electronics. Finally, *Batchelor (1991)* [92] defined the apparent temperature to be

$$\theta = T_\infty + \frac{1}{k_\infty}\int_{T_\infty}^{T} k(\tau)d\tau \tag{88}$$

where T_∞ is the ambient temperature and $k_\infty \equiv k(T_\infty)$.

Although all of these definitions lead to the correct solution when applied properly, there may be some confusion when applying the Kirchhoff transform to heat conduction problems. In this paper, we derive the Kirchhoff transform in a general way in order to give a more clear explanation of the transform and its application to a variety of heat conduction problems. We review the transformation of the boundary conditions associated with the Kirchhoff transform and show that the transform can be applied with convection boundary conditions. This is an important aspect of solving thermal spreading problems in electronics chips and packages in closed-form with high accuracy.

4.1 Mathematical Theory

4.1.1 General Form of the Kirchhoff Transform

The most essential feature of the Kirchhoff transform is that the non-linear equation

$$\nabla \cdot (k \nabla T) = 0 \tag{89}$$

is transformed into the linear equation

$$\nabla^2 \theta = 0 \tag{90}$$

in terms of the apparent temperature. This is true when

$$k_0 \nabla \theta = k \nabla T \tag{91}$$

or

$$k_0 \mathrm{d}\theta = k \mathrm{d}T \tag{92}$$

where k_0 is a constant, representing the thermal conductivity of the medium evaluated at a convenient reference temperature T_0. Integration of the above equation from the reference temperature T_0 to θ (or T_0 to T) yields

$$\int_{T_0}^{\theta} k_0 \mathrm{d}\theta' = k_0(\theta - T_0) = \int_{T_0}^{T} k(T')dT'. \tag{93}$$

Rearranging and solving for θ results in the definition given by *Joyce (1974)* [92] as

$$\theta = K\{T\} = T_0 + \frac{1}{k_0} \int_{T_0}^{T} k(\tau)d\tau. \tag{94}$$

Note that when $T = T_0$, the apparent temperature is $\theta = T_0$. As described in detail by *Vernotte (1944)* [93] and referenced by the footnote in *Carslaw and Jaeger (1959)* [88], the apparent temperature θ is a scalar potential that is proportional to the heat flux in the solid. Thus, one can add or subtract a constant from θ without changing the validity of the previous equations.

In thermal spreading problems, we are interested in solving for the temperature rise (or thermal spreading resistance) in a system with a surface flux in the source plane and fixed temperature or convection in the sink plane as shown in Figure 56.

Prescribed heat flux

Single- or multi-layered chip

Fixed temperature or convection at the sink
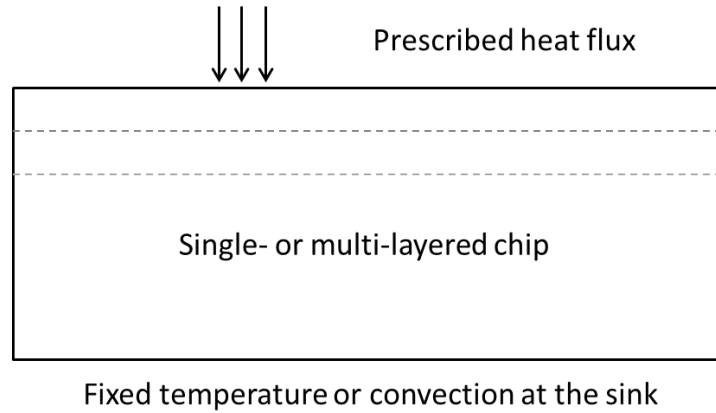
Figure 56: Schematic of layout for common thermal spreading problem in electronics

Many solutions for temperature rise and thermal spreading resistance are formulated in terms of the temperature rise above the sink plane (cf. [71] and references contained therein). Therefore, it is usually convenient to define T_0 as the prescribed temperature at the sink boundary when a first kind BC is used at the sink. In this case, the apparent temperature also has the value T_0 in the sink plane. It will be shown later that in cases with a convective BC in the sink plane, the average temperature of the sink plane is the most appropriate choice for T_0.

4.1.2 Transformation of the Boundary Conditions

Ozisik (1969) [94] provides the most in-depth explanation of how the boundary conditions of the original problem in terms of T can be transformed into those in terms of θ for boundary conditions of the first and second kind. In the following section, the derivation in this reference is followed for the transformation of the boundary conditions.

First Kind Boundary Conditions

Boundary conditions of the first kind in terms of T take the form

$$T = f(\vec{r}) \tag{95}$$

where $\vec{r}$ denotes the position vector, the boundary condition on T transforms into the boundary condition on θ specified by

$$\theta = f^*(\vec{r}). \tag{96}$$

The form of the transformed boundary condition, f^*, depends on the functional form of the thermal conductivity relationship $k(T)$. For example, in the case of the simple linear thermal conductivity relationship $k(T) = a + bT$, the apparent temperature is given by

$$\theta = T_0 + \frac{1}{k_0} \int_{T_0}^{T} k(\tau) d\tau$$

$$= T_0 + \frac{1}{k_0} \int_{T_0}^{T} (a + b\tau) d\tau \qquad (97)$$

$$= T_0 + \frac{1}{k_0} \left(aT + \frac{1}{2} bT^2 \right).$$

Therefore, the transformed boundary condition is given by

$$\theta = f^*(\vec{r}) = T_0 + \frac{1}{k_0} \left(af(\vec{r}) + \frac{1}{2} b[f(\vec{r})]^2 \right). \qquad (98)$$

One should note that it is desirable to solve the linearized problem for θ using standard analytical techniques. Thus, the boundary condition on T should not be too complicated as to make the transformed boundary condition θ even more complicated and prohibitively challenging to solve with standard analytical techniques.

Second Kind Boundary Conditions

For boundary conditions of the second kind (prescribed heat flux),

$$\vec{q} = -\vec{n} \cdot k\nabla T = g(\vec{r}) \qquad (99)$$

where $\vec{n}$ is the outward unit normal vector, transformation of the boundary condition on T is more straightforward to boundary conditions on θ. Since

$$k_0 \nabla \theta = k\nabla T \qquad (100)$$

then

$$-\vec{n} \cdot k_0 \nabla \theta = -\vec{n} \cdot k\nabla T \qquad (101)$$

and so the boundary condition on θ is equivalent to the boundary condition on T and is given by

$$\vec{q} = -\vec{n} \cdot k_0 \nabla \theta = g(\vec{r}). \qquad (102)$$

Transformation of boundary conditions of the second kind does not add any additional difficulty to the problem.

Third Kind Boundary Conditions

For boundary conditions of the third kind, which usually represent convection to a fluid, the boundary condition on T is given by

$$-\vec{n} \cdot k\nabla T = h(T - T_\infty) \qquad (103)$$

where h is the heat transfer coefficient and T_∞ is the ambient temperature. The left hand side of the equation above is equal to $-\vec{n} \cdot k_0 \nabla\theta$ but the right hand side does not transform as desired into a condition on θ. This leads to the transformed boundary condition

$$-\vec{n} \cdot k_0 \nabla\theta = h(K^{-1}\{\theta\} - T_\infty) \neq h(\theta - T_\infty) \tag{104}$$

which is not a linear boundary condition on θ due to the non-linearity in the inverse Kirchhoff transform $T = K^{-1}\{\theta\}$.

4.2 Application for Different Thermal Conductivity Relationships

In general, the thermal conductivity of a solid can vary with temperature in a number of ways depending upon the material and the temperature regime of interest.

1. Linear: It is common for the thermal conductivity of metals above ambient temperature to increase approximately linearly, *i.e.,*

$$k(T) = a + bT \tag{105}$$

where a and b are constants determined empirically and the temperature T is often ambiguously defined. For copper (Cu), a = 421.429 W/m-K and b = -0.069 W/m-K^2 when T is defined as the absolute temperature [95]. At lower temperatures, the thermal conductivity of Cu depends non-linearly on the temperature due to the dominance of different phonon scattering mechanisms. In the older heat transfer literature, it is common to define the temperature used in the thermal conductivity relationships of metals in degrees Celsius rather than kelvin. For consistency, it is recommended to perform all calculations in absolute temperature units and convert to the desired units at the end of the calculation. The Kirchhoff transform and its inverse are thus given by

$$\theta = T_0 + \frac{1}{k_0} \int_{T_0}^{T} (a + b\tau)d\tau = \frac{b}{2k_0}\left(T^2 - T_0{}^2\right) + \frac{a}{k_0}(T - T_0) + T_0 \tag{106}$$

$$T = \frac{-a + \sqrt{a^2 + 2bk_0\theta - 2bk_0T_0 + 2abT_0 + b^2T_0{}^2}}{b} \tag{107}$$

Often, the thermal conductivity decreases with increasing temperature ($b < 0$) so that the temperature rise is higher when variable thermal conductivity is taken into account than for constant thermal conductivity values.

2. Inversely proportional: The thermal conductivity of many semiconductors shows an approximately inverse dependence on absolute temperature around ambient temperature (300 K) and above where phonon-phonon scattering dominates. In this case, the thermal conductivity relationship is expressed as

$$k(T) = k_{ref} \left(\frac{T_{ref}}{T} \right) \tag{108}$$

where T_{ref} is a reference temperature (typically 300 K) and k_{ref} is the thermal conductivity of the material evaluated at T_{ref}. In practice, the thermal conductivity of semiconductor materials varies as $\sim T^{-n}$ with $1 < n < 1.5$. Because the antiderivative of T^{-n} is very different from that of T^{-1}, the former case is treated in the following section. Sapphire, however, has been shown to follow the $\sim T^{-1}$ relationship from 3ω thermal conductivity measurements in the temperature range 300 K to 600 K [96]. The Kirchhoff transform and its inverse are given by the following formulas in this case

$$\theta = T_0 + \frac{1}{k_0} \int_{T_0}^{T} k_{ref} \left(\frac{T_{ref}}{\tau} \right) d\tau = T_0 + \frac{k_{ref}}{k_0} T_{ref} \ln \frac{T}{T_0} \tag{109}$$

$$T = T_0 \exp \left[\frac{k_0 (\theta - T_0)}{k_{ref} T_{ref}} \right] \tag{110}$$

where k_{ref} and T_{ref} have been explicitly distinguished from k_0 and T_0 for generality. Although many reports and the literature choose and ambient temperature as the reference temperature for the transform which is equal to the reference temperature for the thermal conductivity relation, this may not always be the case. The expressions for the apparent and actual temperature can be further simplified from the relation $k_0 = k_{ref} T_{ref} / T_0$ to

$$\theta = T_0 \left(1 + \ln \frac{T}{T_0} \right) \tag{111}$$

$$T = T_0 \exp \left[\frac{\theta}{T_0} - 1 \right]. \tag{112}$$

The reference temperature for the thermal conductivity relationship T_{ref} has canceled out of the formulas above, indicating that the choice of this reference temperature is not significant.

3. Power dependence: As noted above, the thermal conductivity of a semiconductor crystalline material is often given by

$$k(T) = k_{ref} \left(\frac{T_{ref}}{T} \right)^n \tag{113}$$

where $1 < n < 1.5$ and is determined by experiment [63]. The Kirchhoff transform and inverse transform are given by

$$\theta = T_0 + \frac{1}{k_0} \int_{T_0}^{T} k_{ref} \left(\frac{T_{ref}}{\tau}\right)^n d\tau$$

$$= T_0 + \frac{k_{ref}}{k_0} \frac{T_{ref}{}^n\left(T^{1-n} - T_0{}^{1-n}\right)}{1-n} \qquad (114)$$

$$= T_0 + \frac{T_0{}^n\left(T^{1-n} - T_0{}^{1-n}\right)}{1-n}$$

$$T = \left[T_0{}^{1-n} + \frac{k_0}{k_{ref}} \frac{(\theta - T_0)(1-n)}{T_{ref}{}^n}\right]^{\frac{1}{1-n}}$$

$$= \left[T_0{}^{1-n} + \frac{(\theta - T_0)(1-n)}{T_0{}^n}\right]^{\frac{1}{1-n}} \qquad (115)$$

$$= T_0 \left[1 + \frac{(\theta - T_0)(1-n)}{T_0}\right]^{\frac{1}{1-n}}$$

where the dependence on the values of T_{ref} and k_{ref} has also been eliminated through the algebraic simplifications [97]. Thus, only a consistent choice of the reference temperature T_0 for the Kirchhoff transform is required.

4.3 Application to Multi-finger GaN HEMTs

The Kirchhoff transform provides a very computationally efficient way of calculating the temperature rise in multi-finger GaN HEMTs, particularly when analytical thermal models are available that closely match the physical situation. In this demonstration, the influence of temperature-dependent thermal conductivity on the temperature rise is demonstrated for the multi-finger HEMT layout shown in the previous chapter of this thesis on analytical thermal modeling. The structure consists of a compound substrate (GaN and SiC) with 10 heat sources for the half-model (representing the full structure with 20 heat sources) as shown in Figure 57.
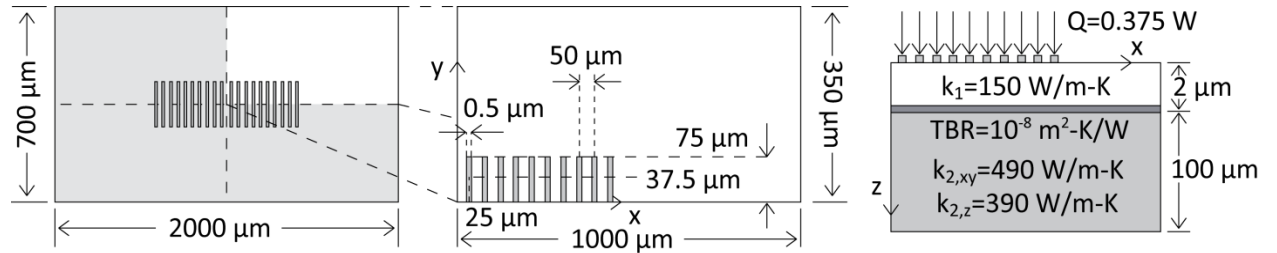
Figure 57: Layout of multi-finger GaN HEMT model for Kirchhoff transform application

For simplicity at first, the expressions for the thermal conductivity of GaN and SiC are assigned the same temperature-dependent functional form ($\sim T^{-1.3}$) because heterogeneous temperature-dependent relationships in multi-layered structures introduce additional complications [97]-[98]. Similarly, the base of the SiC substrate is set to a fixed temperature of 20 °C in order to avoid the

difficulty associated with a finite base conductance boundary condition, which will be treated in detail in the following section. Thus, the thermal conductivity relationships of the layers are

$$k_{GaN}(T) = 150 \left(\frac{300}{T}\right)^{1.3} \tag{116}$$

$$k_{SiC}(T) = 420 \left(\frac{300}{T}\right)^{1.3} \tag{117}$$

where an isotropic value of SiC thermal conductivity (420 W/m-K) has been used and gives a very similar result as when orthotropic values are used. In the past, the effective GaN-SiC thermal interfacial resistance associated with the AlN nucleation layer and true interfacial resistances between the various layers was measured to have a much stronger temperature dependence ($\sim T^{-2.76}$) than the thermal conductivity of GaN and SiC [32]. However, for the purposes of demonstration, the interfacial conductance (inverse of resistance) is assigned the same temperature-dependent functional form as the thermal conductivity of GaN and SiC

$$h_c(T) = 10^8 \left(\frac{300}{T}\right)^{1.3} \tag{118}$$

where h_c is the interfacial conductance in W/m^2-K.

Non-linear 3D FEA thermal simulations were performed in COMSOL Multiphysics 4.3a to validate the accuracy of the analytical thermal model based on infinite series techniques and the application of the Kirchhoff transformation. The 3D surface temperature plot and the 1D temperature slice through the center of the model at the surface is shown in Figure 58. The data points computed by the analytical model and Kirchhoff transform are indicated at dots while the FEA solution is shown by a solid line.
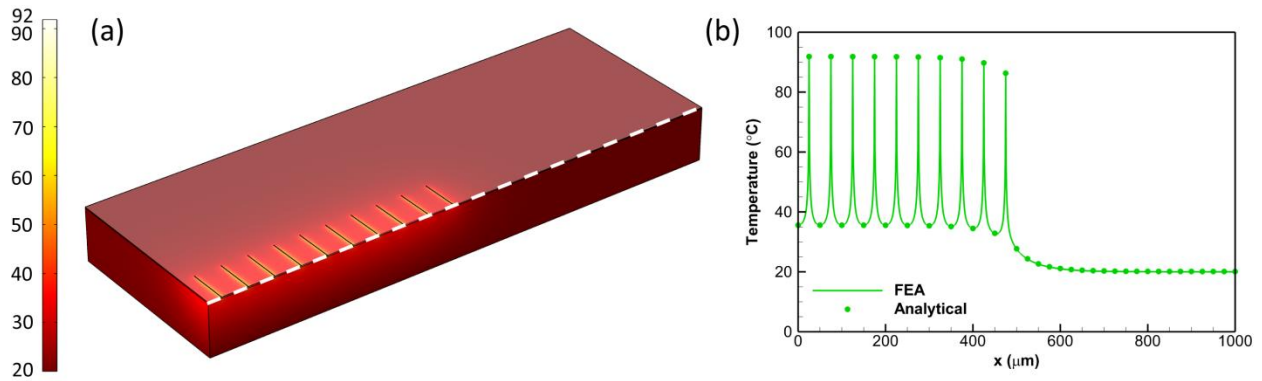
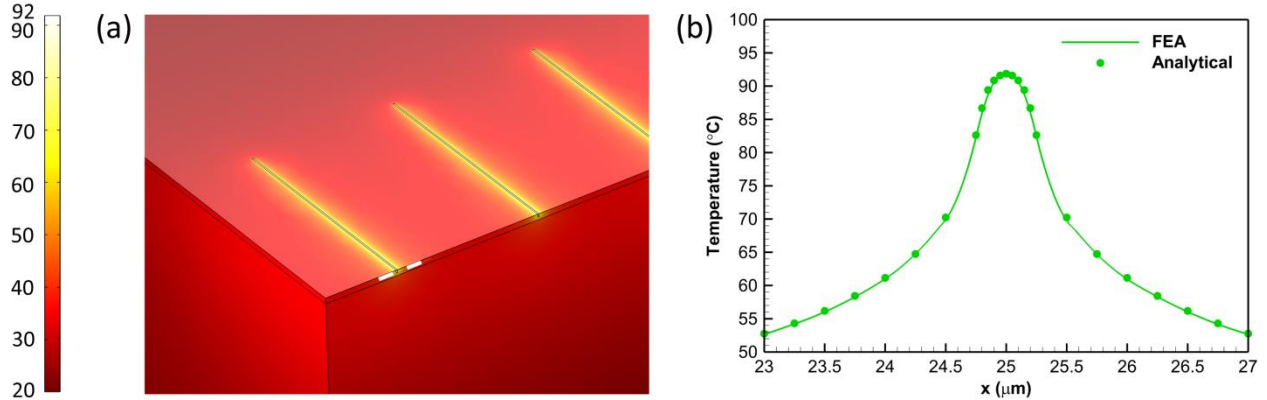Figure 58: FEA and analytical temperature distribution (°C) for multi-finger HEMT with temperature-dependent thermal conductivity
(a) 3D FEA model temperature plot and (b) 1D temperature slice with FEA/analytical comparison

Overall, excellent agreement is obtained between the FEA and Kirchhoff transform methods because the Kirchhoff transform and infinite series solution represent the exact analytical

solution to the non-linear heat conduction equation. The temperature at the center of the heat sources agrees to within 0.8% between the non-linear FEA and Kirchhoff transformed analytical solutions. This includes the region around the 0.5 µm long heat source in which the temperature changes sharply over a few microns as seen in Figure 59.

Figure 59: FEA and analytical temperature distribution (°C) for multi-finger HEMT in the region around the innermost heat source
(a) 3D FEA model temperature plot and (b) 1D temperature slice with FEA/analytical comparison

The device temperature field is strongly affected by the fixed base temperature boundary condition, causing the temperature to rise significantly in the area around the heat source but remain near the base temperature elsewhere. It is also interesting to note that the temperature-dependence of the thermal conductivity has a relatively small impact on the peak temperature rise because of the influence of the base boundary condition and high thermal conductivity 100 µm SiC substrate (increases by ~2 °C from the constant thermal conductivity case). A comparison of the source centroid temperatures calculated from the FEA and analytical models are shown in the table below.

Table 5: Heat source centroid temperatures for multi-finger HEMT with variable thermal conductivity and uniform base temperature

Source no.	FEA Temperature (°C)	Analytical Temperature (°C)
1	91.7615	91.8585
2	91.7095	91.8561
3	91.7440	91.8496
4	91.7262	91.8341
5	91.6927	91.7977
6	91.6215	91.7120
7	91.3744	91.5078
8	90.9209	91.0114
9	89.6913	89.7607
10	86.2148	86.3111

The differences between the FEA and analytical solution are likely due to numerical errors in the FEA solver and/or truncation error in the MATLAB implementation of the infinite series

105

solution. The temperature differences of ~0.1 °C or ~1% of the temperature rise above the heat sink base serves as a comparison metric for the non-linear FEA and Kirchhoff transformed analytical solutions for a finite base conductance.

4.4 Application to Problems with Third Kind Boundary Conditions

Although the Kirchhoff transform cannot rigorously be applied to heat conduction problems with boundary conditions of the third kind, we have observed that in some particular problems the Kirchhoff transform gives a very accurate result. In typical thermal spreading problems in electronics with one second kind (source plane) and one third kind boundary condition (sink plane), the actual temperature T and the apparent temperature θ are approximately equal in the sink plane. This is a condition in which

$$-\vec{n} \cdot k_0 \nabla \theta = h(T - T_\infty) \approx h(\theta - T_\infty) \tag{119}$$

at the sink boundary. When this relation holds, multi-dimensional conduction solutions for a medium with constant thermal conductivity can be applied to solve for the apparent temperature θ and the inverse Kirchhoff transform can be calculated to find the actual temperature T.

4.4.1 One-dimensional Example

As an introductory example, one-dimensional conduction through a plane wall of length L with a heat flux boundary condition one side ($x = 0$) and a convection boundary condition on the other side ($x = L$) is considered. The schematic of the model layout is shown in Figure 60.
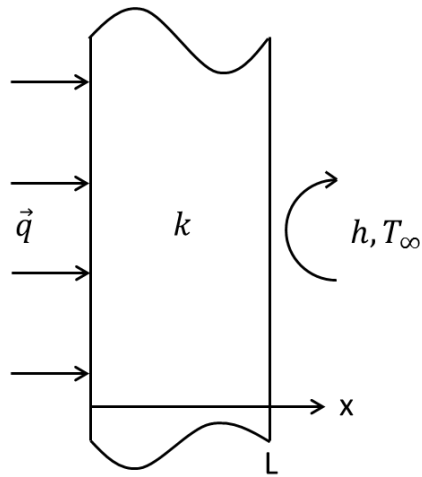

Figure 60: One-dimensional plane wall example for Kirchhoff transformation

The temperature on the right side of the boundary at $x = L$ can be calculated *a priori* because the total heat flux through the plane wall is known, *i.e.,*

$$T(x = L) = \frac{q}{h} + T_\infty. \tag{120}$$

Then, the convective boundary condition on T can be thought of as a first kind boundary condition on T such that $T(x = L) = T_0$. Then, using the previously derived definition of the apparent temperature θ, it follows that $\theta(x = L) = T_0$. The solution to the ordinary linear differential equation

$$\frac{d^2\theta}{dx^2} = 0 \tag{121}$$

is given by

$$\theta(x) = \frac{q}{k_0}(L - x) + T_0 \tag{122}$$

or

$$\theta(x) = q\left(\frac{L - x}{k_0} + \frac{1}{h}\right) + T_\infty \tag{123}$$

where $k_0 = k(T_0)$ for $T_0 = \frac{q}{h} + T_\infty$. This is in fact the exact solution for the apparent temperature θ as if the convective boundary condition were applied to θ instead of T. For the values of $L = 10^{-4}$ m, $q = 3$ x 10^4 W/m^2, $h = 1000$ W/m^2-K, $T_\infty = 350$ K, and $k(T) = a + bT$ with $a = 0.25$ W/m-K and $b = -5$ x 10^{-4} W/m-K^2,

$$T_0 = \frac{3 \times 10^4}{10^3} + 350 = 380 \, K \tag{124}$$

$$k_0 = 0.25 + (-5 \times 10^{-4})(380) = 0.06 \, \frac{W}{mK}. \tag{125}$$

the values of the end points of the apparent temperature are $\theta(x = 0) = 380 \, K$ and $\theta(x = L) = 430 \, K$. When the inverse Kirchhoff transform is applied, the actual temperature of the end points are $T(x = 0) = 380$ K and $T(x = L) = 451.0102$ K. When compared to a non-linear finite element solution computed in COMSOL Multiphysics, the temperature at the right boundary $T(x = L) = 450.9878$ K is in excellent agreement with the Kirchhoff transformed analytical solution.

4.4.2 Two-dimensional Example

The situation in multi-dimensional conduction with convection at the sink is more complicated that the one-dimensional case because convection at the boundary effects the temperature in the medium non-uniformly. In one-dimensional conduction problems, the convection at the boundary raises the boundary temperature uniformly. However, in multi-dimensional conduction problems, the sink temperature becomes non-uniform when a convection boundary condition is used instead of a fixed temperature boundary condition. Thus, one would expect that the Kirchhoff transform cannot always be applied with convection at the sink because the convection

boundary condition cannot be treated as a first kind boundary condition with the sink temperature changed by an appropriate amount.

For multi-dimensional conduction problems in which the substrate is relatively thick (to be quantified in more detail later), the Kirchhoff transform may be applied with convection at the base using the procedure outlined below:

1. Compute the average temperature at the base where convection occurs and set this temperature equal to the Kirchhoff transform reference temperature (T_0).
2. Find the apparent temperature by solving the linearized heat equation using thermal conductivity values at T_0 and a heat transfer coefficient h_s equal to the heat transfer coefficient in the non-linear problem.
3. Apply the inverse Kirchhoff transform to find the actual temperature, which accounts for temperature-dependent thermal conductivity.
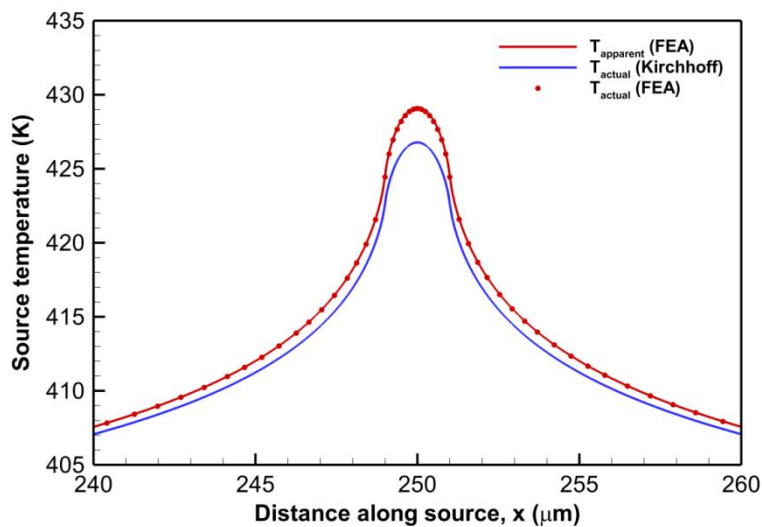
As a first example, the temperature distribution in a 2D isotropic, rectangular substrate of 500 μm length and 100 μm thickness with the thermal conductivity of Si and a centrally-located heat source is calculated using the proposed method. A heat flux of 10^9 W/m² was applied to the 2 μm long heat source. The heat transfer coefficient at the base of the substrate was $h_s = 10^5$ W/m²-K and the ambient temperature was $T_\infty = 350$ K. The thermal conductivity of Si was described by the relation

$$k_{Si}(T) = 150 \left(\frac{300}{T} \right)^{1.3} \tag{126}$$

where the thermal conductivity is given in W/m-K [32]. In this case, the average temperature of the sink was calculated to be $T_0 = 390$ K from the prescribed heat flux at the source, the heat transfer coefficient, the ambient temperature, and the geometry of the substrate. The thermal conductivity at the reference temperature of 390 K is therefore $k_0 = 106.6510$ W/m-K. An FEA model computed in COMSOL Multiphysics 4.2a (2D free triangular mesh with 132,296 elements) was used to calculate the linearized (apparent) and actual (variable thermal conductivity) temperature distributions. The 1D temperature slice along the source plane and in the region close to the heat source is shown in Figures 61 and 62.
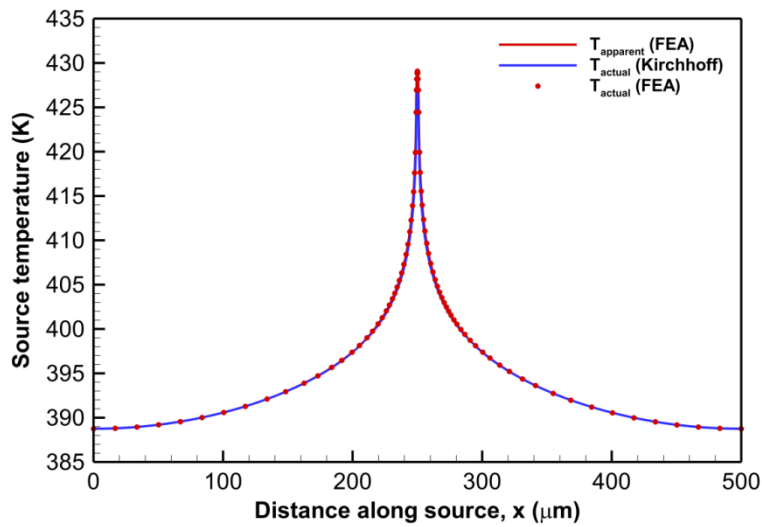
Figure 61: 1D source plane temperature distribution for 2D Kirchhoff example with third kind BC

Figure 62: 1D source plane temperature distribution for 2D Kirchhoff example with third kind BC in the region around the heat source

The very good agreement between the non-linear FEA and Kirchhoff transformed solution in the source plane is related to the close match between the temperature predicted by the two methods in the sink plane. The temperature distribution along the sink plane is shown in Figure 63 to demonstrate this close agreement.
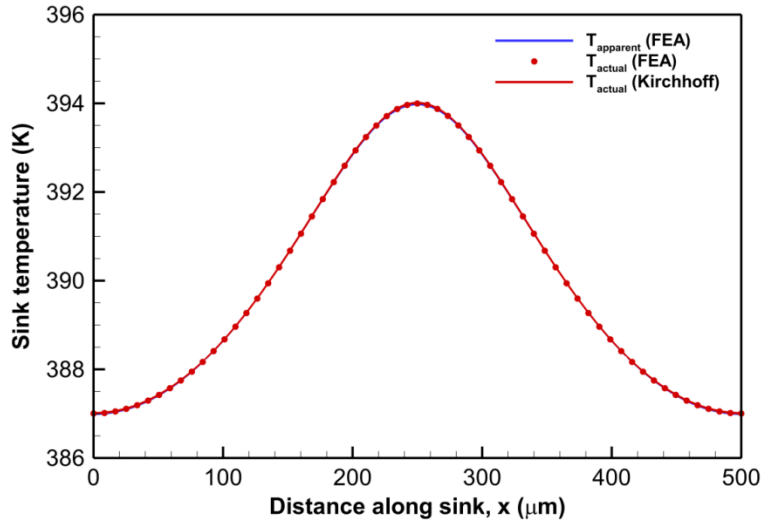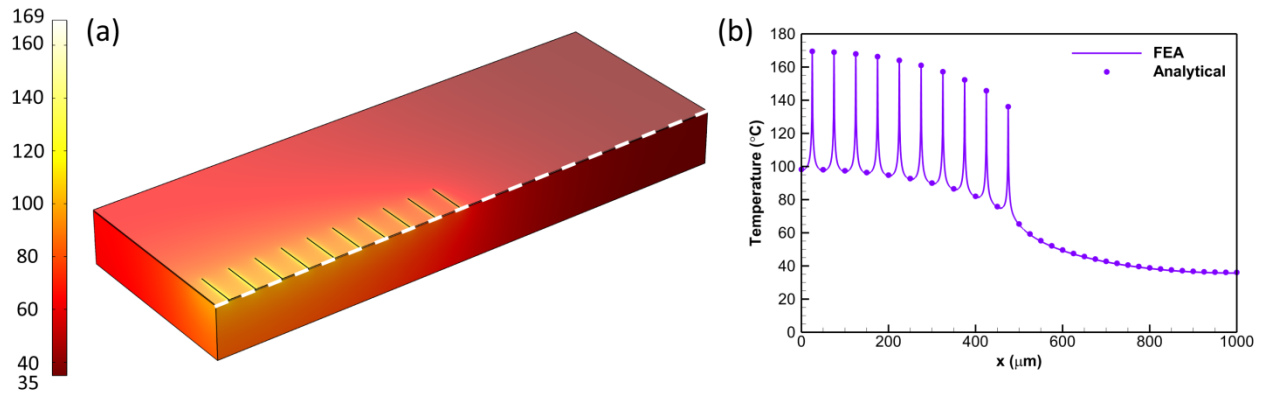
Figure 63: 1D sink plane temperature distribution for 2D Kirchhoff example with third kind BC

In this example, the Kirchhoff transformed temperature is very close to the non-linear FEA solution. The peak temperature of the source from the non-linear FEA is 429.0708 K while the peak temperature from the Kirchhoff transform of the linear FEA solution is 429.1524 K. The maximum difference between the non-linear and Kirchhoff transformed temperatures in the source plane is 0.081605 K and the average difference between these two temperatures is 0.017949 K. I think the reason for this very good agreement is that the sink temperature is fairly uniform and that the apparent sink temperature with the thermal conductivity evaluated at k_0 is very close to the actual sink temperature with variable thermal conductivity.

4.4.3 Three-dimensional Example

As the final example of the Kirchhoff transform and the problem of interest in this thesis, the steady-state temperature rise in a multi-finger GaN HEMT with convection at the base. It was demonstrated in the previous chapter on thermal modeling that the substrate base thermal boundary condition plays a crucial role in shaping the device temperature distribution. Thus, capturing the finite base conductance is critical in accurately evaluating the near-junction temperature rise of GaN HEMTs in RF and high frequency applications. As before, a heat transfer coefficient of $h_s = 3.2748 \times 10^5$ W/m^2-K and an ambient temperature of $T_\infty = 20$ °C was used to represent spreading in a SnAg die attach layer and sinking to a micro-channel cooler [15]. The technique for applying the Kirchhoff transform with convection BCs described here was applied with a reference temperature of $T_0 = 325.86737$ K. All other parameters of the model were the same as in the previous Kirchhoff transform demonstration for a multi-finger GaN HEMT. The 3D surface temperature plot from the non-linear FEA model is shown in Figure 64 along with a 1D source plane temperature comparison in the center of the device.

110

**Figure 64: Temperature distribution for 3D multi-finger GaN HEMT example with finite base conductance
(a) Non-linear FEA surface temperature and (b) FEA/analytical comparison**

As in the previous cases, very good agreement to within ~1% between the non-linear FEA and Kirchhoff transformed analytical solution was obtained for the centroidal heat source temperatures. Unlike the case with the uniform base temperature, the use of a finite heat transfer coefficient at the substrate base causes a significant temperature rise throughout the SiC substrate. In addition, the temperature varies between adjacent heat sources. A closer view of the innermost heat source is shown in Figure 65, also demonstrating the excellent agreement between the Kirchhoff transformed solution and the non-linear FEA model.
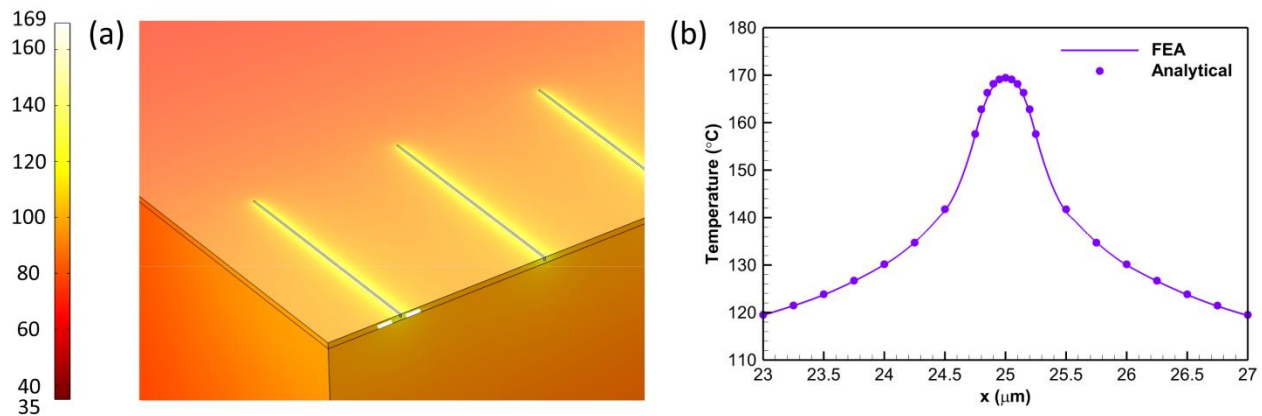
**Figure 65: Close-up view of temperature distribution for 3D multi-finger GaN HEMT example with finite base conductance
(a) Non-linear FEA surface temperature and (b) FEA/analytical comparison**

These examples demonstrate the suitability of solving non-linear heat conduction problems in GaN-based electronics with the Kirchhoff transform while accounting for heat transfer across a finite conductance at the base. Analytical models based on infinite series techniques coupled with the Kirchhoff transform are much faster for evaluating the peak temperature rise of the source centroids.

5. Conclusions and Future Work

The field of GaN-based electronics is continually expanding and the need for cost-effective, high performance thermal management solutions will continue to increase. It was demonstrated in 2004 that 40 W/mm RF output power is possible with single finger GaN HEMTs [1] but commercial GaN devices have been limited to relatively low values of 2 to 4 W/mm by a combination of thermal, circuit design, and other issues. GaN-based electronics have been specifically targeted by the DARPA Microsystems Technology Office's Near-junction Thermal Transport (NJTT) call in 2011 [99] and the ICECool Fundamentals call in 2012 [43]. Thermal management of GaN-based electronics will continue to be at the forefront of the consciousness of the heat transfer community over the next decade as it represents one of the most challenging thermal management application areas.

The work presented in this thesis is a helpful starting point in understanding the source of high device temperatures in GaN-based electronics and some basic metrics that thermal management strategies must meet to be effective on a device-level. Electro-thermal modeling was performed and the results explained in the context of thermal spreading resistance, which is the key concept to understanding the importance of the heat source distribution. Both numerical and analytical thermal models were used to evaluate the device temperature distribution and to explain the dependence of device temperature on key parameters. Finally, an extension to the Kirchhoff transform was developed to account for non-uniform substrate base temperatures due to a finite heat transfer coefficient associated with the resistance of the package. Much of the modeling work presented in this thesis requires further experimental validation with high spatial resolution temperature measurement techniques.

Many aspects of thermal issues and thermal management in GaN-based electronics need further investigation; the initial research into the topics presented herein raise many more questions that have not been discussed adequately in the literature. Some of these issues include:

- Thermal conductivity of thin (~10 to 100 nm) epitaxial layers of AlN, GaN, and $Al_xGa_{1-x}N$ alloys due to classical size effects
- Accurate thermal conductivity measurements for AlxGa1-xN based on the presence of alloy scattering effects
- The necessity of non-continuum thermal modeling of GaN HEMTs with the Boltzmann Transport Equation (BTE) or other methods versus the classical (Fourier's law) heat conduction equation
- Optimization of GaN-on-diamond structures and transfer processes
- Materials for low cost RF packages
- Low cost thermal management solutions for GaN-based power electronics
- Potential for top-side thermal management with high thermal conductivity materials without compromising electrical performance and reliability
- Very high (~0.1 μm) spatial resolution temperature measurements
- Role of field plates and other device structures and electron trapping effects on the heat source distribution in GaN HEMTs under DC bias

In the future, it will be important to pursue a variety of thermal management strategies that address thermal issues in different application areas of GaN-based electronics, such as continuous wave RF PAs, pulse mode RF PAs, and power-switching electronics for voltage conversion and control. The strategies ultimately chosen for a particular application area will be dependent upon the device layout, packaging scheme, cost, volume, and other factors. Some possible device-level thermal management strategies are shown in Figure 66.
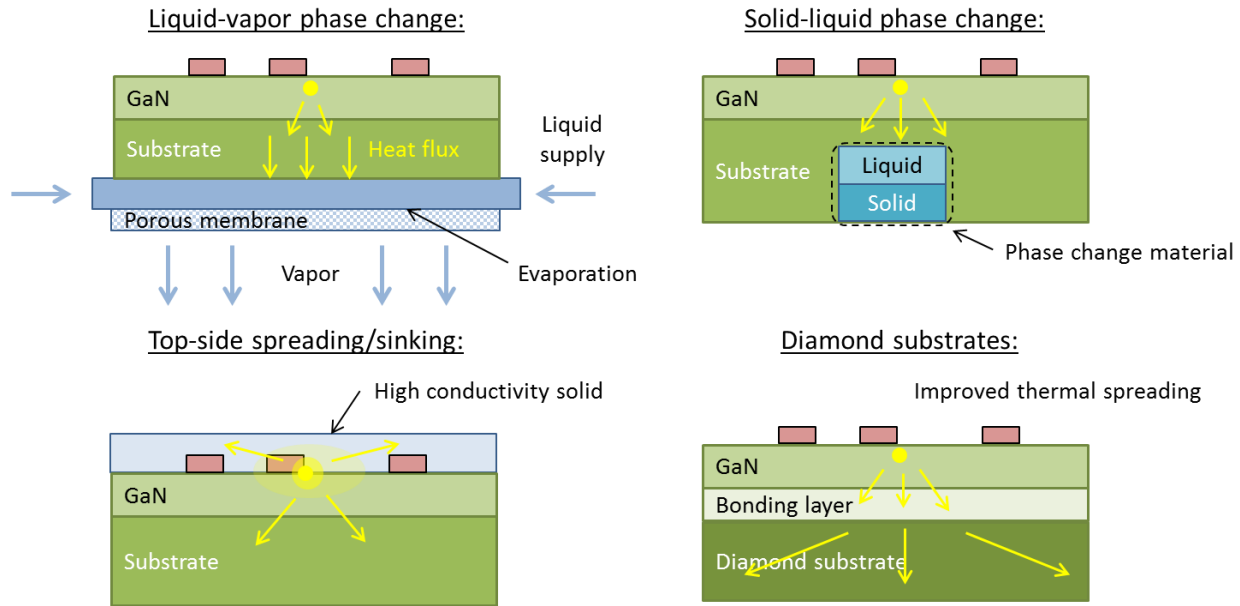
Figure 66: Possible future thermal management strategies for GaN-based electronics

The wide variety of challenges that has emerged in thermal management of GaN-based electronics motivates new creativity in developing novel solutions. The success of future thermal management strategies will certainly require interdisciplinary collaboration and new fabrication technologies to meet the present and future need. This is truly one of the most exciting times to work in the development of microelectronics and high performance thermal management technologies.

6. Bibliography

[1] Y.-F. Wu, M. Moore, A. Saxler, T. Wisleder, and P. Parikh, "40-W/mm double field-plated GaN HEMTs," *Proc. IEEE Device Research Conference*, 2006, pp. 151-152.

[2] M. Yanagihara, Y. Uemoto, T. Ueda, T. Tanaka, and D. Ueda, "Recent advances in GaN transistors for future emerging applications," *Physica Status Solidi A*, vol. 206, no. 6, Feb. 2009, pp. 1221-1227.

[3] D. I. Babic, Q. Diduck, P. Yenigalla *et al.,* "GaN-on-diamond field-effect transistors: from wafers to amplifier modules," *Proc. 33rd Int. Convention MIPRO,* Opatija, Croatia, 2010, pp. 60-66.

[4] R. Chu, A. Corrion, M. Chen *et al.,* "1200-V normally off GaN-on-Si field-effect transistors with low dynamic on-resistance," *IEEE Electron Device Lett.,* vol. 32., no. 5, May 2011.

[5] S. M. Sze and K. K. Ng, *Physics of Semiconductor Devices*, Hoboken, NJ: Wiley, 2007, p. 747.

[6] S. R. Bahl and J. A. del Alamo, "A new drain-current injection technique for the measurement of off-state breakdown voltage in FET's," *IEEE Trans. Electron Devices*, vol. 40, no. 8, Aug. 1993, pp. 1558-1560.

[7] J. D. Albrecht, T.-H. Chang, A. S. Kane, and M. J. Rosker, "DARPA's nitride electronic NeXt generation technology program," *Proc. IEEE CSICS*, Monterey, CA, Oct. 2010.

[8] A. D. Kraus and A. Bar-Cohen, *Thermal Analysis and Control of Electronic Equipment*, Washington, D.C.: Hemisphere, 1983.

[9] M. Rosker, C. Bozada, H. Dietrich *et al.,* "The DARPA wide band gap semiconductors for RF applications (WBGS-RF) program: phase II results," *Proc. CS MANTECH Conference*, Tampa, FL, May 2009.

[10] J. Joh, J. A. del Alamo, U. Chowdhury *et al.,* "Measurement of channel temperature in GaN high-electron mobility transistors," *IEEE Trans. Electron Devices*, vol. 56, no. 12, Dec. 2009, pp. 2895-2901.

[11] Silvaco, Inc., "Silvaco ATLAS Device Simulation Framework." [Online]. Available: http://www.silvaco.com/products/device_simulation/atlas.html.

[12] Synopsys, Inc., "Sentaurus Device: An advanced multidimensional (1D/2D/3D) device simulator." [Online]. Available: http://www.synopsys.com/Tools/TCAD/Device Simulation/Pages/SentaurusDevice.aspx.

[13] "Minimos-NT." [Online]. Available: http:// http://www.iue.tuwien.ac.at/index.php? id=205.

[14] K. J. Negus and M. M. Yovanovich, "Thermal computations in a semiconductor die using surface elements and infinite images," *Proc. Int. Symp. Cooling Technol. for Electronic Equipment*, 1987, pp. 474-485.

[15] M. Garven and J. P. Calame, "Simulation and optimization of gate temperatures in GaN-on-SiC monolithic microwave integrated circuits," *IEEE Trans. Components and Packaging Technol.*, vol. 32, no. 1, Mar. 2009, pp. 63-72.

[16] J. P. Calame, R. E. Myers, S. C. Binari *et al.,* "Experimental investigation of microchannel coolers for the high heat flux thermal management of GaN-on-SiC semiconductor devices," *Int. J. Heat and Mass Transfer*, vol. 50, Apr. 2007, pp. 4767-4779.

[17] F. Bertoluzza, N. Delmonte, R. Menozzi, "Three-dimensional finite-element thermal simulation of GaN-based HEMTs," *Microelectronics Reliability*, vol. 49, Mar. 2009, pp. 468-473.

[18]   E. A. Douglas, F. Ren, S. J. Pearton, "Finite-element simulations on the effect of device design on channel temperature for AlGaN/GaN high electron mobility transistors," *J. Vac. Sci. Technol. B*, vol. 29, no. 2, Mar. 2011.

[19]   H. C. Nochetto, N. R. Jankowski, A. Bar-Cohen, "The impact of GaN/substrate thermal boundary resistance on a HEMT device," *Proc. ASME 2011 IMECE*, Denver, CO, USA, Nov. 2011, pp. 1-9.

[20]   U. Lindefelt, "Heat generation in semiconductor devices," *J. Appl. Phys*, vol. 75, no. 2, Jan. 1994, pp. 942-957.

[21]   V. O. Turin, and A. A. Balandin, "Electrothermal simulation of the self-heating effects in GaN-based field-effect transistors," *J. Appl. Phys.*, vol. 100, Sep. 2006.

[22]   E. R. Heller and A. Crespo, "Electro-thermal modeling of multifinger AlGaN/GaN HEMT device operation including thermal substrate effects," *Microelectronics Reliability*, vol. 48, Apr. 2007, pp. 45-50.

[23]   A. M. Darwish, A. J. Bayba, H. A. Hung, "Thermal resistance calculation of AlGaN-GaN Devices," *IEEE Trans. Microwave Theory and Techniques*, vol. 52, no. 11, Nov. 2004, pp. 2611-2620.

[24]   B. Benbakhti, A. Soltani, K. Kalna, M. Rousseau, and J.-C. De Jaeger, "Effects of self-heating on performance degradation in AlGaN/GaN-based devices," *IEEE Trans. Electron Devices*, vol. 56, no. 10, Oct. 2009, pp. 2178-2185.

[25]   S. Vitanov, V. Palankovski, S. Maroldt, and R. Quay, "High-temperature modeling of AlGaN/GaN HEMTs," *Solid-state Electronics*, vol. 54, Jun. 2010, pp. 1105-1112.

[26]   J. Worman and Y. Ma, "Thermal performance of EPC eGaN FETs," *Efficient Power Conversion Application Note*, 2011.

[27]   M. S. Liu, L. A. Bursill, S. Prawer *et al.,* "Temperature dependence of Raman scattering in single crystal GaN films," *Appl. Phys. Lett.*, vol. 74, no. 21, May 1999, pp. 3125-3127.

[28]   M. Kuball, J. M. Hayes, M. J. Uren et al., "Measurement of temperature in active high-power AlGaN/GaN HFETs using Raman spectroscopy," *IEEE Electron Device Lett.,* vol. 23, no. 1 Jan. 2002, pp. 7-9.

[29]   M. Kuball, S. Rajasingam, A. Sarua *et al.,* "Measurement of temperature distribution in multifinger AlGaN/GaN heterostructure field-effect transistors using micro-Raman spectroscopy," *Appl. Phys. Lett.*, vol. 82, no. 1, Jan. 2003, pp. 124-126.

[30]   Z Yan, G. Liu, J. M. Khan, and A. A. Balandin, "Graphene quilts for thermal management of high-power GaN transistors," *Nature Comm.*, May 2012, pp. 1-8.

[31]   S. Rajasingam, J. W. Pomeroy, M. Kuball *et al.,* "Micro-Raman temperature measurements for electric field assessment in active AlGaN-GaN HFETs," *IEEE Electron Device Lett.*, vol. 25, no. 7, Jul. 2004, pp. 456-458.

[32]   A. Sarua, H. Ji, K. P. Hilton *et al.,* "Thermal boundary resistance between GaN and substrate in AlGaN/GaN electronic devices," *IEEE Trans. Electron Devices*, vol. 54, no. 12, Dec. 2007, pp. 3152-3158.

[33]   A. Manoi, J. W. Pomeroy, N. Killat, and M. Kuball, "Benchmarking of thermal boundary resistance in AlGaN/GaN HEMTs on SiC substrates: implications of the nucleation layer microstructure," *IEEE Electron Device Lett.*, vol. 31, no. 12, Dec. 2010, pp. 1395-1397.

[34]   G. J. Riedel, J. W. Pomeroy, K. P. Hilton *et al.,* "Nanosecond timescale thermal dynamics of AlGaN/GaN electronic devices," *IEEE Electron Device Lett.*, vol. 29, no. 5, May 2008, pp. 416-418.

[35] S. Nuttinck. R. Mukhopadhyay, C. Loper *et al.,* "Direct on-wafer non-invasive thermal monitoring of AlGaN/GaN power HFETs under microwave large signal conditions," *Proc. 12th GAAS Symposium*, Amersterdam, NL, 2004, pp. 79-82.

[36] A. Sarua, H. Ji, M. Kuball, "Integrated micro-Raman/infrared thermography probe for monitoring of self-heating in AlGaN/GaN transistor structures," *IEEE Trans. Electron Devices*, vol. 53, no. 10, Oct. 2006, pp. 2438-2447.

[37] P. W. Webb, "Thermal imaging of electronic devices with low surface emissivity," *Proc. IEE*, vol. 138, no. 3, Jun. 1991, pp. 390-400.

[38] G. C. Albright, J. A. Stump, and J. D. McDonald, "True temperature measurements on microscopic semiconductor targets," *Proc. SPIE Conf. Thermosense XXI*, Orlando, FL, 1999, pp. 245-250.

[39] J. Kuzmik, P. Javorka, A. Alam *et al.,* "Determination of channel temperature in AlGaN/GaN HEMTs grown on sapphire and silicon substrates using DC characterization method," *IEEE Trans. Electron Devices*, vol. 49, no. 8, Aug. 2002, pp. 1496-1498.

[40] J. Kuzmik, S. Bychikhin, M. Neuburger *et al.,* "Transient thermal characterization of AlGaN/GaN HEMTs grown on silicon," *IEEE Trans. Electron Devices*, vol. 52, no. 8, Aug. 2005, pp. 1698-1705.

[41] J. Kuzmik, S. Bychikhin, D. Pogany *et al.,* "Investigation of the thermal boundary resistance at the III-nitride/substrate interface using optical methods," *J. Appl. Phys.*, vol. 101, pp. 1-7.

[42] J. G. Felbinger, M. V. S. Chandra, Y. Sun *et al.,* "Comparison of GaN HEMTs on diamond and SiC substrates," *IEEE Electron Device Lett.*, vol. 28, no. 11, Nov. 2007, pp. 948-950.

[43] "Intrachip/Interchip Enhanced Cooling Fundamentals (ICECool Fundaments)", DARPA Broad Agency Announcement DARPA-BAA-12-50, Jun. 2012.

[44] N. Tsurumi, H. Ueno, T. Murata *et al.,* "AlN passivation over AlGaN/GaN HFETs for surface heat spreading," *IEEE Trans. Electron Devices*, vol. 57, no. 5, May 2010.

[45] M. Tadjer, T. J. Anderson, K. D. Hobart *et al.,* "Reduced self-heating in AlGaN/GaN HEMTs using nanocrystalline diamond heat-spreading films," *IEEE Electron Device Lett.*, vol. 33, no.1, Jan. 2012, pp. 23-25.

[46] K. Hirama, M. Kasu, and Y. Taniyasu, "RF high-power operation of AlGaN/GaN HEMTs epitaxially grown on diamond," *IEEE Electron Device Lett.*, vol. 33, no. 4, Apr. 2012.

[47] N. Otsuka, S. Nagai, M. Yanagihara *et al.,* "Low-pressure direct-liquid-cooling technology for GaN power transistors," *Japanese J. Appl. Phys.*, vol. 50, Apr. 2011, pp. 1-5.

[48] R. Quay, *Gallium Nitride Electronics*, Berlin, Germany: Springer, 2008.

[49] M. Albulet, *RF Power Amplifiers*, Atlanta, GA: Noble, 2001.

[50] Silvaco Inc., *ATLAS User's Manual: Device Simulation Software*, Aug. 2011.

[51] J. C. Freeman and W. Mueller, "Channel temperature determination for AlGaN/GaN HEMTs on SiC and sapphire," NASA Technical Report NASA/TM-2008-215444, Dec. 2008.

[52] G. K. Wachutka, "Rigorous thermodynamic treatment of heat generation and conduction in semiconductor device modeling," *IEEE Trans. CAD*, vol. 9, no. 11, Nov. 1990, pp. 1141-1149.

[53]   O. I. Saadat, K. R. Bagnall, T. Fujishima, D. Piedra, J. R. Lachapelle, E. N. Wang, and T. Palacios, "Schottky diode based in-situ temperature sensors for AlGaN/GaN HEMTs," *Proc. International Workshop on Nitride Semiconductors*, Sapporo, Japan, Oct. 2012.

[54]   D. Brunner, H. Angerer, E. Bustarret *et al.,* "Optical constants of epitaxial AlGaN films and their temperature dependence," *J. Appl. Phys.*, vol. 82, no. 10, Nov. 1997, pp. 5090-5096.

[55]   O. Ambacher, J. Smart, J. R. Shealy *et al.,* "Two-dimensional electron gases induced by spontaneous and piezoelectric charges in N- and Ga-face AlGaN/GaN heterostructures," *J. Appl. Phys.*, vol. 85, no. 6, Mar. 1999, pp. 3222-3233.

[56]   S. K. Cheung and N. W. Cheung, "Extraction of Schottky diode parameters from forward current-voltage characteristics," *Appl. Phys. Lett.*, vol. 49, no. 2, Jul. 1986, pp. 85-87.

[57]   Z. Lin, W. Lu, J. Lee, D. Liu, "Barrier heights of Schottky contacts on strained AlGaN/GaN heterostructures: determination and effect of metal work functions," *Appl. Phys. Lett.,* vol. 82, no. 24, Jun. 2003, pp. 4364-4366.

[58]   Department of Electrical and Computer Engineering, Brigham Young University. (2009). "Measuring metal-semiconductor junction contact resistance using the transmission line method (TLM)" [Online]. Available: http://www.cleanroom.byu.edu/contact_resistance .phtml.

[59]   M. Farahmand, C. Garetto, E. Bellotti *et al.,* "Monte Carlo simulation of electron transport in the III-nitride wurtzite phase materials system: binaries and ternaries," *IEEE Trans. Electron Devices*, vol. 48, no. 3, Mar. 2001, pp. 535-542.

[60]   X.-D. Wong, W.-D. Hu, X.-S. Chen, and W. Lu, "The study of self-heating and hot-electron effects in AlGaN/GaN double-channel HEMTs," *IEEE Trans. Electron Devices*, vol. 59, no. 5, May 2012, pp. 1393-1401.

[61]   N. Maeda, K. Tsubaki, T. Saitoh, and N. Kobayashi, "High-temperature electron transport properties in AlGaN/GaN heterostructures," *Appl. Phys. Lett.*, vol. 79, no. 11, Sep. 2001, pp. 1634-1636.

[62]   T.-H. Yu and K. F. Brennan, "Monte Carlo calculation of two-dimensional electron dynamics in GaN-AlGaN heterostructures," *J. Appl. Phys.*, vol. 91, no. 6, Mar. 2002, pp. 3730-3736.

[63]   G. Chen, *Nanoscale Energy Transport and Conversion*, New York, NY: Oxford UP, 2005.

[64]   A. Sarua, H. Ji, K. P. Hilton *et al.,* "Thermal boundary resistance between GaN and substrate in AlGaN/GaN electronic devices," *IEEE Transactions on electron devices*, vol. 54, no. 12, pp. 3152-3158, Dec. 2007.

[65]   Cree Inc. (2011). *Silicon carbide substrates and epitaxy* [Online]. Available: http://scn.cree.com/_onelink_/cree/en2zh/products/pdf/MAT-CATALOG.pdf.

[66]   W. Liu and A. A. Balandin, "Thermal conduction in AlxGa1-xN alloys and thin films," *J. Appl. Phys.*, vol. 97, Mar. 2005, p. 073710.

[67]   H. F. Cooke, "Precise technique finds FET thermal resistance," *Microwaves and RF*, vol. 25, Aug. 1986, pp. 85-87.

[68]   P. W. Webb and I. A. D. Russell, "Thermal resistance of gallium-arsenide field-effect transistors," *Proc. IEE*, vol. 136, no. 5, Oct. 1989, pp. 229-234.

[69]   R. Anholt, *Electrical and thermal characterization of MESFETs, HEMTs, and HBTs*, Norwood, MA: Artech House, 1995, pp. 56-59.

[70] M. Hosch, J. W. Pomeroy, A. Sarua *et al.,* "Field dependent self-heating effects in high-power AlGaN/GaN HEMTs," *Proc. CS MANTECH Conf.*, Tampa, FL, May 2009.

[71] Y. S. Muzychka, K. R. Bagnall, and E. N. Wang, "Thermal spreading resistance and heat source temperature in compound orthotropic systems with interfacial resistance," accepted in *IEEE Trans. Components, Packaging, and Manufacturing Technology*, Mar. 2013.

[72] A. G. Kokkas, "Thermal analysis of multiple-layer structures," *IEEE Transactions on Electron Devices*, vol. ED-21, no. 11, pp. 674-681, Nov. 1974.

[73] F. N. Masana, "A closed form solution of junction to substrate thermal resistance in semiconductor chips," *IEEE Transactions on Components, Packaging, and Manufacturing Technology – Part A*, vol. 19, no. 4, pp. 539-545, Dec. 1996.

[74] M. M. Yovanovich, "Conduction and thermal contact resistances (conductances)," in *Handbook of Heat Transfer*, 3rd ed., W. M. Rohsenow, J. P. Hartnett, and Y. I. Cho, Eds. New York, NY: McGraw-Hill, 1998, pp. 3.34-3.35.

[75] J. Cho, E. Bozorg-Grayeli, D. H. Altman *et al.,* "Low thermal resistances at GaN-SiC interfaces for HEMT technology," *IEEE Electron Device Letters*, vol. 33, no. 3, pp. 378-380, Mar. 2012.

[76] G. J. Riedel, J. W. Pomeroy, K. P. Hilton *et al.,* "Reducing thermal resistance of AlGaN/GaN electronic devices using novel nucleation layers," *IEEE Electron Device Lett.,* vol. 30, no. 2, Feb. 2009, pp. 103-105.

[77] R. Chu, A. Corrion, M. Chen *et al.*, "1200-V normally off GaN-on-Si field-effect transistors with low dynamic on-resistance," *IEEE Electron Device Lett.,* vol. 32, no. 5, May 2011.

[78] D. Berning, J. Reichl, A. Hefner *et al.*, "High speed IGBT module transient thermal response measurements for model validation," *Industry Applications Conference, 2003. 38th IAS Annual Meeting. Conference Record of the*, vol. 3, pp. 1826- 1832, Oct. 2003.

[79] J. D. Li, S. S. Choi, and D. M. Vilathgamuwa, "Impact of voltage phase shift on motor loads and series custom power devices including converter thermal effects," *IEEE Trans. Power Delivery*, vol. 19, no. 4, Oct. 2004.

[80] J. M. Ortiz-Rodriguez, M. Hernandez-Mora, T. H. Duong *et al.*, "Thermal network component models for 10 kV SiC power module packages," *IEEE Power Electronics Specialists Conference 2008*, pp.4770-4775, Jun. 2008.

[81] E. Kohn, I. Daumiller, M. Kunze *et al.,* "Switching behavior of GaN-based HFETs: thermal and electronic transients," *Electronics Letters*, vol. 38, no. 12, Jun. 2002.

[82] J. Kuzmik, S. Bychikhin, M. Neuburger *et al.,* "Transient thermal characterization of AlGaN/GaN HEMTs grown on silicon," *IEEE Trans. Electron Devices*, vol. 52, no. 8, Aug. 2005.

[83] J. Xu, W.-Y. Yin, and J. Mao, "Transient thermal analysis of GaN heterojunction transistors (HFETs) for high-power applications," *IEEE Microwave and Wireless Components Lett.,* vol. 17, no. 1, Jan. 2007.

[84] W. Saito, Y. Saito, H. Fujimoto *et al.*, "Switching controllability of high voltage GaN-HEMTs and the cascode connection." *Proc. 2012 24th Intl. Symp. On Power Semicond. Devices and ICs, Bruges*, 2012, pp. 229-232.

[85] W. Saito, Semiconductor Company, Toshiba Corporation, Kawasaki, Japan, private communication, Aug. 2012.

[86] Y. K. Koh, Y. Cao, D. G. Cahill *et al.*, "Heat-transport mechanisms in superlattices," *Advanced Functional Materials,* vol. 19, no. 4, Feb. 2009.

[87]   G. Kirchhoff, *Vorlesungen uber die theorie der varme*, Leipzig: Teubner, 1894, p. 13.

[88]   H. S. Carslaw and J. C. Jaeger, *Conduction of heat in solids*, New York, NY: Oxford UP, 1959, p. 11.

[89]   V. S. Arpaci, *Conduction heat transfer: abridged edition*, Needham Heights, MA: Ginn Press, 1991.

[90]   S. Kakac and Y. Yener, *Heat conduction*, Ankara, Turkey: Middle East Technical University, 1979, pp. 60-62.

[91]   W. B. Joyce, "Thermal resistance of heat sinks with temperature-dependent thermal conductivity," *Solid-state Electronics*, vol. 18, Sep. 1974, pp. 321-322.

[92]   A. R. Batchelor, "Steady-state heat sink analysis for two-terminal microwave oscillator devices with temperature dependent thermal conductivity," *Proc. Math. Phys. Sci.*, vol. 441, no. 1911, Apr. 1993, pp. 181-189.

[93]   P. Vernotte, "Formes diverses de l'equation de la chaleur," *Comptes Rendus*, vol. 218, Jan. 1944, pp. 39-41.

[94]   M. N. Ozisik, *Boundary value problems of heat conduction*, Scranton, PA: International Textbook, 1968, pp. 353-356.

[95]   G. W. C. Kaye and T. H. Laby, *Tables of physical and chemical constants*, London: Longman, 1986, p. 62.

[96]   D. G. Cahill, S.-M. Lee and T. I. Selinder, "Thermal conductivity of κ-Al2O3 and α-Al2O3 wear-resistant coatings," *J. Appl. Phys.*, vol. 83, no. 11, Jun. 1998, pp. 5783-5786.

[97]   P. Hui and H. S. Tan, "On the effect of nonlinear boundary conditions for heat conduction in diamond heat spreaders with temperature-dependent thermal conductivity," *IEEE Trans. Components, Packaging, and Manufacturing Technology – Part A*, vol. 20, no. 4, Dec. 1997, pp. 537-540.

[98]   F. Bonani and G. Ghione, "On the application of the Kirchhoff transformation to the steady-state thermal analysis of semiconductor devices with temperature-dependent and piecewise inhomogeneous thermal conductivity," *Solid-State Electronics*, vol. 38, no. 7, Sept. 1994, pp. 1409-1412.

[99]   "Near Junction Thermal Transport (NJTT)", DARPA Broad Agency Announcement DARPA-BAA-11-09, Nov. 2011.