Nonlinear Electrothermal Monte Carlo Device Simulation

by

Ky E. Merrill

A Dissertation Presented in Partial Fulfillment
of the Requirement for the Degree
Doctor of Philosophy

Approved April 2020 by the
Graduate Supervisory Committee:

Marco Saraniti, Chair
Stephen Goodnick
David Smith
Robert Wang

ARIZONA STATE UNIVERSITY

May 2020

ABSTRACT

A model of self-heating is incorporated into a Cellular Monte Carlo (CMC) particle-based device simulator through the solution of an energy balance equation (EBE) for phonons. The EBE self-consistently couples charge and heat transport in the simulation through a novel approach to computing the heat generation rate in the device under study. First, the moments of the Boltzmann Transport Equation (BTE) are discussed, and subsequently the EBE of for phonons is derived. Subsequently, several tests are performed to verify the applicability and accuracy of a nonlinear iterative method for the solution of the EBE in the presence of convective boundary conditions, as compared to a finite element analysis solver as well as using the Kirchhoff transformation. The coupled electrothermal characterization of a GaN/AlGaN high electron mobility transistor (HEMT) is then performed, and the effects of non-ideal interfaces and boundary conditions are studied.

The proposed thermal model is then applied to a novel Π-gate architecture which has been suggested to reduce hot electron generation in the device, compared to the conventional T-gate. Additionally, small signal ac simulations are performed for the determination of cutoff frequencies using the thermal model as well.

Finally, further extensions of the CMC algorithm used in this work are discussed, including 1) higher-order moments of the phonon BTE, 2) coupling to phonon Monte Carlo simulations, and 3) application to other large-bandgap, and therefore high-power, materials such as diamond.

ACKNOWLEDGEMENTS

Firstly, I would like to thank my advisor Dr. Marco Saraniti for his patience and continual support. Especially for allowing me the time to sort through the interests I wanted to pursue, both personally and professionally, and the leeway to make mistakes in order to learn from them. Furthermore, especially for his guidance and suggestions leading to the completion of this dissertation, without which it would not have been possible.

I would also like to express my gratitude to the members of my advisory committee: Dr. Stephen Goodnick, Dr. David Smith, and Dr. Robert Wang for their time and expertise in evaluating this work, as well as their comments and suggestions which have greatly improved it.

Additionally, I would like to thank Dr. Tom Heydt for more interesting conversations that I can count and his advice over the years both personally and professionally.

Finally, I want to thank Daisy for her patience and immeasurable help in creating this work, and always reminding me that I have to sleep at some point.

TABLE OF CONTENTS

CHAPTER Page

LIST OF TABLES

LIST OF FIGURES

Figure Page

Figure Page

Figure Page

Chapter 1

INTRODUCTION

The semiconductor industry is constantly shrinking devices in order to increase performance and yield. This comes at the cost of increased power density losses, which in turn causes a large amount of heat to accumulate within the devices. Thermal effects can lead to the degradation of electrical characteristics as well as reliability issues [1–4] resulting in device breakdown. As devices are pushed even further, both in terms of smaller size and higher performance, these problems are only expected to become more severe. In addition, new architectures such as the FinFET and 3-D integrated circuit layouts may exacerbate thermal problems.

These problems illustrate the necessity of bringing the simulation and study of thermal characteristics closer to the current state of electrical characteristics in the context of a comprehensive electrothermal device simulation, as the two are closely intertwined. Coupled electrothermal simulations offer the possibility of obtaining temperature maps within a device, allowing the "hot spot" in a device to be more accurately predicted. Hence, the potential effects of self-heating may be better understood prior to fabrication. Furthermore, these results can be used to mitigate the negative effects associated with self-heating either through more efficient heat dissipation and/or modified device layouts leading to a reduction in heat generation and peak temperatures.

Monte Carlo (MC) techniques have been widely utilized to study electrical transport in semiconductor materials and devices in a number of seminal works [5–8] which have been built upon in the following decades. Although drift-diffusion and hyrdrodynamic simulation models offer shorter simulation times than MC methods,

they are generally based on low-order moments of the Boltzmann Transport Equation (BTE) and become inaccurate as devices are pushed smaller in size and correspondingly closer to material limits. Furthermore, assumptions made in the derivation of these models are often only valid near-equilibrium. On the other hand, MC methods provide a full stochastic solution to the BTE, and hence improved accuracy, at the expense of being more expensive in simulation time, hardware requirements, or both.

However, the use of MC methods in thermal device simulations has been less widespread until more recently. A rigorous and comprehensive treatment of self-heating and heat transport in devices, from a thermodynamics viewpoint, is given in [9]. As for simulation approaches, the method of incorporating self-heating effects using a set of hydrodynamic equations, *i.e.*, the first three moments of the BTE, was developed in [10–12], and subsequently used in more recent studies [13, 14]. This approach uses relaxation time approximations (RTAs) for both the energy transfer between electrons and optical phonons as well as that between optical and acoustic phonons. In addition, it assumes a specified energy decay path starting from electrons to optical phonons and then to acoustic phonons. While these methods take into account temperature-dependent scattering rates, the temperature dependence of material properties, such as thermal conductivity, is often neglected.

One approach to incorporate these temperature-dependent thermal conductivities at the device scale is that of Bonani and Ghione [15] based on the Kirchhoff transformation, which implicitly maps the temperature dependence into a new "apparent temperature" variable. The Kirchhoff transformation is discussed in detail in Sec. 3.3. Elsewhere, Pilgrim *et al.* [16–19] used the Kirchhoff transformation along with a net phonon emission approach to obtain internal temperature maps of high electron mobility transistor (HEMT) devices. Here, it is assumed that all optical phonons emitted and/or absorbed correspond to only a single energy value, *i.e.*, the optical

phonons are assumed to be dispersionless. An analytical temperature distribution was then computed using a thermal resistance matrix method developed in [20]. It can be seen, however, that issues arise when using the Kirchhoff transformation in the presence of non-ideal interfaces [21, 22], as discussed in Section 3.3 and in Chapter 4.

The rather comprehensive model of Hao *et al.* [23, 24] improves upon many of these aspects by directly coupling electron and phonon MC simulations in particular regions of interest within a device. The reason for limiting its use to specific regions is that the phonon MC algorithm is quite expensive computationally, and hence in the majority of the device a simpler, faster Fourier's law solution is used. In these works, the electron MC is used to extract parameters such as the electron density, energy, and drift velocity in order to determine the energy exhanged between electrons and phonons.

The aim of the work detailed here is to provide electrothermal simulation functionality within the Cellular Monte Carlo (CMC) framework [25], which offers several performance advantages to the more traditional Monte Carlo techniques. Throughout this work, a full electronic bandstructure and full phonon dispersion representation is used, including the full dispersion for the optical phonon modes rather than assuming them to be dispersionless. In addition, a novel method of computing the heat generation rate in the heat transport equation in real time by tracking the individual scattering events is utilized, rather than by extracting parameters such as electron density, energy, and drift velocity to be used within a relaxation time approximation. This new approach allows us to move away from the RTA for electrons and provide a more direct self-consistent coupling between charge and heat transport.

In Chapter 2 an overview of the method of moments for the solution of the BTE is given, as well as the drift-diffusion and hydrodynamic methods obtained with them. These are then contrasted with Monte Carlo methods used in the Cellular Monte Carlo

approach. The energy balance equation for phonons is derived from the moments of the phonon BTE, and its manipulation into the form of a linear elliptical partial differential equation (PDE) through the use of the Kirchhoff transformation and an iterative approach to its full solition are introduced in Chapter 3.

Subsequently, the structure and properties of the computational grid and the boundary conditions are discussed in Chapter 3. Benchmark simulation results for both the Kirchhoff transformation and the iterative approach are shown in Chapter 4. In particular, the iterative approach overcomes two limitations of using the Kirchhoff transformation: it allows for 1) the straightforward implementation of non-ideal interfaces, and 2) the inclusion of temperature dependence of properties such as the thermal conductivity within a finite-difference framework. It will be shown that the Kirchhoff transformation requires *a priori* knowledge of the solution in the case of non-ideal interfaces, while the iterative method does not.

Additionally, fully coupled nonlinear electrothermal simulations performed on a simple GaAs diode as well as an experimentally characterized GaN high electron mobility transistor (HEMT) using both approaches are detailed in Chapters 5 and 6, respectively.

Finally, these results are then built upon in Chapter 7 to study a novel gate structure and the outcome of extending the electrothermal model to small-signal AC simulations.

Chapter 2

SIMULATION METHODS

Generally speaking, an electron device simulation includes the coupling of the charge transport and scattering with the solution of Poisson's equation in addition to the application of suitable boundary conditions. Mathematically, a differential equation must include a Dirichlet boundary condition at some point in the domain for a unique solution to exist. In the case of a problem with only Neumann boundary conditions by constrast, one obtains an infinite family of solutions differing from each other by a constant. Fortunately, from a physical perspective, an electrical device always includes contacts where a voltage is applied which corresponds to a Dirichlet condition on the electrostatic potential computed as a solution of Poisson's equation.

For the case of charge transport, the majority of simulation approaches used for semiconductor devices are based on the BTE. Rather than solving Newton-like equations for every particle, the BTE instead describes the particle *ensemble* with what is known as the particle distribution function, $f(\mathbf{r}, \mathbf{k}, t)$. This distribution function represents the probability of finding a particle at a specific time t, position $\mathbf{r}$, and momentum $\mathbf{k}$. Hence, $f(\mathbf{r}, \mathbf{k}, t)$ describes the time evolution of the particle distribution in phase space.

The BTE therefore represents an equation which simply accounts for all the ways in which the distribution function of the carrier ensemble may undergo changes and is written as [26–29]

$$\frac{\partial f}{\partial t} = -\nabla_{\mathbf{r}} \cdot \left[\frac{d\mathbf{r}}{dt} f\right] - \nabla_{\mathbf{k}} \cdot \left[\frac{d\mathbf{k}}{dt} f\right] + \left.\frac{\partial f}{\partial t}\right|_{coll}. \tag{2.1}$$

This equation says that the time variation of the distribution function, the left-hand side (LHS) of Eq. (2.1), is given by the contributions from the flux in position space, the first term on the right-hand side (RHS), and the flux in momentum space, the second term on the RHS. The latter two of these are the so called "drift terms", and are entirely classical in nature. Additionally, the final term on the RHS, often called the collision term, describes changes in the distribution function due to any collisions the carrier might undergo and is computed including quantum mechanical considerations. Hence, these methods are described as the semi-classical approach.

Using the relations for velocity, $\frac{d\mathbf{r}}{dt} = \mathbf{v}$ and force, $\mathbf{F} = \hbar\frac{d\mathbf{k}}{dt}$, Eq. (2.1) can then be written as

$$\frac{\partial f}{\partial t} = -\frac{1}{\hbar}\mathbf{F}\cdot\nabla_k f - \mathbf{v}\cdot\nabla_r f + \left.\frac{\partial f}{\partial t}\right|_{coll}, \tag{2.2}$$

which is a commonly seen form of the BTE [27, 30].

2.1 Moments of a Distribution Function

The drift-diffusion and hydrodynamic models can be derived from Eq. (2.2) by using the first two (the zeroeth and first) moments or the first three (zeroeth, first, and second) moments of the distribution function, respectively. In general, if X is a continuous random variable, and $g(x)$ its continuous probability distribution function, then the nth central moment, $\langle X^n\rangle$, of X is defined as [31, 32]

$$\langle X^n\rangle = \int_{-\infty}^{\infty} x^n g(x)dx, \tag{2.3}$$

where $\langle X^n\rangle$ is also called the expected value of X^n. On the other hand, the nth moment about the mean is given by $\left\langle (X-\mu)^2\right\rangle$

$$\left\langle (X-\mu)^2\right\rangle = \int_{-\infty}^{\infty} (x-\mu)^n g(x)dx \tag{2.4}$$

where μ is the expected value or mean of the distribution. In fact, μ is simply the first moment corresponding to $n = 1$ in Eq. (2.3).

The nth moment of a distribution is said to exist if the integral in Eq. (2.3) is absolutely convergent. That is:

$$\begin{aligned}\int_{-\infty}^{\infty} x^n g(x)dx &= \int_{-\infty}^{0} x^n g(x)dx + \int_{0}^{\infty} x^n g(x)dx \\ &= \lim_{a\to-\infty}\int_{a}^{0} x^n g(x)dx + \lim_{b\to\infty}\int_{0}^{b} x^n g(x)dx\end{aligned} \tag{2.5}$$

where both of the limit terms must converge to finite values. Subsequently, if the nth moment exists then all moments of an order lower than n also exist [33].

The moments of a distribution (provided they exist) are, in a qualitative sense, analogous to the terms of a Taylor series expansion of a function. For a Taylor expansion, each successive higher-order term gives a more accurate approximation to the true function itself. In the same way, successive higher-order moments for a probability distribution give a more accurate representation of its shape. In statistics, in addition to the mean it is common to use the moments for $n = 2, 3$, and 4 corresponding to $\langle X^2\rangle = \sigma^2$ being the variance of $g(x)$, $\langle X^3\rangle$ a measure of its skewness (or asymmetry), and $\langle X^4\rangle$ its kurtosis which contains information about both the tailedness and peakedness of the distribution. Note, however, that the information in the kurtosis is separate from that contained in the variance. This is discussed here as some hydrodynamic models do take into account the 4th moment as well as even higher-order moments [34].

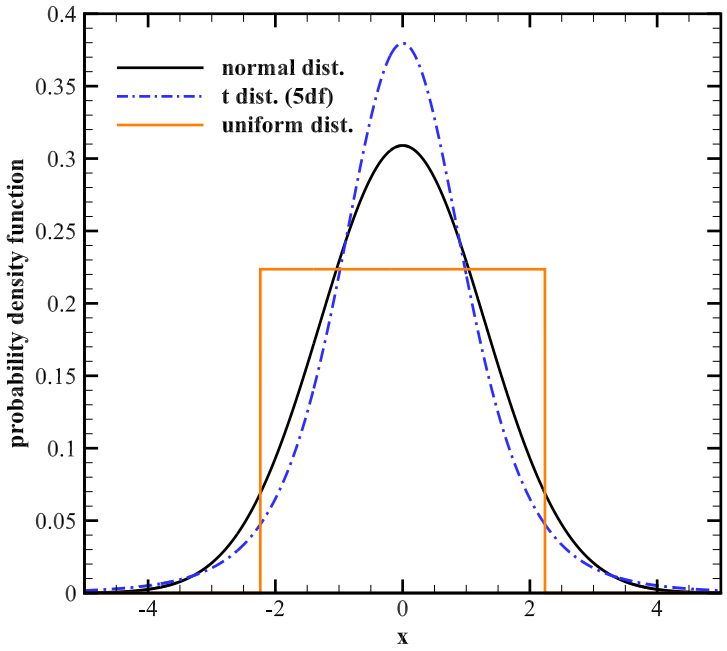
A popular measure of kurtosis is normalized such that the normal distribution possesses a kurtosis of zero, and that of any other distribution is relative to the normal distribution. This value is called the excess kurtosis [35]. In general, a positive excess kurtosis corresponds to a distribution with both a sharper peak and larger tails than

a normal distribution, while a negative excess kurtotsis corresponds to a distribution with both a flatter peak and smaller tails [35]. In Fig. 2.1a, all of the distributions shown are scaled to a variance of 5/3 to match that of the student t-distribution [36].

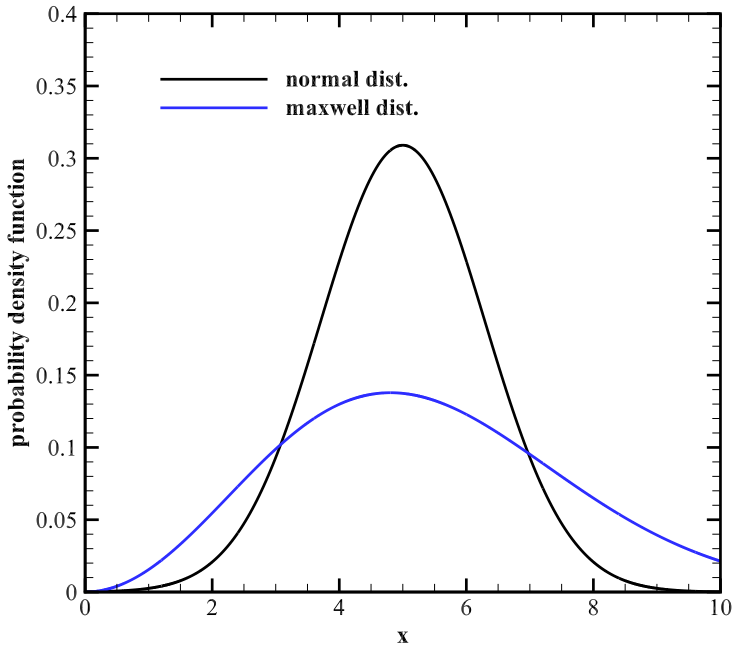
The student t-distribution with 5 degrees of freedom is shown to represent a positive excess kurtosis distribution, while a uniform distribution is shown to represent negative excess kurtosis [35]. Figure 2.1b shows the Maxwell-Boltzmann distribution often used in semiconductor theory, which has only a slightly positive excess kurtosis of ≈ 0.1. This may be surprising initially, but looking at Fig 2.1, it is seen that while for the Maxwell distribution the tails are clearly "heavier" the peak is much flatter. In comparison, the t-distribution, with a much larger excess kurtosis of 3, also has "heavier" tails than the normal distribution, but in addition has a sharper peak and the "shoulder" area in between dips below that of the normal distribution. The reason changes in both the tails and peak are involved in a change in kurtosis is that it represents a change in the distribution function that does not affect its variance [35]. Hence, if the tails of the distribution become "heavier" they must be offset by the distribution becoming sharper near the peak in order to leave the variance unchanged.

In terms of an electron distribution function, the kurtosis would physically correspond to additional information about the "hot" electron population in the distribution, which are those that have a much higher energy than the mean, *i.e.*, the effects of the outliers of the distribution function.

As a further note, while including successive higher-order moments results in a more accurate representation of the actual distribution function, it should be clear that whatever information is contained in the discarded higher-order moments is lost.



(a)



(b)

Figure 2.1: (a) The normal distribution (zero kurtosis) compared to the student t-distribution with 5 degrees of freedom (positive kurtosis) and the uniform distribution (negative kurtosis). Reproduced from [35]. (b) Maxwell-Boltzmann distribution compared to normal distribution. All distributions shown are set to a variance of 5/3.

2.2 Moments of the BTE

Starting with Eq. (2.3), if we instead use the carrier distribution function of Eq. 2.2 rather than $g(x)$, and use orders of the wavevector $\mathbf{k}$ rather than x, where $\phi(\mathbf{k}^n)$ is some function of order n, the nth moment of the carrier distribution function, n_ϕ, is given by [6, 27, 37]

$$n_\phi = \int \phi(\mathbf{k}^n) f\left(\mathbf{r}, \mathbf{k}, t\right) dk. \tag{2.6}$$

For the zeroeth moment in particular, $n = 0$ and we obtain

$$n_\phi = \int \phi(\mathbf{k}^0) f\left(\mathbf{r}, \mathbf{k}, t\right) dk. \tag{2.7}$$

which simply gives the total number of carriers. It follows that the carrier density, n, is obtained by dividing Eq. (2.7) by a unit volume of the reciprocal space, and hence the carrier density is found from the zeroeth moment of the distribution function,

$$n = \frac{1}{\Omega} \int \phi(\mathbf{k}^0) f\left(\mathbf{r}, \mathbf{k}, t\right) dk. \tag{2.8}$$

Similarly, from the first moment one obtains the average carrier velocity, or correspondingly the current density, and from the second moment one obtains the average particle energy.

Just as we can take moments of the distribution function, we can also take the moments of the BTE itself. This is done by multiplying Eq. (2.2) by $\phi(\mathbf{k})/\Omega$ and summing (or integrating) over momentum space. Rearranging terms, the general moment equation can then be written as [37–39]

$$\frac{1}{\Omega}\frac{\partial}{\partial t}\sum_{\mathbf{k}} \phi(\mathbf{k}) f = -\frac{1}{\Omega}\nabla_{\mathbf{r}} \cdot \sum_{\mathbf{k}} \phi(\mathbf{k}) \mathbf{v} f + \frac{\mathbf{F}}{\Omega\hbar} \sum_{\mathbf{k}} f \nabla_{\mathbf{k}} \phi(\mathbf{k}) + \frac{1}{\Omega} \sum_{\mathbf{k}} \phi(\mathbf{k}) \left.\frac{\partial f}{\partial t}\right|_{coll}. \tag{2.9}$$

Taking the zeroeth moment using $\phi(\mathbf{k}^0) = 1$ and using Eq. (2.8), the term on the LHS of Eq. (2.9) is just the time derivative of the carrier density from Eq. 2.8, $\partial n/\partial t$. The first term on the RHS is a carrier flux, denoted by $F_{\phi,0}$,

$$F_{\phi,0} = \frac{1}{\Omega}\sum_{\mathbf{k}} \phi(\mathbf{k}^0)\mathbf{v}f = \frac{1}{\Omega}\sum_{\mathbf{k}} \mathbf{v}f = n\mathbf{v} = -\frac{\mathbf{J}_n}{q}, \tag{2.10}$$

where q and $\mathbf{J}_n$ are the usual carrier charge and the current density, respectively. As for the second term on the RHS, if $\phi(\mathbf{k})$ is a constant then $\nabla_{\mathbf{k}}\phi(\mathbf{k}) = 0$ and hence this entire term is zero. Finally, if we let $\mathcal{S}_n$ denote the change in the distribution function due to collisions we obtain from the zeroeth moment of the BTE:

$$\frac{\partial n}{\partial t} = \nabla_{\mathbf{r}} \cdot F_{\phi,0} + \mathcal{S}_n, \tag{2.11}$$

which is the charge continuity equation. However, if we consider the velocity to be some general function of $\mathbf{k}$ such that $\mathbf{v} = u(\mathbf{k})$, then from Eq. (2.10)

$$F_{\phi,0} = -\frac{\mathbf{J}_n}{q} = \frac{1}{\Omega}\sum_{\mathbf{k}} u(\mathbf{k})f, \tag{2.12}$$

which is a first order moment of the distribution function proportional to $\mathbf{k}^1$. Hence, by taking the zeroeth moment of the BTE we end up with an equation which includes the first moment of the distribution function, and therefore the equation is not in closed form.

In turn, for the first moment of the electron BTE, corresponding to $n = 1$, we assume $\phi(\mathbf{k}^1) = u(\mathbf{k})$. That is, we assume that $\phi(\mathbf{k})$ is some general function of the carrier velocity as in Eq. (2.12). Here, a parabolic bandstructure is often assumed in order to simplify the derivation. In that case, the relation $u(\mathbf{k}_i) = \hbar\mathbf{k}_i/m^*$ relates the carrier momentum to the velocity where m^* is the effective mass [1]. Then, we can use Eq. (2.10) to obtain the time variation of the current density as the term on the

LHS:

$$\frac{1}{\Omega}\frac{\partial}{\partial t}\sum_{\mathbf{k}}\mathbf{u}(\mathbf{k}_i)f = -\frac{1}{q}\frac{\partial \mathbf{J}_n}{\partial t}. \tag{2.13}$$

Since we assumed a parabolic bandstructure, the energy is of the form $E(\mathbf{k}) = \hbar^2\mathbf{k}_i^2/2m^*$, and we can also express the velocity in the BTE as $\mathbf{v} = \hbar\mathbf{k}_j/m^*$

$$-\frac{1}{\Omega}\nabla_r \cdot \sum_{\mathbf{k}} u(\mathbf{k}_i)\cdot\mathbf{v}f = -\frac{1}{\Omega}\nabla_{\mathbf{r}}\cdot\sum_{\mathbf{k}}\frac{\hbar^2\mathbf{k}_i\mathbf{k}_j}{m^{*2}} = -\frac{2}{m^*}\nabla_{\mathbf{r}}\cdot\mathbf{W}_n, \tag{2.14}$$

where $\mathbf{W}_n = 1/\Omega\sum_{\mathbf{k}}E(\mathbf{k})f$ is the total energy of the carrier ensemble, and hence Eq. (2.14) represents the energy flux of the carrier ensemble. In general, $\mathbf{W}_n$ is a tensor whose trace is the average energy density, however, it is often assumed that the tensor is diagonal [27].

For the second term on the RHS, one obtains:

$$\frac{\mathbf{F}}{\hbar}\frac{1}{\Omega}\sum_{\mathbf{k}} f\nabla_{\mathbf{k}}\mathbf{v} = \frac{\mathbf{F}}{\hbar}n = -\frac{qn\mathbf{E}}{m^*}, \tag{2.15}$$

where $\mathbf{E}$ is the electric field. Lastly, the collision term represents the rate at which momentum is lost by the carriers due to collisions, as collisions result in a change in the state $\mathbf{k}$, and for the time being can simply be denoted by $\mathcal{S}_m$.

The final result for the first moment of Eq. (2.3) is

$$\frac{\partial \mathbf{J}_n}{\partial t} = \frac{2q}{m^*}\nabla_{\mathbf{r}}\cdot\mathbf{W}_n + \frac{nq^2\mathbf{E}}{m^*} - \mathcal{S}_m, \tag{2.16}$$

which says the time variation in the current density is given by contributions from the energy flux, the electric field, and the change of the momentum due to scattering, respectively. Just as the zeroeth moment of the BTE included a term containing the first moment, this equation from the first moment of the BTE contains a term, $\mathbf{W}_n$, which is the expected value of the second moment of the BTE.

In fact, due to the first term on the RHS of Eq. (2.9) containing $\mathbf{v}(\mathbf{k})$, when the BTE is multiplied by the function $\phi(\mathbf{k}^n)$, it always results in a term of the order $\phi(\mathbf{k}^{n+1})$, and so the nth moment of the BTE always has a term involving the $(n+1)$st moment. Furthermore, an infinite number of moments exist for the BTE [40] so that one could obtain an infinite number of equations. In order to close the system of moment equations, a closure relation must be devised to simplify the nth order equation such that it does not depend on the $(n+1)$st order equation. This is done either by simply ignoring a higher-order term [38, 41, 42], by using phenomenological reasoning [27, 37, 43], or by an assumption on the electron distribution function itself [38, 44], which is usually assumed to be a Maxwellian distribution seen in Fig. 2.1b. Additionally, in Eqs. (2.11) and (2.16), a relaxation time approximation (RTA) is used to evaluate the collision terms, $\mathcal{S}_n$ and $\mathcal{S}_m$, respectively.

2.3 The Drift Diffusion Model

The drift-diffusion model is obtained directly from Eqs. (2.11) and (2.16) coupled with Poisson's Equation for the electrostatic potential. To derive the drift-diffusion equation, if we assume that the total energy density is given by only the thermal energy, then

$$W = \frac{3}{2} n k_B T_C, \tag{2.17}$$

where n is the carrier density, k_B is Boltzmann's constant, T_C is the carrier temperature, and where W has a contribution of $W/3$ from each degree of freedom, x, y, and z. Then, taking the divergence of the diagonal $\mathbf{W}$ tensor results in $\frac{1}{3}\nabla_r W$, and Eq. (2.16) becomes

$$\frac{\partial \mathbf{J}_n}{\partial t} = \frac{2}{3}\frac{1}{m^*}\nabla_r W + \frac{nq^2}{m^*}\mathbf{E} - \mathcal{S}_m. \tag{2.18}$$

Next, let the momentum relaxation term $\mathcal{S}_m$ be represented by a RTA of the form

$$\mathcal{S}_m = \frac{\mathbf{J}_n}{\tau_m}, \tag{2.19}$$

where τ_m is a characteristic relaxation time. If we assume that it is short enough that $\mathbf{J}_n$ does not change during τ_m, the time derivative on the LHS of Eq. (2.18) is zero, and we obtain

$$\mathbf{J}_n = \frac{\tau_m n q^2}{m^*}\mathbf{E} + \frac{2}{3}\frac{\tau_m}{m^*}\nabla_r W = \frac{\tau_m n q^2}{m^*}\mathbf{E} + \frac{\tau_m}{m^*}k_B T_C \nabla n + \frac{\tau_m}{m^*}k_B n \nabla T_C. \tag{2.20}$$

Defining the carrier mobility, $\mu_n = (q\tau_m)/(m^*)$, the diffusion coefficient, $D_n = \mu_n k_B T_c/q$, the thermal diffusivity, $D_T = \mu_n k_B n/q$, and subsequently substituting them into Eq. (2.20) results in the drift-diffusion equation for electrons [27, 45]

$$\mathbf{J}_n = q\mu_n n\mathbf{E} + qD_n\nabla_r n + qD_T\nabla T_C. \tag{2.21}$$

Additionally, a similar equation can be found for the holes in terms of $\mathbf{J}_p$, μ_p, D_p, and p. Hence, the drift-diffusion model is given by the following set of equations:

$$\begin{aligned} \frac{\partial n}{\partial t} &= \frac{1}{q}\nabla_{\mathbf{r}}\cdot\mathbf{J}_n + \mathcal{S}_n \\ \frac{\partial p}{\partial t} &= -\frac{1}{q}\nabla_{\mathbf{r}}\cdot\mathbf{J}_p + \mathcal{S}_n \end{aligned} \tag{2.22}$$

$$\begin{aligned} \mathbf{J}_n &= q\mu_n n\mathbf{E} + qD_n\nabla_{\mathbf{r}} n + qD_T\nabla T_C \\ \mathbf{J}_p &= q\mu_n n\mathbf{E} - qD_n\nabla_{\mathbf{r}} p - qD_T\nabla T_C \end{aligned} \tag{2.23}$$

$$\nabla_r\cdot\epsilon\nabla V = -P, \tag{2.24}$$

where P is the forcing function of Poisson's equation, *i.e.*, the total charge density.

In this case, as a closure relation it was assumed that $W = \frac{3}{2}nk_BT_C$, and implicitly that $T_C = T_n = T_p = T_L$ where T_n and T_p are the electron and hole temperatures,

respectively, and T_L is the lattice temperature. In other words, it was assumed that the carrier ensembles n and p are in thermal equilibrium with the lattice. Due to this assumption, non-equilibrium thermal effects cannot be accounted for in this model. This consequence is the principle motivation for the development of the hydrodynamic model from higher-order moments. However, the simple drift-diffusion model is still useful in some cases due to its suitable computational cost.

2.4 Hydrodynamic Models

The hydrodynamic models of charge transport, referred to as such due to their similarity to the equations of fluids, extend and improve upon the drift-diffusion models by including the $n = 2$ moment of the BTE, which allows them to include some non-equilibrium effects and hence improved thermal modeling. This is due to the fact that the balance equation obtained from the additional moment includes information on the average carrier energy [45]. The second moment of the BTE is found by letting $\phi(\mathbf{k}) = E(\mathbf{k}) \propto k^2$. The moment equation is then written as:

$$\frac{1}{\Omega}\frac{\partial}{\partial t}\sum_{\mathbf{k}} E(\mathbf{k})f = -\frac{1}{\Omega}\nabla_r \cdot \sum_{\mathbf{k}} E(\mathbf{k})\mathbf{v}f + \frac{\mathbf{F}}{\Omega\hbar}\sum_{\mathbf{k}} f\nabla_k E(\mathbf{k}) + \frac{1}{\Omega}\sum_{\mathbf{k}} E(\mathbf{k})\left.\frac{\partial f}{\partial t}\right|_{coll} \tag{2.25}$$

The LHS of Eq. (2.25) gives the time variation of the energy density:

$$\frac{1}{\Omega}\frac{\partial}{\partial t}\sum_{\mathbf{k}} E(\mathbf{k})f = \frac{\partial W}{\partial t}, \tag{2.26}$$

where W is the total energy density of the carrier ensemble. The first term on the RHS represents the energy density multiplied by a velocity, or the "flow" of the energy density:

$$-\frac{1}{\Omega}\nabla_{\mathbf{r}} \cdot \sum_{\mathbf{k}} E(\mathbf{k})\mathbf{v}f = -\nabla_{\mathbf{r}} \cdot \mathbf{F}_W, \tag{2.27}$$

where $\mathbf{F}_W$ is an energy flux. This again is a term in the equation for the $n = 2$ moment which includes information from the $n = 3$ moment, just as in the previous cases for the $n = 1$ and $n = 0$ moments, and as such it will require a closure relation. For the second term on the RHS we use the general relation between the carrier energy and velocity, $\nabla_k \frac{1}{\hbar} E(\mathbf{k}) = \mathbf{v}$, as well as the fact that, for an electron, the force due to the electric field is given by $\mathbf{F} = -q\mathbf{E}$ to obtain

$$\frac{\mathbf{F}}{\Omega\hbar} \sum_{\mathbf{k}} f \nabla_{\mathbf{k}} E(\mathbf{k}) = -q\mathbf{E} \sum_{\mathbf{k}} \mathbf{v} f = \mathbf{E} \cdot \mathbf{J}_n, \tag{2.28}$$

which represents the increase in energy from the electric field, and has been used as a first approximation for the heat generation rate in some cases [9].

The final term is simply denoted here as

$$\frac{1}{\Omega} \sum_{\mathbf{k}} \phi(\mathbf{k}) \frac{\partial f}{\partial t}\bigg|_{coll} = \frac{\partial W}{\partial t}\bigg|_{coll}, \tag{2.29}$$

which represents the change in the ensemble energy density due to collisions. As in the drift-diffusion equation, this is often approximated using the RTA of the form

$$\frac{\partial W}{\partial t}\bigg|_{coll} = \frac{(W - W_0)}{\tau_W}, \tag{2.30}$$

where W_0 is the equilibrium energy density and τ_W is the characteristic relaxation time of W towards equilibrium, *i.e.*, towards W_0. Putting these terms together, the final electron energy balance equation is then

$$\frac{\partial W}{\partial t} = -\nabla_r \cdot \mathbf{F}_W + \mathbf{E} \cdot \mathbf{J}_n + \frac{\partial W}{\partial t}\bigg|_{coll}. \tag{2.31}$$

The general set of hydrodynamic equations (for electrons) from the first 3 moments of the BTE is therefore given by:

$$\frac{\partial n}{\partial t} = \frac{1}{q}\nabla_{\mathbf{r}} \cdot \mathbf{J}_n + \mathcal{S}_n \tag{2.32}$$

$$\frac{\partial \mathbf{J}_n}{\partial t} = \frac{2q}{m^*}\nabla_{\mathbf{r}} \cdot \mathbf{W}_n + \frac{nq^2\mathbf{E}}{m^*} - \mathcal{S}_m \tag{2.33}$$

$$\frac{\partial W}{\partial t} = -\nabla_r \cdot \mathbf{F}_W + \mathbf{E} \cdot \mathbf{J}_n + \left.\frac{\partial W}{\partial t}\right|_{coll}. \tag{2.34}$$

In general, a similar equation from each moment would also be obtained for holes as well, as in the drift-diffusion model of Eqs. (2.22), (2.23), and (2.24). While the hydrodynamic model has the advantage of being less computationally expensive than MC, it includes parameters such as the carrier mobilities, μ_C, and effective masses, m_C^*, whose values must be taken as input from elsewhere. Furthermore, it is not uncommon for these values to be computed from MC where they are themselves an output of the simulation. [38].

The hydrodynamic models have been widely used in the literature [38, 41, 46, 47]. An excellent review of the various hydrodynamic models as well as their similarities and differences is given in [45]. These models are discussed here to highlight the differences between transport simulations based on the moments of the BTE and MC methods, as well as the similarity of the electron energy balance equation with the energy balance equation developed for phonons, Eq. (3.8) in Sec. 3.2.

2.5 Monte Carlo Methods

In contrast to the drift-diffusion and hydronamic methods mentioned previously, the MC approach provides a full statistically exact solution to the BTE. In other words, the MC solution is a statistical one which includes all moments of the BTE [5, 7, 8]. Moreover, the MC method uses fewer assumptions and its accuracy does not depend on the macroscopic transport parameters used in the closure relations mentioned previously. To the contrary, these parameters are an output of the MC

simulation.

Monte Carlo simulations were first developed by Ulam and Metropolis for the study of differential equations [48, 49], along with Von Neumann, while working on the Manhattan Project at Los Alamos [50]. Since then, MC techniques have found use in a wide range of fields from finance [51] to the social sciences [52, 53]. In the context of semiconductors, they have been consistently employed since the 1980s to study charge transport, especially under high field conditions where, due to systems being far-from-equilibrium, other methods may fail [5–8].

As depicted in Fig. 2.2, the overall process of particle-based Monte Carlo device simulation generally consists of five distinct stages: 1) initializing the system, 2) free flight and particle tracking, 3) scattering algorithm, 4) the solution of field equations, and 5) recording average quantities.

At the start of the simulation, the particle populations are initialized and distributed within the geometry under the assumptions of charge neutrality and thermal equilibrium. That is, the carriers are placed according to the defined doping levels and any surface charges that may be defined so that the overall charge is neutral. Other approaches are possible to obtain an initial particle distribution. In addition, the carriers are given an initial Maxwellian velocity distribution corresponding to the initial temperature as defined in the simulation setup. As the simulation progresses, a charge density needs to be computed from the positions of the individual charge carriers. The standard charge assignment schemes are the nearest grid point (NGP), the cloud-in-cell (CIC) scheme, or the triangular shaped cloud (TSC) scheme [54].

By using the charge density, Poisson's equation can then be solved to find the electrostatic potential and hence the electric field at each point in space. Note that Poisson's equation does not need to be solved after every free flight and scattering step, but its solution is required in time-steps short enough to resolve plasma oscillations

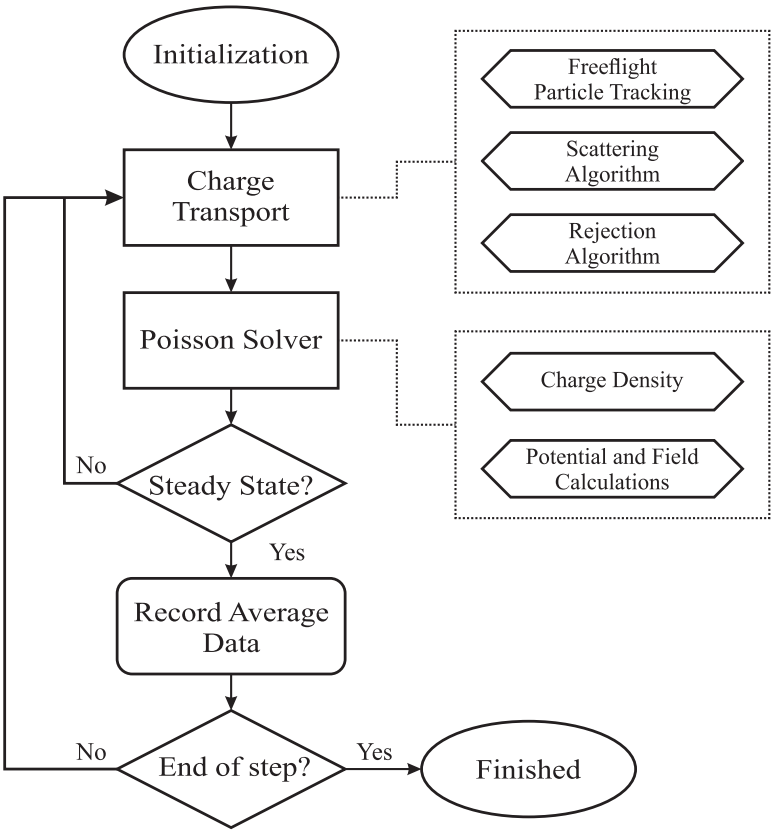


Figure 2.2: Simulation flowchart showing the stages of a general Monte Carlo device simulation.

[54].

Once the force on the carriers is known, their positions are updated using a ray tracing algorithm along with a time integration scheme and taking into account possible interactions with the layout geometry, *e.g.*, from specular or diffuse reflections against a boundary or "wall" of the device. The two stages discussed thus far account for the "drift terms" Eqs. (2.1) and (2.2).

A stochastic selection is then performed in the scattering algorithm to decide if a carrier undergoes a scattering event based on a pre-computed scattering rate. Scattering mechanisms, as a whole, represent a process through which the carrier distribution function evolves towards steady-state [55]. Since the scattering rate is a function of energy, "lucky" electrons which become highly energetic are more likely

to scatter, therefore dissipating excess energy. This stochastic model of scattering corresponds to the evaluation of the collision term in Eq. (2.2), and is generally the most expensive to evaluate. Simpler approaches based on carrier mobility values and/or relaxation time approximations are often used to model the effects of the collision term in the drift-diffusion and hydrodynamic models. This evaluation is carried out through a process in which the momentum vector, $\mathbf{k}$, of the particle is assumed to change instantaneously and independently of its previous state, *i.e.*, as a Markovian process.

The scattering mechanisms taken into account in our simulations include those with the lattice resulting in the absorption or emission of acoustic or optical phonons, scattering with ionized impurities, dislocation scattering, and impact ionization. In particular, those resulting in the absorption or emission of a phonon are taken into account using both deformation potentials and, when needed, polar scattering mechanisms for both acoustic and optical phonon modes. The calculation methods used for these individual scattering rates can be found in the literature [5, 8, 25, 27]. Here, only two separate approaches for selecting the scattering mechanisms are discussed: 1) the Ensemble Monte Carlo (EMC) method [5], and 2) the Cellular Monte Carlo (CMC) method [25].

2.6 Ensemble Monte Carlo and Cellular Monte Carlo

The EMC and CMC methods mainly differ in the manner of selecting the final state of a carrier after the occurrence of a scattering event. When a scattering event is modeled in a traditional EMC simulation, a scattering mechanism is chosen, then the First Brillouin Zone (BZ1) [26] is searched for all possible final states which are energy- and momentum-conserving. A scattering rate is then computed for each of these potential final states, and subsequently the final state is chosen among all candidates

with a stochastic selection [5]. As scattering events are extremely common, this is very computationally expensive and has a significant contribution to total simulation time.

In comparison, prior to a device simulation, the CMC method performs the scattering rate calculations from every initial momentum state, $\mathbf{k}_i$, in BZ1 to every final state, $\mathbf{k}_f$, for each scattering mechanism taken into account. This scattering information is then stored in a large look-up table which is loaded into fast computer memory (RAM) at the start of a simulation, *i.e.*, during the initialization stage of Fig. 2.2. Given the initial state of a carrier, $\mathbf{k}_i$, the selection of the final state in the original CMC algorithm is therefore reduced to the generation of a random number [25].

A scattering table look-up approach is much less expensive than the process performed in the EMC method, however the table itself can be large in size, but reasonably so with modern hardware. Hence, this represents a trade-off between computational cost and hardware requirements. The size of the look-up table in a device simulation is generally about 1 - 20 GB, although in some extreme cases the table can be larger, or more than one table is necessary, as in devices composed of different materials.

2.6.1 Adapting the CMC: Energy Conservation

In the current implementation of the CMC, the process of choosing a final state represents a highly efficient hybrid between the EMC and the CMC.

In our CMC algorithm, the full-band energy dispersion relationship is tabulated over momentum space, corresponding to the crystal wavevectors $\mathbf{k}_i$ at the center of cell i in the momentum space grid are computed and stored. The energy at every point $\mathbf{k} \neq \mathbf{k}_i$ is computed by the sum of the contributions of cell i and seven of its neighboring cells in a trilinear interpolation scheme as [56]

$$E(\mathbf{k}) = \sum_{i=0}^{7} w_i(\mathbf{k}) \cdot E_i(\mathbf{k}), \tag{2.35}$$

where w_i is a geometric weighting factor. Both energy and momentum should be conserved during a scattering event. When a scattering event occurs, finding a final state which conserves energy means finding the $\mathbf{k}'$ corresponding to a target energy, E', within the final momentum-space cell. This means that one would need to invert Eq. (2.35) which does not have a closed-form solution. Hence, a series of approximations is made, in a non-iterative manner, which seek to minimize the error made in the conservation of energy. The mathematical details of performing this inversion can be found in [57], where it is shown that the error made in energy conservation is smaller than the single precision machine in 80% of cases.

2.6.2 Adapting the CMC: Rejection

Distinct from the issue of energy conservation, is the issue of adapting the pre-computed scattering table to account for runtime conditions at each point in the real-space computational grid. For example, let's assume that we have computed a scattering table for a given material at a temperature of $T = 300$K. In the case of an electrothermal simulation, the temperature in the entire device is initialized to 300K which corresponds to a computed scattering rate, $\Gamma(300\text{K})$.

As we allow the simulation to evolve under an applied electrical bias, some amount of electron-phonon scattering events, which transfer energy from the electrons to the lattice, in turn generating heat, will occur at a rate of $\Gamma_{ph}(300\text{K})$. As this causes the temperature to increase in some parts of the device, say we find a location in the device with a temperature of $T = 350$K. Now we would need a scattering rate corresponding to this temperature, $\Gamma(350\text{K})$.

The calculation of the scattering table is a lengthy process and it cannot be per-

formed at runtime. Moreover, several scattering tables would be necessary to account for the entire range of temperatures within the device. One option is to use multiple scattering tables in a simulation corresponding to increasing temperatures, say every 25K, and then using some sort of interpolation to find scattering rates at interlying temperatures. However, each scattering table is rather large, and hence loading several tables into memory is unfeasible. A better approach uses only a single table and is based on the rejection algorithm [58, 59], which comes at only a slight cost in terms algorithmic performance [57].

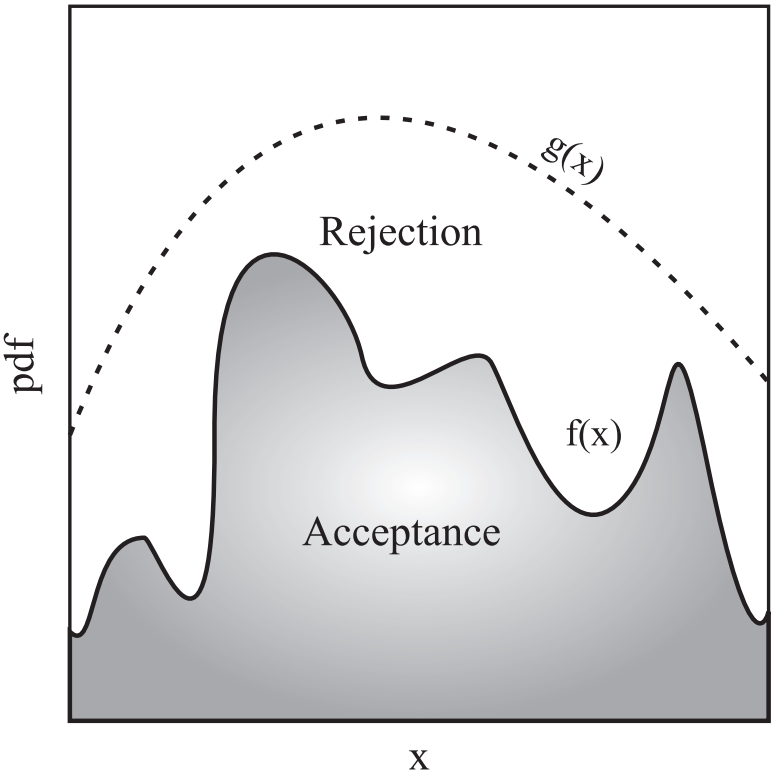


Figure 2.3: Illustration of the rejection method. The probability distribution function $f(x)$ is the desired pdf, and $g(x) > f(x)$ is a proposed pdf. The region below $f(x)$ is accepted while the region between $g(x)$ and $f(x)$ is rejected, reproducing $f(x)$ from $g(x)$ [57].

Rejection is an effective technique when the desired probability distribution, $f(x)$, is not known *a priori*. In this case, another distribution, $g(x)$, is taken as a proposed distribution subject to the condition that $g(x)$ is greater than $f(x)$ everywhere. Fig.

2.3 illustrates the rejection method being used to reproduce $f(x)$ from $g(x)$ where the regions of acceptance and rejection are labeled. Assume first that we are at a given point, a, on the x-axis, from which we know the probabilities $f(a)$ and $g(a)$. The rejection probability is then given by $P_r = f(a)/(g(a))$, which must be between 0 and 1. Subsequently, a uniform random number, R is generated, also between 0 and 1. A given event is rejected if $R > P_r$ and accepted if $R \leq P_r$ [57].

This idea intuitively maps over to the case where $g(x)$ is a scattering rate corresponding to a maximum expected temperature, and $f(x)$ is the one for a local temperature found at runtime. As long as the scattering rates monotonically increase with temperature,the condition $g(x) \geq f(x)$ is satisfied. For example, the total electron-phonon scattering rates $g(x) = \Gamma(800\text{K})$, corresponding to $T = 800\text{K}$, and $\Gamma(300\text{K})$, corresponding to $T = 300\text{K}$, are shown in Fig. 2.4, where it is also seen the condition that $\Gamma(800\text{K}) \geq \Gamma(300\text{K})$ everywhere is satisfied.

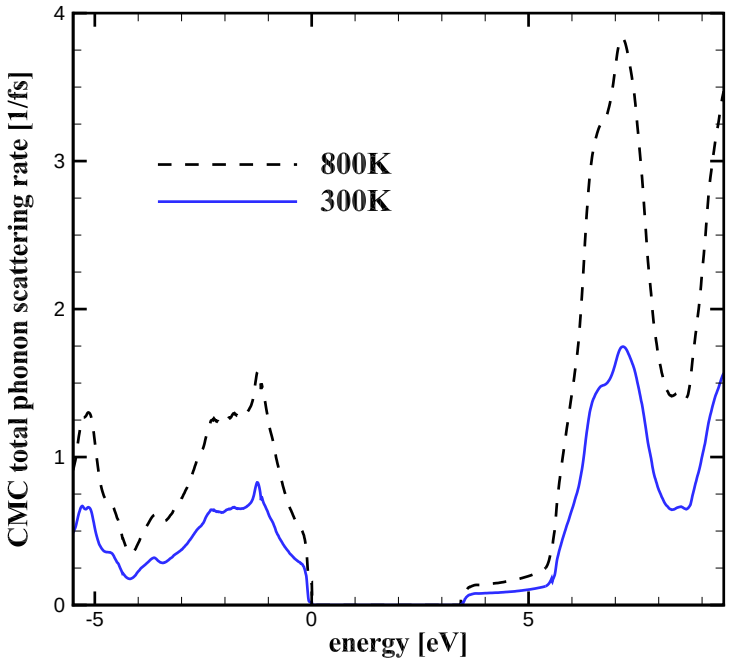


Figure 2.4: Phonon scattering rates for Wurtzite GaN used in the CMC rejection algorithm for $T = 800\text{K}$ and $T = 300\text{K}$.

Of course, in the context of the electrothermal simulations discussed in this document, the scattering rate corresponding to $f(x)$ is not known until runtime, while the scattering rate corresponding to $g(x)$ for the maximum expected temperature is pre-tabulated in the look-up table. Therefore, in these flux-based CMC simulations an electron-phonon scattering event will tentatively "occur" at the rate corresponding to $\Gamma_{ph}(\mathrm{T}_{max})$, where in this example $T_{max} = 800\mathrm{K}$, and the potential final state is chosen to be $\mathbf{k}'$. The change in the carrier wavevector is known from $\mathbf{k} = \mathbf{k}' - \mathbf{k}_i$, and hence the energy associated with the phonon, E_{ph}, can be found from Eq. (2.35).

The expected value of the phonon occupation number, assuming local thermal equilibrium, is given by the Bose-Einstein distribution

$$n\left(E_{ph}, T\right) = \left(\exp\left(\frac{E_{ph}}{k_B T} - 1\right)^{-1}\right), \tag{2.36}$$

Hence, the probability for T_{max} and T_{loc} in the electron-phonon rejection algorithm is computed as

$$P_{max} = n_{ph}\left(E'_{ph}, \mathrm{T}_{max}\right) \tag{2.37}$$

$$P_{loc} = n_{ph}\left(E'_{ph}, \mathrm{T}_{loc}\right). \tag{2.38}$$

A stochastic decision is subsequently made by generating a uniform random number in the range from zero to one, $R_{0,1}$, and the scattering event is accepted if:

$$R_{0,1} P_{max} \leq P_{loc} \tag{2.39}$$

and rejected otherwise.

It should be noted that this approach is only valid in the cases where phonons are modeled via energy balance equations. The more general case where electrons and phonons are modeled as particles is more complex, but conceptually similar [57].

Conclusion

The meaning of the moments of a distribution, especially that corresponding to its kurtosis, have been discussed as well as how they physically relate to an electron distribution. Subsequently, the moments of the electron BTE were considered which lead to the ubiquitous drift-diffusion model as well as the more accurate hydrodynamic models. These models based on lower order moments of the BTE have the advantage of being less expensive computationally, at the expense information contained in the higher-order moments of the electron distribution function being lost. Furthermore, since the BTE has an infinite number of moments it requires a closure relation in order to obtain solutions for a system made of a finite number of equations. Monte Carlo methods, on the other hand, are more computationally expensive but yield a statistically exact solution to the BTE.

The algorithmic differences between the EMC and the CMC were then discussed. The speed advantage of the CMC comes principally from pre-tabulating all of the scattering rates in the simulation. Additionally, using the rejection algorithm, the pre-tabulated scattering table can be adapted to include local runtime conditions, akin to the EMC algorithm.

Chapter 3

THEORETICAL AND COMPUTATIONAL FOUNDATIONS

3.1 Introduction

In this chapter, we will first discuss the derivation of the energy balance equation for phonons and the conditions which must satisfied relating to the continuity of the solution. Subsequently, two separate approaches will be detailed to solve the energy balance equation in a finite difference framework, including nonlinear thermal effects, based on transforming it into a Poisson-like equation. This is accomplished through the use of the Kirchhoff transformation [15, 21, 60] as well as by an iterative approach which overcomes many of the difficulties imposed by the Kirchhoff transformation, as discussed in Secs. 3.3 and 3.4, respectively. Finally, the numerical details for the computational framework used to represent the Laplacian operator and the boundary conditions will be derived.

3.2 Energy Balance Equation

The particle- and flux-based electrothermal simulator discussed here solves an energy balance equation (EBE) for phonons, derived directly from the phonon Boltzmann Transport Equation (BTE), in lieu of the Heat Transport Equation (HTE). It is easily seen in Eq. (3.10) that the final form of the EBE is the same as that of the HTE [60, 61]. However, the EBE is better suited for self-consistent coupling with the electron dynamics modeled by CMC simulations and it allows for greater accuracy by giving a separate solution for each respective phonon mode (or groups of modes, *i.e.*, all of the acoustic or optical modes grouped together).

The derivation for the EBE from the phonon BTE used here is similar to the derivation of the balance equations from the electron BTE offered in, *e.g.*, [38] or [27]. Again, the well-known electron BTE is written as

$$\frac{\partial f}{\partial t} = -\nabla_{\mathbf{r}} \cdot \left[\frac{d\mathbf{r}}{dt} f\right] - \nabla_{\mathbf{k}} \cdot \left[\frac{d\mathbf{k}}{dt} f\right] + \left.\frac{\partial f}{\partial t}\right|_{coll}. \tag{3.1}$$

However, the traditional electron BTE is not valid for phonons. This is because in its derivation Boltzmann made the assumption that all collisions were binary, *i.e.*, they involved only two particles, while this can not hold for phonons. In fact, in order to obtain a finite thermal conductivity, three-phonon processes must be included. Fortunately, Rudolf Peierls used a similar approach by assuming that the occupation number includes all possible information about the system and that phonon states are quantized [62]. The resulting phonon BTE has the same form as that for electrons, but excluding the term due to external forces

$$\frac{\partial f}{\partial t} = -\mathbf{v} \cdot \nabla_{\mathbf{r}} f + \left.\frac{\partial f}{\partial t}\right|_{coll}. \tag{3.2}$$

Similar to the moment equation for the electron BTE, the moment equation for the phonon BTE is obtained by multiplying Eq. (3.2) by some function of $\mathbf{k}$, $\phi(\mathbf{k})$, and dividing by a volume of reciprocal space, Ω. The nth moment then is given by

$$n_\phi = -\frac{1}{\Omega}\sum_{\mathbf{k}} \phi(\mathbf{k}^n)\frac{\partial f}{\partial t} - \frac{1}{\Omega}\sum_{\mathbf{k}} \phi(\mathbf{k}^n)\mathbf{v} \cdot \nabla_r f + \frac{1}{\Omega}\sum_{k} \phi(\mathbf{k}^n)\left.\frac{\partial f}{\partial t}\right|_{coll}. \tag{3.3}$$

Before working with the moments of Eq. (3.2), first we will define the total energy density of the ensemble, which will be used later in the derivation of the EBE for phonons. In order to find the total energy density of the phonon ensemble, we need to weight the energy of each state in momentum space, $E(\mathbf{k})$, by the probability

the state is occupied and sum over all possible states. Therefore, assuming that $E(\mathbf{k}) \propto \mathbf{k} \cdot \mathbf{k}$, we find the second moment of the distribution function

$$n_\phi = (\mathbf{r}, t) = \frac{1}{\Omega} \sum_{\mathbf{k}} E(\mathbf{k}) f(\mathbf{r}, \mathbf{k}, t) = W, \tag{3.4}$$

where W is used to denote the total ensemble energy density in the volume Ω, having units of [J/m^3]. It follows that the derivative of Eq. (3.4) with respect to time is

$$\frac{\partial W}{\partial t} = \frac{1}{\Omega} \sum_{\mathbf{k}} E(\mathbf{k}) \frac{\partial f(\mathbf{r}, \mathbf{k}, t)}{\partial t}, \tag{3.5}$$

representing the rate of change of the ensemble energy density in the volume Ω.

Now, letting $\phi(\mathbf{k}) = E(\mathbf{k})$ in Eq. (3.3), yields

$$\frac{1}{\Omega} \sum_{\mathbf{k}} E(\mathbf{k}) \frac{\partial f}{\partial t} + \frac{1}{\Omega} \sum_{\mathbf{k}} E(\mathbf{k}) \mathbf{v} \cdot \nabla_r f = \frac{1}{\Omega} \sum_{\mathbf{k}} \left. \frac{\partial f}{\partial t} \right|_{coll}, \tag{3.6}$$

where the first term is simply $\partial W / \partial t$, while the second term is the energy flux and can be denoted by

$$\mathbf{F}_W (\mathbf{r}, t) = \frac{1}{\Omega} \sum_{\mathbf{k}} E(\mathbf{k}) \mathbf{v} \cdot \nabla_r f(\mathbf{r}, \mathbf{k}, t). \tag{3.7}$$

Here, we will approximate this term by using Fourier's law [63] for the heat flux, $\mathbf{F}_W = -\kappa \nabla T$, where κ is the thermal conductivity of the material, ∇T is the temperature gradient, and $\mathbf{F}_W$ is the heat flux. Finally, the collision term on the RHS can be split into two distinct processes which contribute to changes in the phonon energy density through changes in the phonon populations themselves: 1) the increase (or decrease) in the phonon energy density due to an electron-phonon scattering event where phonons are emitted or absorbed, and 2) the change due to phonon-phonon interactions causing the recombination of acoustic phonons to a high energy optical phonon or the decay of optical phonons to low energy acoustic phonons. In other

words, they denote the heat generation rates, for the particular phonon mode due to interaction with electrons, $\partial W/\partial t|_{e-p}$, and other phonons, $\partial W/\partial t|_{p-p}$, respectively. Using these substitutions in Eq. (3.6), we can write the time-dependent EBE for phonons as

$$\frac{\partial W}{\partial t} = -\nabla \cdot (-\kappa \nabla T) + \left.\frac{\partial W}{\partial t}\right|_{e-p} + \left.\frac{\partial W}{\partial t}\right|_{p-p}. \tag{3.8}$$

In order to account for the different phonon modes, one equation in the form of Eq. (3.8) is obtained for each phonon mode, or group of modes, being considered. Generalizing the time-dependent EBE for phonons to a system of equations, where μ denotes the particular mode, we therefore have

$$\frac{\partial W_\mu}{\delta t} = -\nabla \cdot \mathbf{F}_\mu + \left.\frac{\partial W_\mu}{\delta t}\right|_{e-p} + \left.\frac{\partial W_\mu}{\delta t}\right|_{p-p}. \tag{3.9}$$

Under steady-state conditions the time derivative on the LHS is zero, and if P_μ is used to represent the sum of the electron-phonon and phonon-phonon contributions, then Eq. (3.9) becomes

$$\nabla \cdot (\kappa_\mu (\mathbf{r}, \mathrm{T}) \nabla \mathrm{T}) = -\left(\left.\frac{\partial W_\mu}{\delta t}\right|_{e-p} + \left.\frac{\partial W_\mu}{\delta t}\right|_{p-p} \right) = -P_\mu. \tag{3.10}$$

One should note that ballistic heat transfer effects are not explicitly included in our model as a consequence of using Fourier's law to approximate the heat flux. This is because Fourier's law is diffusive in nature, implying the presence of scattering. However, it has been suggested that ballistic heat transfer effects can in fact be captured either through a choice of boundary conditions [64] or by solving a generalized form of Fourier's law derived using the Peierls-Boltzmann Equation [23].

Equation (3.10) is subject to the conditions that both the temperature (in the absence of a nonideal interface) and the heat flux in the normal direction are continuous

across a boundary [15], Σ. That is,

$$\mathrm{T} \text{ continuous across } \Sigma \tag{3.11}$$

$$\kappa_1(\mathrm{T})\hat{n} \cdot \nabla \mathrm{T}|_\Sigma = \kappa_2(\mathrm{T})\hat{n} \cdot \nabla \mathrm{T}|_\Sigma, \tag{3.12}$$

which must also be satisfied in the derivation of the computational grid as discussed in Sec. 3.5.

Equations of the form of Eq. (3.10) can be solved directly with a nonlinear finite element analysis (FEA) approach, while such an equation is non-trivial to solve using finite difference methods. Although, in principle, a MC device simulator could include a field solver using finite element methods, the process of particle tracking would become more complicated when using non-rectangular grids, as is often the case in FEA simulation.

An efficient finite-difference solver for the Poisson equation, a linear elliptical PDE, is present in any MC electrical device simulator. Hence, manipulating Eq. (3.10) into the form of a linear elliptical PDE would allow for a more straightforward implementation including temperature-dependent effects within the finite difference framework. Simply put, if we are able to pull $\kappa_\mu(\mathbf{r}, \mathrm{T})$ outside of the divergence operator on the LHS of Eq. (3.10), we can then work with a Poisson-like linear equation of the form:

$$\kappa_\mu \nabla^2 \mathrm{T} = -P_\mu, \tag{3.13}$$

The main issue in performing this manipulation is the dependence of $\kappa_\mu(\mathbf{r}, \mathrm{T})$ on both position and temperature. For the former, we assume that the thermal conductivity is piecewise constant. In other words, the thermal conductivity is constant within any particular cell, C, but allowed to vary when moving from one cell to an-

Table 3.1: Thermal conductivity α parameter values for selected materials [69]

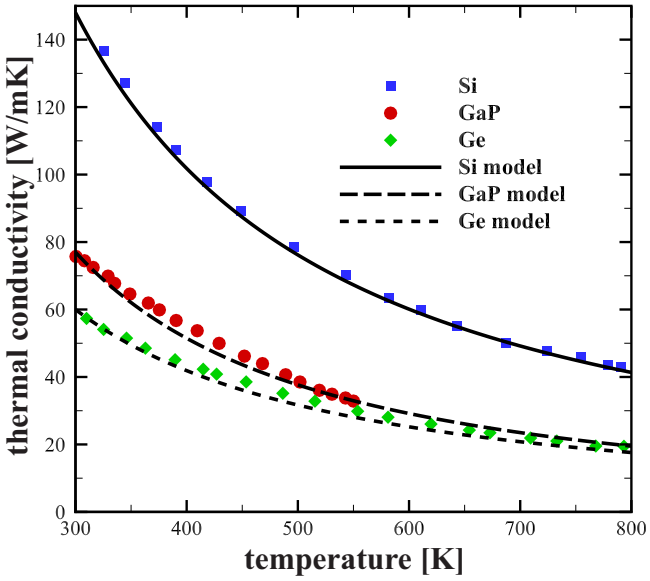
Material	k_{300} [W/K m]	α
Si	148	-1.65
Ge	60	-1.25
GaAs	46	-1.25
AlAs	80	-1.37
InAs	27.3	-1.1
InP	68	-1.4
GaP	77	-1.4

other. This is perfectly analogous to the approach used for the dielectric constant in the solution of Poisson's Equation. Hence, we will express this restricted position dependency with the cell index C rather than via a full functional dependence on the position vector $\mathbf{r}$. For the latter case of the temperature dependence, two separate approaches have been taken to perform this manipulation and are discussed in Sec. 3.3 and Sec. 3.4, respectively: 1) using the Kirchhoff transformation, and 2) an iterative approach which overcomes many of the restrictions imposed by the Kirchhoff transformation.

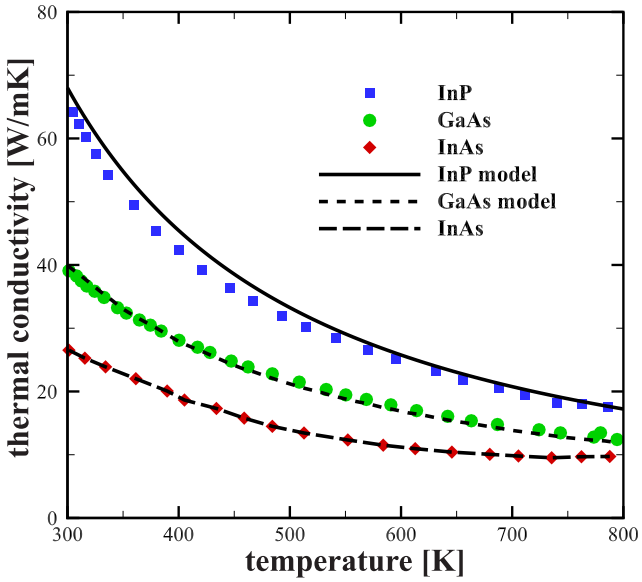
As for the functional form used for the temperature dependence of the thermal conductivity itself, it has been found that a simple power law provides a good fit to experimental measurements for a large number of materials [15, 65–69]

$$\kappa_{\mu,C}\left(\mathrm{T}\right) = \kappa_{ref}\left(\frac{\mathrm{T}}{\mathrm{T}_{ref}}\right)^{\alpha}, \tag{3.14}$$

where k_{ref} is the independently known thermal conductivity at the reference temperature T_{ref}, often taken as 300K. The values used for k_{300} and α taken from [69] are shown in Table 3.1, while a comparison between experimentally measured thermal



(a)



(b)

Figure 3.1: Experimental results compared with the power law model for temperature dependent thermal conductivities of Si, GaP, Ge, InP, GaAs, and InAs [15, 65–69].

conductivities and Eq. (3.14) is shown in Fig. 3.1. Due to the good agreement with experimental data seen here for the materials of interest, the power law fit has been

adopted as the general form for the material thermal conductivity throughout this work.

3.3 Kirchhoff Transformation

The Kirchhoff transformation is defined as [15, 21, 60]

$$\Theta_{\mu,C}\left(\mathrm{T}\right) = \mathrm{T}_{ref} + \frac{1}{\kappa_{\mu,C}\left(\mathrm{T}_{ref}\right)} \int_{\mathrm{T}_{ref}}^{\mathrm{T}} \kappa_{\mu,C}\left(\tau\right) \partial\tau, \tag{3.15}$$

where Θ denotes a new "apparent" temperature which implicitly includes the temperature dependence of the thermal conductivity. T_{ref} is a chosen reference temperature, often 300K, at which the thermal conductivity value, $\kappa_{\mu,C}\left(\mathrm{T}_{ref}\right)$, is independently known.

Using the apparent temperature, Θ, we can now linearize the EBE in the form of a Poisson-like equation as

$$\nabla^2 \Theta_{\mu,C} = -\frac{P_{\mu}\left(\mathbf{r}\right)}{\kappa_{\mu,C}\left(\mathrm{T}_{ref}\right)}. \tag{3.16}$$

However, a few significant restrictions are imposed by the Kirchhoff transformation. First, Eq. (3.16) requires that both the temperature and its derivative in the normal direction, *i.e.*, the heat flux, are continuous across a boundary [15]. These are in fact the same conditions of the standard heat equation, however they must be transformed as well. The heat flux is invariant under the transformation, and so the second condition is satisfied [15]. For the temperature continuity condition, say from the left (L) side of an interface to the right (R), from Eq. (3.15)

$$\Theta_L - \Theta_R = \frac{1}{\kappa_L\left(\mathrm{T}_{ref}\right)} \int_{\mathrm{T}_{ref}}^{\mathrm{T}_L} \kappa_L\left(\tau\right) \partial\tau - \frac{1}{\kappa_R\left(\mathrm{T}_{ref}\right)} \int_{\mathrm{T}_{ref}}^{\mathrm{T}_R} \kappa_R\left(\tau\right) \partial\tau. \tag{3.17}$$

In order for the apparent temperature to be continuous across a boundary, $\Theta_L - \Theta_R = 0$, the two materials must have the same functional form for the thermal conductivity. If we use the power law fit discussed previously, this means that the two relationships can differ only by the multiplying constant κ_{ref}, while the exponent α must be the same for both materials.

However, this clearly introduces a further limitation if the materials have a differing temperature dependence. Indeed, while this approximation may be reasonable for most semiconductor materials as the values of α generally lie between -1.25 and -1.4 [69], metals and oxides display vastly different temperature dependent behavior which is relevant at their respective semiconductor interfaces.

Secondly, functions representing Dirichlet and simple Neumann boundary conditions, corresponding to a constant temperature or heat flux, respectively, are linear under the Kirchhoff transformation. In contrast, convective boundary conditions become nonlinear under the Kirchhoff transformation [21, 60]. Convective boundary conditions describe thermal contacts that have an inherent thermal resistance, R_{th}, or a finite heat transfer coefficient, H, where $H = 1/R_{th}$. Clearly, as $R_{th} \to 0$, $H \to \infty$ and the perfect thermal contact of a Dirichlet boundary condition is recovered. In addition, convective conditions can also occur at material interfaces, which often introduce a resistance to the heat passing from one material to the other.

Convective boundary conditions are expressed as

$$-\kappa \nabla \mathrm{T} = H \left(\mathrm{T}_L - \mathrm{T}_R \right), \tag{3.18}$$

which relates the heat flux approaching the interface, $-\kappa \nabla \mathrm{T}$, with the heat transport coefficient and the temperature jump across the interface, $(\mathrm{T}_L - \mathrm{T}_R)$.

Since the Kirchhoff transformation is nonlinear for any nonlinear functional form of the thermal conductivity, it is generally not true that $H\left(\Theta_L - \Theta_R\right) = H\left(\mathrm{T}_L - \mathrm{T}_R\right)$.

Bagnall *et al.* [21] overcame this non-linearity in large-scale thermal spreading problems by choosing the reference temperature and thermal conductivity in the Kirchhoff transformation such that the convective boundary condition is roughly linear, *i.e.*, $H\left(\Theta_L - \Theta_R\right) \approx H\left(\mathrm{T}_L - \mathrm{T}_R\right)$. However, this requires some *a priori* knowledge of the temperature distribution expected at the convective boundary in order to choose T_{ref} around which to linearize the transformation, so as to reduce the error due to the nonlinearity. This approach poses problems for coupled electrothermal device simulations where the heat generation is not necessarily known ahead of time. In addition, Bagnall demonstrated that in cases where the temperature distribution along the convective boundary is less uniform this approach becomes less accurate. This is simply because a single temperature is chosen around which to linearize the Kirchhoff transformation, and calculations at temperatures far away from the chosen value become more inaccurate, again due to the nonlinearity.

3.4 Iterative Method

To overcome the discussed restrictions related to the Kirchhoff transformation, a classical iterative approach to solving the EBE has also been implemented. This takes advantage of the fact that the thermal conductivity is allowed to vary, in a piece-wise constant way, from one cell to another to incorporate the temperature-dependent values, and a sequence of constant-conductivity problems is solved where the thermal conductivity in each cell is updated with each new solution according to the local temperature. In other words, we solve a system of the form:

$$\nabla^2 \mathrm{T}^i_{\mu,C} = -\frac{P_\mu\left(\mathbf{r}\right)}{\kappa_{\mu,C}\left(\mathrm{T}^{i-1}\right)}. \tag{3.19}$$

where the new temperature T^i, for iteration i, on the LHS is obtained using the temperature-dependent thermal conductivity in each cell obtained from the previous

temperature solution from iteration $i-1$. The temperatures on the LHS are solved self-consistently, and the thermal conductivities are subsequently updated, until the change for both the acoustic and optical temperature between two successive iterations has become sufficiently small and the procedure has converged, as depicted in Fig. 3.2.

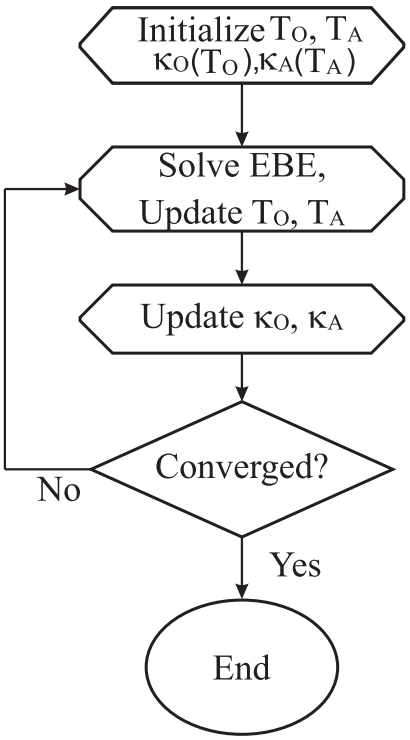


Figure 3.2: Flowchart for the self-consistent thermal solution using the iterative method.

This iterative approach uses the same power law fit previously discussed for the thermal conductivity, however it is not subject to the requirement that the exponent α in the temperature dependence be the same for different materials, as that requirement is due to the Kirchhoff transformation and not Eq. (3.10) itself.

In addition, while the convective boundary condition in Eq. (3.18) is linear in and of itself, it becomes non-linear under the Kirchhoff transformation as mentioned previously. However, as seen in Sec. 3.5, convective boundary conditions are accounted

for in a straightforward manner under the iterative approach, *i.e.*, in the same manner Dirichlet and Neumann conditions are handled. Therefore, this approach allows us to accurately simulate heat transport across interfaces possessing non-ideal effects, including realistic convective heat sinks and thermally-resistive material interfaces.

This is significant for devices such as a High Electron Mobility Transistor (HEMT) which is studied in a later section, where material interfaces are used to create a 2-D electron gas (2DEG) conduction channel, as well as in layouts attempting to use a high thermal conductivity material to transport heat away from the active device, where the interface can have an associated thermal resistance.

It should be noted, however, that the thermal conductivity is likely to assume different values in every computational cell in Eq. (3.19), and so classical multigrid methods are likely to become very inefficient. One possible approach to mitigate this if multigrid methods are desired is to attempt to average the thermal conductivity, assuming the values fall within a reasonable range, between cell which are merged together, however this could be dubious. Fortunately, in an electrothermal simulation the solution to Poisson's equation, where the iterative method is not used, is required orders of magnitude more times than that of the phonon EBE.

3.5 Computational Difference Scheme

The computational grid developed for the solution of Eq. (3.10) must satisfy the heat flux condition in Eq. 3.12. A schematic of a two-dimensional grid showing the nearest neighbors and denoting the relevant distances between points of the grid is shown in Fig. 3.3.

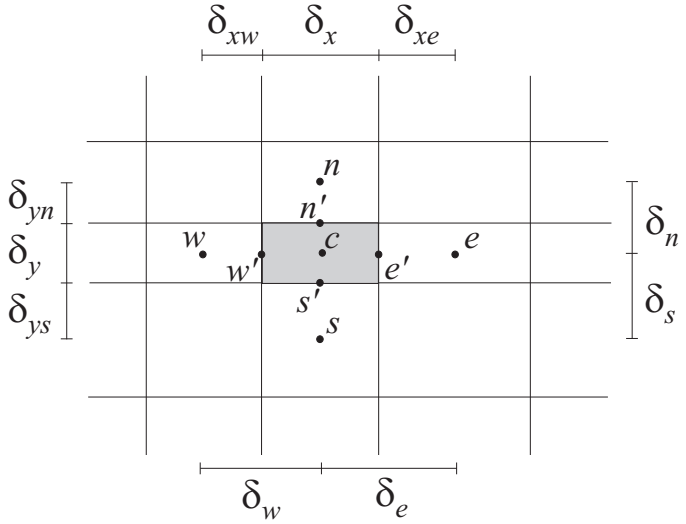


Figure 3.3: A schematic of the computational grid. Points labeled with n,e,w,s are neighbors of the point labeled c. Intermediate points on the cell boundary are used for interpolations, and the various δ's are distances used in the calculation of the Laplacian coefficients.

For a given cell inside the computational domain, labeled c, the northern, eastern, western, and southern neighbors are labeled as n, e, w, and s, respectively. The distances δ_x and δ_y are the x and y dimensions of cell c. The distances δ_n, δ_s, δ_w, and δ_e are the respective distances from the center of cell c to the center of the northern, southern, western, and eastern neighbor cells, respectively. These distances cannot be assumed to be equivalent the δ_x and δ_y for inhomogenous grids. Lastly, the $\delta_{xw,xe}$, $\delta_{yn,ys}$ are the distances from the center of the neighbor cell to the nearest boundary of cell c, which are needed to perform interpolations from cell-to-cell.

The Laplacian operator in Eq. (3.13) can be approximated by the difference

relation

$$\nabla^2 \mathrm{T}_c \approx a_n \mathrm{T}_n + a_e \mathrm{T}_e + a_s \mathrm{T}_s + a_w \mathrm{T}_w + a_c \mathrm{T}_c, + O(\Delta_x \Delta_y)^m, \tag{3.20}$$

where the final term is an error term dependent upon the chosen grid sizing. For uniformly spaced grids $m = 3$, while for inhomogenous grids $m < 3$ [70].

In the most general case, considering an inhomogenous grid with a piece-wise constant thermal conductivity, the coefficients in Eq. (3.20) are given by

$$\begin{aligned}
a_w &= \left[\frac{2}{\delta_w\left(\delta_e + \delta_w\right)}\right]\left[1 + \frac{\delta_{xw}}{\delta_x/2}\right]\left[\frac{\kappa_w \delta_x/2}{\kappa_c \delta_{xw} + \kappa_w \delta_x/2}\right] \\
a_e &= \left[\frac{2}{\delta_e\left(\delta_e + \delta_w\right)}\right]\left[1 + \frac{\delta_{xe}}{\delta_x/2}\right]\left[\frac{\kappa_e \delta_x/2}{\kappa_c \delta_{xe} + \kappa_e \delta_x/2}\right] \\
a_n &= \left[\frac{2}{\delta_n\left(\delta_n + \delta_s\right)}\right]\left[1 + \frac{\delta_{yn}}{\delta_y/2}\right]\left[\frac{\kappa_n \delta_y/2}{\kappa_c \delta_{yn} + \kappa_n \delta_y/2}\right] \\
a_s &= \left[\frac{2}{\delta_s\left(\delta_n + \delta_s\right)}\right]\left[1 + \frac{\delta_{ys}}{\delta_y/2}\right]\left[\frac{\kappa_s \delta_y/2}{\kappa_c \delta_{ys} + \kappa_s \delta_y/2}\right] \\
a_c &= -\frac{2}{\delta_e \delta_w} - \frac{2}{\delta_n \delta_s} + p_w + p_e + p_s + p_n,
\end{aligned} \tag{3.21}$$

where the values of the p_i terms in the center coefficient are further given by

$$\begin{aligned}
p_w &= \left[\left(1 + \frac{\delta_{xw}}{\delta_x/2}\right)\left(\frac{\kappa_c \delta_{xw}}{\kappa_c \delta_{xw} + \kappa_w \delta_x/2}\right) - \frac{\delta_{xw}}{\delta_x/2}\right] \\
p_e &= \left[\left(1 + \frac{\delta_{xe}}{\delta_x/2}\right)\left(\frac{\kappa_c \delta_{xe}}{\kappa_c \delta_{xe} + \kappa_e \delta_x/2}\right) - \frac{\delta_{xe}}{\delta_x/2}\right] \\
p_s &= \left[\left(1 + \frac{\delta_{ys}}{\delta_y/2}\right)\left(\frac{\kappa_c \delta_{ys}}{\kappa_c \delta_{ys} + \kappa_s \delta_y/2}\right) - \frac{\delta_{ys}}{\delta_y/2}\right] \\
p_n &= \left[\left(1 + \frac{\delta_{yn}}{\delta_y/2}\right)\left(\frac{\kappa_c \delta_{yn}}{\kappa_c \delta_{yn} + \kappa_n \delta_y/2}\right) - \frac{\delta_{yn}}{\delta_y/2}\right].
\end{aligned} \tag{3.22}$$

In the case of the Kirchhoff transformation, the temperature dependence of the thermal conductivity is absorbed into the temperature variable Θ, and so the κ values in the above coefficients are simply those at the reference temperature. Hence, the Laplace operator, and all of the associated coefficients, only need be set up once at

the beginning of the simulation and stored as they will be constant throughout the simulation. On the other hand, in the iterative approach the various values of κ in Eqs. (3.21) and (3.22) are updated only after each new temperature solution is found.

3.6 Boundary Conditions

Two kinds of boundary conditions are considered in this work with the context of device simulation in mind. These are the Dirichlet boundary condition [61, 71] of a prescribed constant temperature (Eq. (3.23), and the convective (or Robin) boundary condition [61, 71] which is related to a prescribed non-zero Neumann condition on the heat flux (Eq. (3.25)).

The Dirichlet condition on the temperature of the boundary cells, T_b, is expressed as:

$$\mathrm{T}_b = \mathrm{T}. \tag{3.23}$$

To implement the Dirichlet condition on the cell c, the coefficients of Eq. (3.21) reduce in the normal way such that

$$\begin{aligned} a_w &= 0 \\ a_e &= 0 \\ a_n &= 0 \\ a_s &= 0 \\ a_c &= 1. \end{aligned} \tag{3.24}$$

This reduces Eq. (3.20) to $\mathrm{T}_c = \mathrm{T}_b$ and has the effect of fixing the temperature in the Dirichlet cells where the forcing function is set to the prescribed value T_b.

The convective boundary condition is based on Newton's law of cooling [60, 61, 71],

and given by

$$-\kappa \nabla \mathrm{T} = H(\mathrm{T}_b - \mathrm{T}_\infty), \tag{3.25}$$

where T_b is again the interior temperature at the boundary cell, T_∞ is the ambient temperature specified outside of the boundary, which is assumed to be constant, and H is a heat transfer coefficient. In particular, the heat transfer coefficient is expressed in $\frac{W}{m^2 K}$ and is the inverse of an associated thermal resistance, R_{TH}. The problem modeled by Eq. (3.10) with the convective boundary condition of Eq. (3.25) is well-posed as long as the ratio of thermal conductivity and heat transfer coefficient is positive, $\frac{\kappa}{h} > 0$ [71], which represents any physically realistic case, as either a negative thermal conductivity or a negative heat transport coefficient if nonsensical from a physical perspective.

3.6.1 Interface Conditions

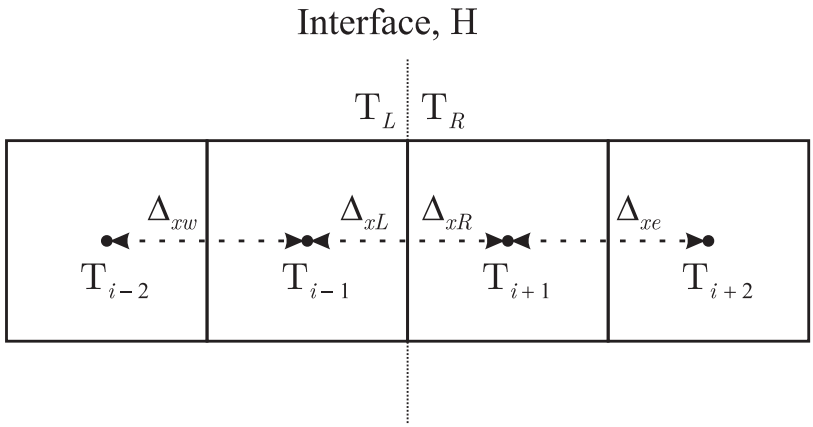


Figure 3.4: Diagram of a 1-D material interface.

To implement the material interface condition, we follow the work of Hickson *et al.* [72]. However, here we are interested in solving the steady-state problem rather than the time-dependent problem studied in [72]. Furthermore, here we use a first order accurate expression for the interface coefficients as the second order expressions

are much more cumbersome [72], and, since we generally work with non-homogeneous grids, we expect only first order accuracy in the Laplacian. Considering only the 1-D case on the x-axis, Fig. 3.4 depicts the cells nearest an interface with an associated heat transfer coefficient H. The temperatures at the edge of each side of the boundary are given by T_L and T_R, while the temperatures at the central cell points are denoted $\mathrm{T}_{i\pm n}$.

The second derivatives on each side of the interface are written as follows [70]:

$$\frac{\partial \mathrm{T}_{i-1}}{\partial x^2} \approx \frac{\Delta_{xw}\mathrm{T}_L - (\Delta_{xL}+\Delta_{xw})\,\mathrm{T}_{i-1}}{\Delta_{xL}\Delta_{xw}(\Delta_{xL}+\Delta_{xw})}, \tag{3.26}$$

$$\frac{\partial \mathrm{T}_{i+1}}{\partial x^2} \approx \frac{\Delta_{xe}\mathrm{T}_R - (\Delta_{xR}+\Delta_{xe})\,\mathrm{T}_{i+1}}{\Delta_{xL}\Delta_{xe}(\Delta_{xL}+\Delta_{xe})}, \tag{3.27}$$

where T_{i-1} denotes the nearest grid point to the left of the interface, T_{i+1} the nearest grid point to the right of the interface, Δ_{xL} the distance between the point $i-1$ and the interface, Δ_{xR} the distance between the point $i+1$ and the interface, and lastly Δ_{xw} the distance between the points $i-1$ and $i-2$ to the left, Δ_{xe} the distance between points $i+1$ and $i+2$ to the right. In Eqs. (3.26) and (3.27) we have 4 unknowns: T_{i-1}, T_{i+1}, T_L, and T_R. The temperatures at the grid points everywhere in the simulation, such as T_{i-1} and T_{i+1} are independently solved for using numerical methods such as successive over-relaxation (SOR), and so we need to find expressions to substitute for T_L and T_R. To find these expressions, we use the Taylor series around T_{i-1} and T_{i+1}, the fact that the flux must be continuous on each side of the interface, and the boundary condition itself. The Taylor series around T_{i-1} and T_{i+1} are given by

$$\mathrm{T}_{i-1} \approx \mathrm{T}_L - \Delta_{xL}\frac{\partial \mathrm{T}_L}{\partial x} \tag{3.28}$$

$$\mathrm{T}_{i+1} \approx \mathrm{T}_R + \Delta_{xR}\frac{\partial \mathrm{T}_R}{\partial x}, \tag{3.29}$$

while the flux continuity and boundary conditions are

$$\kappa_L \frac{\partial \mathrm{T}_{i-1}}{\partial x} = \kappa_R \frac{\partial \mathrm{T}_{i+1}}{\partial x} \tag{3.30}$$

$$\kappa_L \frac{\partial \mathrm{T}_{i-1}}{\partial x} = H\left(\mathrm{T}_R - \mathrm{T}_L\right). \tag{3.31}$$

Using Eqs. (3.28 - 3.31), T_L and T_R can now be obtained as

$$\mathrm{T}_L = \frac{\left(H\Delta_{xR} + \kappa_R\right)\kappa_L}{\left(H\Delta_{xR} + \kappa_R\right)\kappa_L + H\Delta_{xL}\kappa_R}\mathrm{T}_{i-1} + \frac{H\Delta_{xL}\kappa_R}{\left(H\Delta_{xR} + \kappa_R\right)\kappa_L + H\Delta_{xL}\kappa_R}\mathrm{T}_{i+1} \tag{3.32}$$

$$\mathrm{T}_R = \frac{H\Delta_{xR}\kappa_L}{\left(H\Delta_{xR} + \kappa_R\right)\kappa_L + H\Delta_{xL}\kappa_R}\mathrm{T}_{i-1} + \frac{\left(H\Delta_{xL} + \kappa_L\right)\kappa_R}{\left(H\Delta_{xR} + \kappa_R\right)\kappa_L + H\Delta_{xL}\kappa_R}\mathrm{T}_{i+1} \tag{3.33}$$

which can be substituted back into Eqs. (3.26) and (3.27) to obtain the Laplacian in the neighborhood of the interface. On a uniform grid where $\Delta_{xR} = \Delta_{xL} = \Delta_x$, and letting $H \to \infty$ it can be shown that the Laplacian reduces to the straightforward case with no interface resistance.

To verify the robustness of the proposed iterative approach in the presence of interfaces, a 10-layer structure similar to that in [72] has been simulated. The boundary conditions are modified from [72] as we are interested in the steady-state problem, and hence we have imposed a 400K Dirichlet condition on the left and a 300K Dirichlet condition on the right of Fig. 3.4. No heat generation is present inside the device. The alternating layers are labeled and denoted by the light dashed lines. The parameters used are $\kappa_A = 1$ W/mK, $\kappa_B = 0.1$ W/mK, and the heat transport coefficient at each interface is $H_i = 5 \cdot 10^8$ W/m^{2}K, which is a reasonable value for a semiconductor-semiconductor interface.

Interestingly, we see that the temperature-dependent solution here follows that of the constant conductivity quite closely, which is reasonable as each layer is only 1 nm and the total size of the simulation is 10 nm. The Kirchhoff transformation solu-

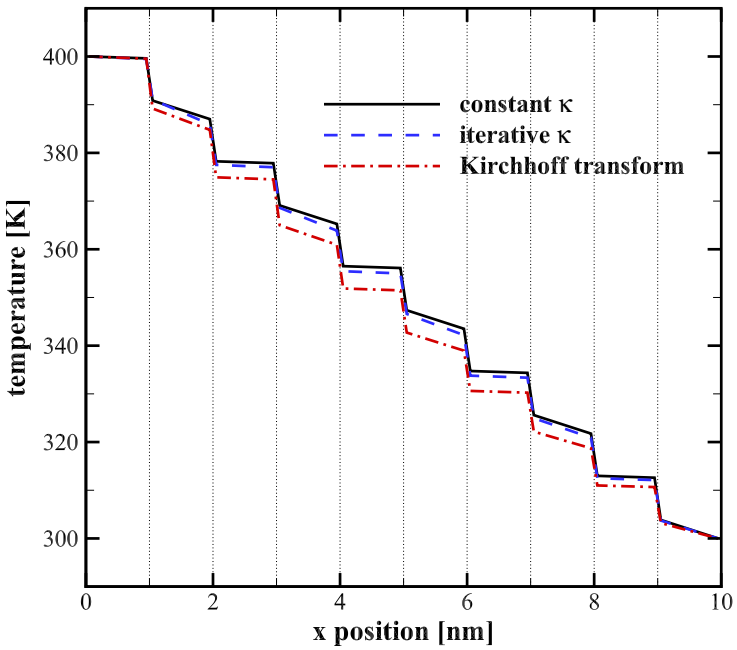


Figure 3.5: A 10-layer structure simulated using 1) a constant thermal conductivity, 2) the iterative algorithm, and 3) the Kirchhoff transformation with the thermal conductivity taken at 300K.

tion, however, becomes more inaccurate the further it moves away from the Dirichlet conditions due to the non-linearity.

Conclusion

The EBE equation for phonons is obtained from the $n = 2$ moment of the phonon BTE, similarly to the EBE for electrons. As any method based on moments of the BTE must use a closure relation, Fourier's law is then used to close the resulting system of equations.

When considering the temperature dependent thermal conductivity, it is found that a relatively simple power law fit shows good agreement with experimental results. The conductivity-vs-temperature relationship can then be used to incorporate temperature dependent effects through either the Kirchhoff transformation or an iterative method. However, the Kirchhoff transformation is ill-suited, compared to the iterative method, to take into account convective boundary conditions which arise from imperfect material interfaces or heat sinks.

On the other hand, in the proposed iterative algorithm convective boundary conditions are implemented in a relatively straightforward way, *i.e.*, in the same manner as a Dirichlet or Neumann condition as coefficients of the Laplacian operator.

Chapter 4

BENCHMARK SIMULATIONS

4.1 Introduction

To verify the computational framework developed in Ch. 3, the simulation tests performed in [21, 73] have been reproduced here using both the Kirchhoff Transformation of Sec. 3.3 and the proposed iterative approach of Sec. 3.4. The results of each approach are compared with the full nonlinear FEA solver in the commercial MATLAB ® 2018b release [74]. As mentioned previously, a full nonlinear FEA solver is capable of solving the full nonlinear EBE in Eq. (3.10) directly.

All of the simulations performed in this chapter have the heat generation set either at a constant rate, or at zero. First, we consider a simple 1-D heat transfer problem using a 100 μm thick piece of silicon with a constant heat generation rate being pumped into the top layer of the material and a convective heat sink located on the bottom layer. Next, we consider a more complicated 2-D heat transfer problem. Here, rather than a constant heat flux across the entire top boundary of the material, a single 2-nm wide heat source is centered at $x = 250\mu$m on the top of the 500 μm long sample. Finally, the most complicated test case of a 11-finger GaN HEMT power amplifier is modeled in Sec. 4.4.

4.2 One-Dimensional Simulations

As an initial test of both the Kirchhoff transformation and the proposed iterative approach, the simple layout of Fig. 4.1 is simulated [21]. In this simplest case of a 1-D simulation, we use the material parameters of a 500 μm long by 100 μm thick

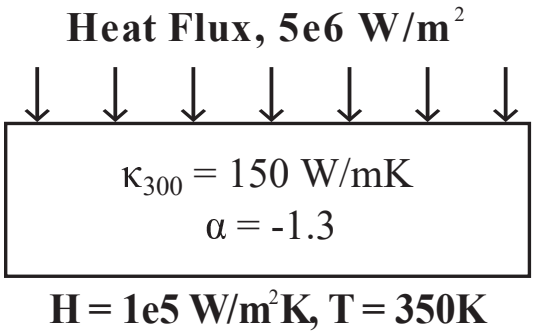


Figure 4.1: Schematic of the 1-D heat transfer layout. A uniform inward heat flux of $H = 5 \cdot 10^6$ W/m^2 is applied uniformly across the top layer

piece of silicon. The thermal conductivity of silicon at 300K is taken as 150 W/m·K, and the exponent in the temperature power law of Eq. (3.14) is taken as $\alpha = -1.3$ [21, 69]. In other words, the temperature dependent thermal conductivity is taken as

$$\kappa\left(\mathrm{T}\right) = 150\left(\frac{\mathrm{T}}{300}\right)^{-1.3}. \tag{4.1}$$

Here, we are simulating only a single bulk temperature as would be measured by a thermometer, and so the subscript μ has been dropped in Eq. (4.1). The boundary conditions used, also indicated in Fig. 4.1, are a constant inward heat flux all along the top boundary of $5{\cdot}10^6$ W/m^2, while a convective boundary is set along the bottom of the material characterized by a heat transfer coefficient of $H = 10^5$ W/m^{2}K and an ambient temperature of $\mathrm{T}_\infty = 350$K. In other words, the boundary condition is

$$-\kappa\nabla\mathrm{T} = 10^5\left(\mathrm{T}_b - 350\mathrm{K}\right). \tag{4.2}$$

The Kirchhoff transformation approach was used to solve Eq. (3.10) with the convective boundary condition of Eq. (4.2) in two cases: 1) effectively linearizing the boundary condition under the Kirchhoff transformation by choosing the reference temperature to be that of the average temperature expected at the boundary, and

2) without linearizing the transformation but instead using $\mathrm{T}_{ref} = 300\mathrm{K}$ and $\kappa_{ref} = \kappa_{300} = 150$ W/mK.

In the simple 1-D case, the temperature expected at the convective boundary can be determined analytically. Since Eq. (3.10) is subject to the condition that the flux is continuous and there is no heat generation inside the material, then the flux leaving the convective boundary at $y = 0$ must be the same as the flux entering the top boundary at $y = L$, *i.e.*, $\Phi = -\kappa\nabla\mathrm{T}$ where Φ denotes the inwards flux at the top boundary. The temperature at the boundary is then [21]

$$\mathrm{T}_b = \frac{\Phi}{H} + \mathrm{T}_\infty. \tag{4.3}$$

If the reference temperature in Eq. (3.15) is chosen to be T_b, then it immediately follows that the apparent temperature, Θ, is also equal to T_b. The nonlinear EBE can then be solved analytically using the Kirchhoff transformation. Because the heat generation rate is zero, Eq. (3.15) becomes:

$$-\kappa_{\mu,C}\left(\mathrm{T}_{ref}\right)\nabla^2\Theta_{\mu,C} = 0, \tag{4.4}$$

which has the solution $\Theta = C_1 y + C_2$. If, as in [21], we then consider the convective boundary condition at at $y = 0$, to be a Dirichlet condition prescribed by Eq. (4.2), then the solution is

$$\Theta\left(y\right) = \frac{\Phi}{\kappa_{ref}} y + \mathrm{T}_b. \tag{4.5}$$

As seen in Fig. 4.2, the solution for the Kirchhoff transformation agrees exactly with the full nonlinear FEA solution when the reference temperature in the Kirchhoff transformation is chosen in this manner, such that the transformation of the boundary condition is effectively linearized around the convective boundary. One should note

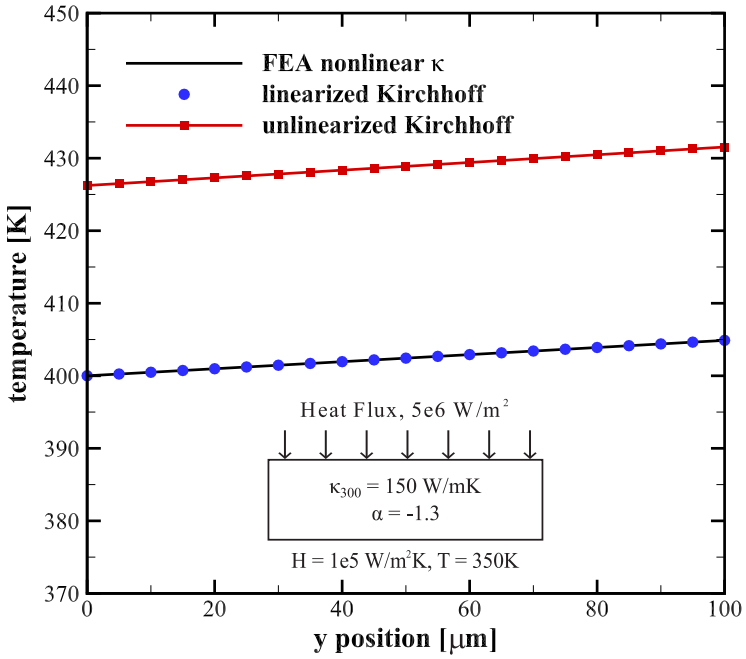


Figure 4.2: 1-D heat transfer solution for a 100μm thick piece of silicon using the expected boundary temperature of T_{ref} = 400K (linearized), and T_{ref} = 300K (unlinearized) in the Kirchhoff transformation [22] ©2020 ASME.

that this also requires using the associated value of the thermal conductivity at this chosen reference temperature. However, this approach requires some knowledge of the solution *a priori*. To illustrate this, an unlinearized Kirchhoff transformation solution, where the reference temperature is taken as 300K, is also shown in Fig. 4.2. In this case, the unlinearized version is seen to be inaccurate by 26K when compared to the full nonlinear FEA and the linearized Kirchhoff transformation, representing a roughly 6.5% error in absolute temperature, but a 52% error in terms of the increase in temperature above ambient.

One advantage of the proposed iterative method is that no *a priori* knowledge of the solution is required. To demonstrate this, the same 1-D simulation is performed with the iterative approach using T_{ref} = 300K and the associated thermal conductivity value. In Fig. 4.3 the iterative solution is shown as compared to that of the full

nonlinear FEA. The iterative approach is seen to agree exactly with the full nonlinear FEA using both a constant and a temperature dependent thermal conductivity, while requiring no information about the solution ahead of time.

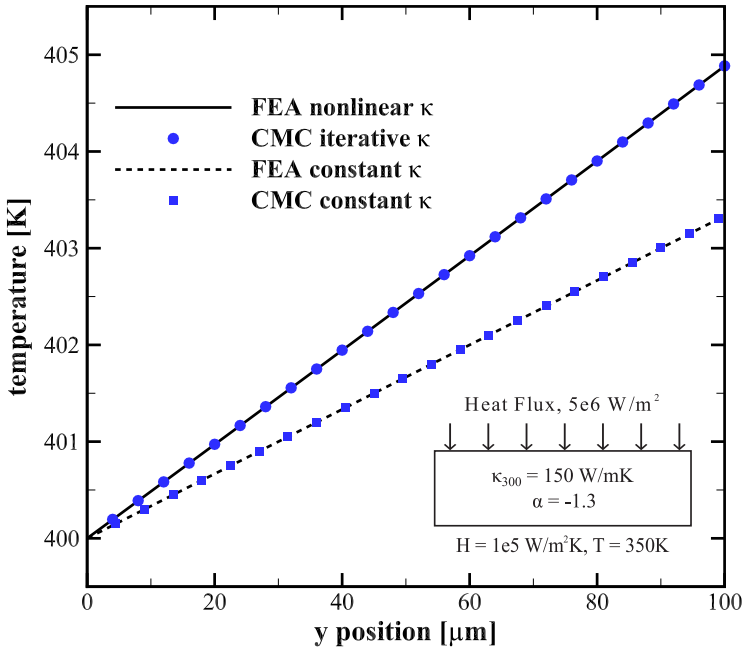


Figure 4.3: 1-D heat transfer solution for a 100 μm thick piece of silicon using both a constant and temperature dependent thermal conductivity in the iterative algorithm [22] ©2020 ASME.

4.3 Two-Dimensional Simulations

Next, a more complicated 2-D heat transfer problem is considered. Here, rather than a constant heat flux across the entire top boundary of the material, a single 2-nm wide heat source is centered on the top surface at $x = 250\mu$m. All other simulation parameters are the same as used previously, *i.e.*, the layout size is 500μm long by 100μm thick and the material properties as shown in the schematic of Fig. 4.4. Here, one cannot manipulate the convective boundary condition into the form of a Dirichlet one as was done in Sec. 4.2 because the temperature distribution at the convective boundary will be non-uniform simply due to the two-dimensional nature of the heat

conduction in this layout.

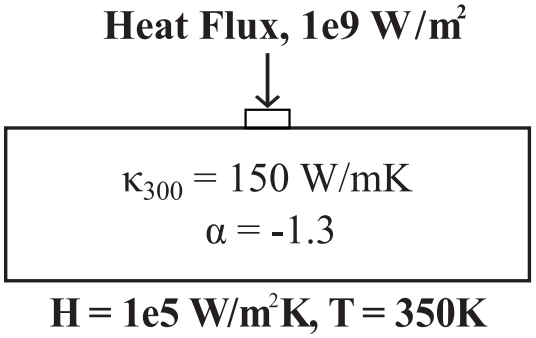


Figure 4.4: Schematic of the 2-D heat transfer layout. An inward heat flux of $H = 1 \cdot 10^9$ W/m^2 is applied over a 2 μm length centered at the middle of the top layer.

However, Bagnall *et al.* found that a reasonable approximation could be made in some instances by choosing the reference temperature in the Kirchhoff transformation to be that of the average temperature at the convective boundary [21]

$$\mathrm{T}_{ref} = \overline{\mathrm{T}_b} = \frac{\Phi \cdot l}{b \cdot H} + \mathrm{T}_\infty \tag{4.6}$$

where $\Phi = 1 \cdot 10^9$ W/m^2 is the heat flux applied to the top boundary, $l = 2\mu$m is the length over which the heat flux is applied, $b = 500\mu$m is the length of the convective boundary, $H = 1 \cdot 10^5$ W/m^{2}K is the heat transfer coefficient, and $\mathrm{T}_\infty = 350$K is the ambient cooling temperature. This corresponds to a reference temperature value of $\mathrm{T}_{ref} = \overline{\mathrm{T}_b} = 390$K.

In Fig. 4.5, the temperature distributions along the top boundary of Fig. 4.4, where the heat flux is applied, are shown as obtained from the Kirchhoff transformation with $\mathrm{T}_{ref} = 390$K, $\mathrm{T}_{ref} = 300$K, and from the full nonlinear FEA solution. When the reference temperature is chosen in this way, the Kirchhoff transformation shows excellent agreement with the full nonlinear FEA solution. However, when the reference temperature is chosen as $\mathrm{T}_{ref} = 300$K there is a temperature discrepancy

of approximately 20K throughout the entire curve.

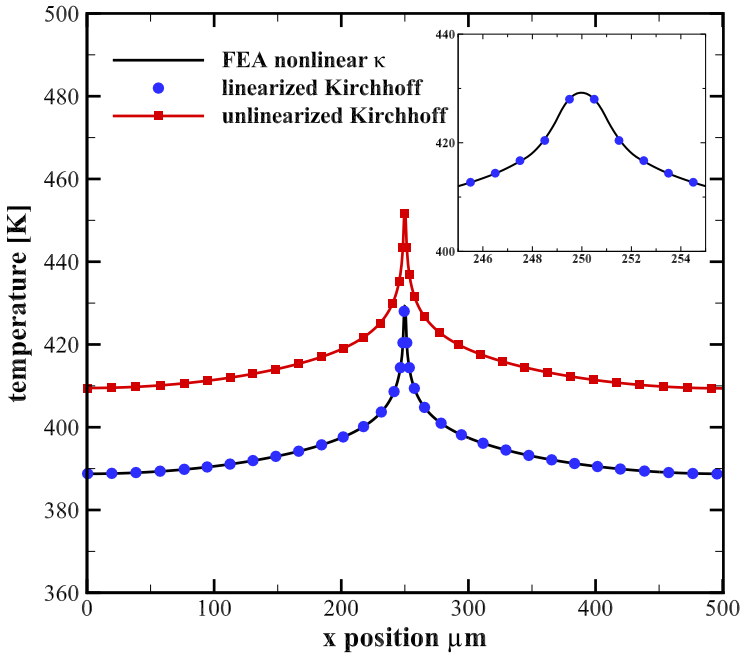


Figure 4.5: 2-D heat transfer solution for a 500μm long x 100μm thick piece of silicon using the average expected boundary temperature of T_{ref} = 390K (linearized), and T_{ref} = 300K (unlinearized) in the Kirchhoff transformation [22] ©2020 ASME.

Moreover, Bagnall *et al.* [21] also found that the approach of effectively linearizing the Kirchhoff transformation around a suitably chosen temperature becomes inaccurate at smaller thicknesses (as compared to the length of the material). For this reason, the Kirchhoff transformation was used to simulate the same 500 μm long silicon slab, but with its thickness reduced to 5 μm and then to 2 μm. Note that the chosen reference temperature, T_{ref} = 390K, does not change since the material thickness does not appear in Eq. (4.3).

From Figs. 4.6 and 4.7 it is clear that the result using the Kirchhoff transformation becomes increasingly inaccurate. In the case of the 5μm thickness the linearized Kirchhoff transformation solution over-estimates the peak temperature by roughly 4K over the FEA ($\approx$ 517K vs $\approx$ 513K), while the unlinearized solution is off by over 50K ($\approx$ 564K). For the 2μm thickness the linearized Kirchhoff transformation

solution over-estimates the peak temperature by roughly 25K over the FEA ($\approx$ 640K vs $\approx$ 615K), while the unlinearized solution is off by over 115K ($\approx$ 732K).

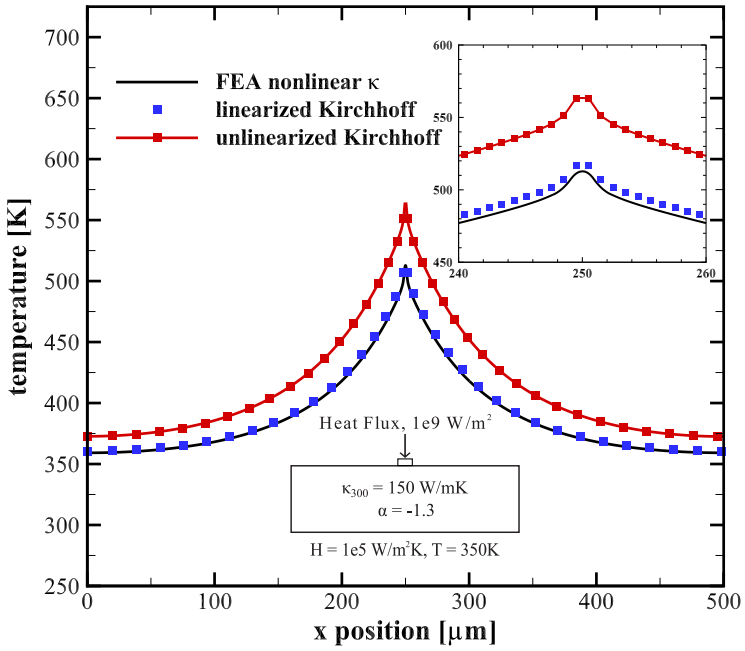


Figure 4.6: 2-D heat transfer solution for a 500μm long x 5μm thick piece of silicon using the average expected boundary temperature of T_{ref} = 390K (linearized), and T_{ref} = 300K (unlinearized) in the Kirchhoff transformation [22] ©2020 ASME.

As in the previous section, the iterative method was again used to simulate the of layout of Fig 4.4 for thicknesses of 100μm, 5μm, and 2μm using a reference temperature of 300K. For the 100μm thick substrate the iterative method agrees extremely well with the nonlinear FEA solution, similarly to the linearized Kirchhoff transformation. However, the iterative method displays far better accuracy than the linearized Kirchhoff transformation in the 5μm and 2μm thick simulations while requiring no *a priori* knowledge, as seen in Figs. 4.8 and 4.9.

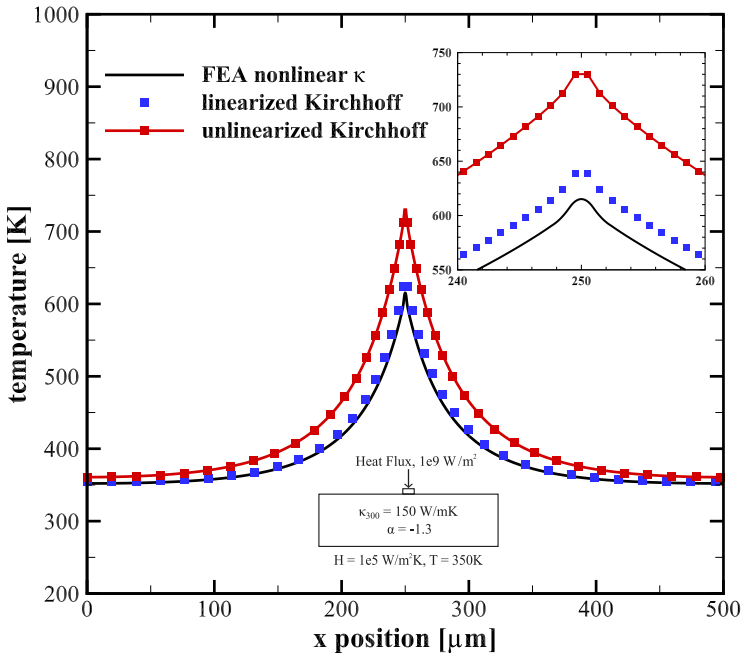


Figure 4.7: 2-D heat transfer solution for a 500μm long x 2μm thick piece of silicon using the average expected boundary temperature of T_{ref} = 390K (linearized), and T_{ref} = 300K (unlinearized) in the Kirchhoff transformation.

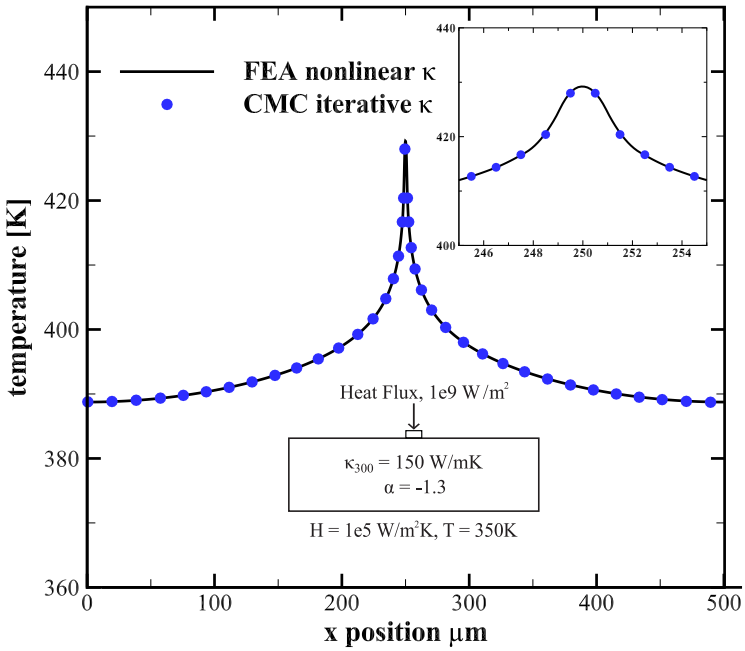
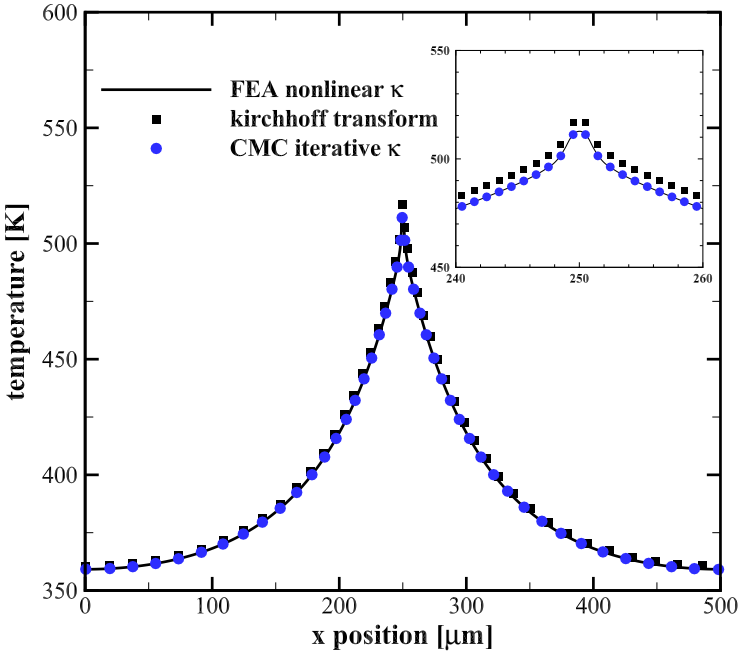
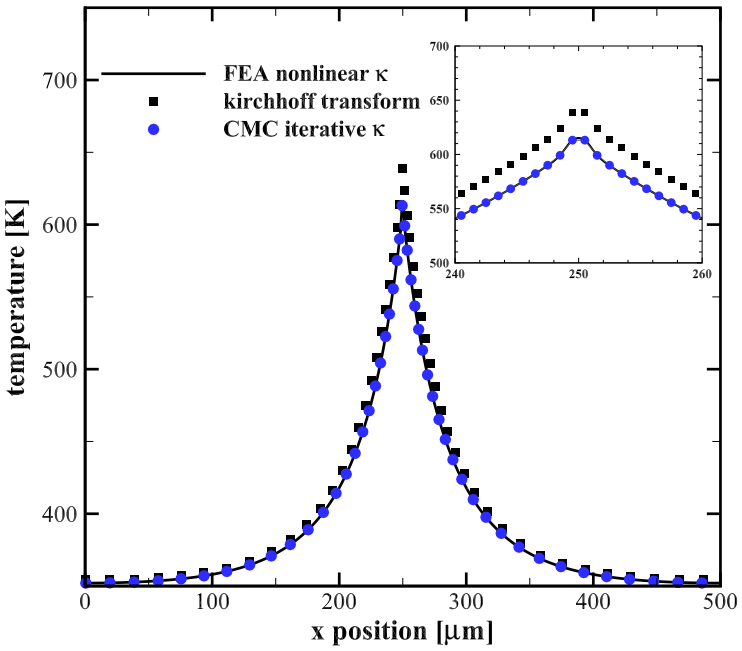


Figure 4.8: 2-D heat transfer solution for a 500μm long x 100μm thick piece of silicon using the proposed iterative method with T_{ref} = 300K [22] ©2020 ASME.



(a)



(b)

Figure 4.9: 2-D heat transfer solution using the proposed iterative method with T_{ref} = 300K for (a) a 500μm long x 5μm thick piece of silicon. (b) a 500μm long x 2μm thick piece of silicon [22] ©2020 ASME.

4.4 Multifinger GaN HEMT

One of the most robust tests found in literature for thermal simulation at the package scale, including convective boundary conditions, is that of a GaN radio frequency (RF) power amplifier (represented by a constant heat flux) on top of a SiC substrate [21, 73]. In [21, 73], the RF power amplifier package's thermal characteristics were simulated in a 3-D layout, whereas here it is simplified to a 2-D layout for practical reasons, namely memory and computational expense. The generalization to the 3-D case is straightforward.

The simulated layout consists of 11 heat sources of 0.5 μm width located at the very top of the material ($y = 102$ μm), each representing the heat generated by a GaN HEMT with an inward flux of 10^{10} W/m^2. The first source is located at $x = 25$ μm, and the sources are separated by 50 μm, with the final one located at $x = 525$ μm. The dimensions of the entire layout are 1000 μm by 102 μm. In the y-direction, the lower 100 μm layer is composed of a high thermal conductivity SiC substrate, while the top 2 μm layer is GaN.

As an initial benchmark, first a constant-conductivity simulation was performed using both the MATLAB® FEA solver in it's Partial Differential Equation (PDE) Toolbox [74] and the iterative approach. This was done to ensure the accuracy of the numerical grid used, and the full convergence of the solutions. The constant-conductivity solution was also used to find the average temperature along the convective boundary, $\mathrm{T}_{ref} = 460$K, to be used in the Kirchhoff transformation. The full nonlinear problem was then solved using the same thermal conductivities and temperature dependencies found in [21, 73] for each material, namely

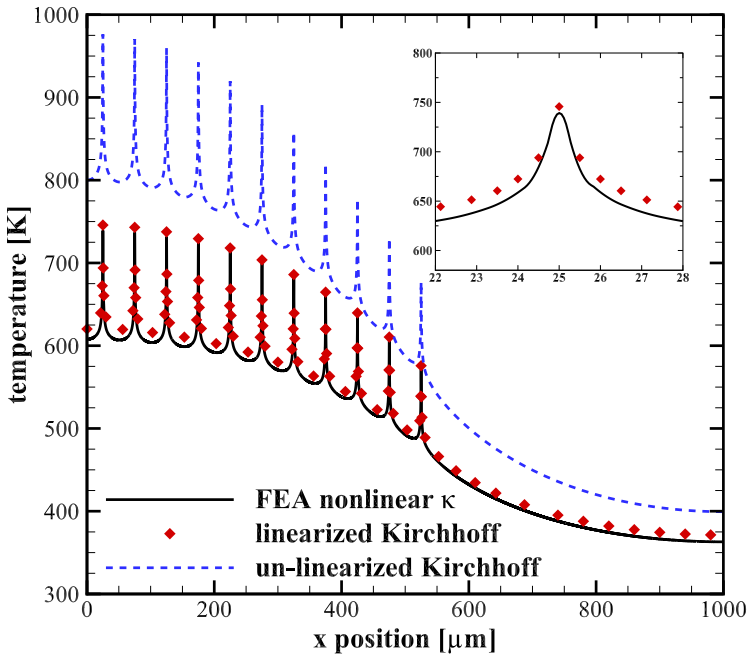
$$\kappa_{\mathrm{GaN}} = \kappa_{ref,\mathrm{GaN}} \left(\frac{\mathrm{T}}{\mathrm{T}_{ref}} \right)^{-1.3}, \kappa_{\mathrm{SiC}} = \kappa_{ref,\mathrm{SiC}} \left(\frac{\mathrm{T}}{\mathrm{T}_{ref}} \right)^{-1.4}, \tag{4.7}$$

where the values of the reference thermal conductivities to be used in the linearized Kirchhoff transformation at $T_{ref} = 460K$ are $\kappa_{ref,GaN} = 85.8$ W/mK and $\kappa_{ref,SiC} = 240.4$ W/mK. On the other hand, in the iterative approach, as well as the unlinearized Kirchhoff transformation, the reference thermal conductivity values taken at $T_{ref} = 300K$ are $\kappa_{ref,GaN} = 150$ W/mK and $\kappa_{ref,SiC} = 420$ W/mK.

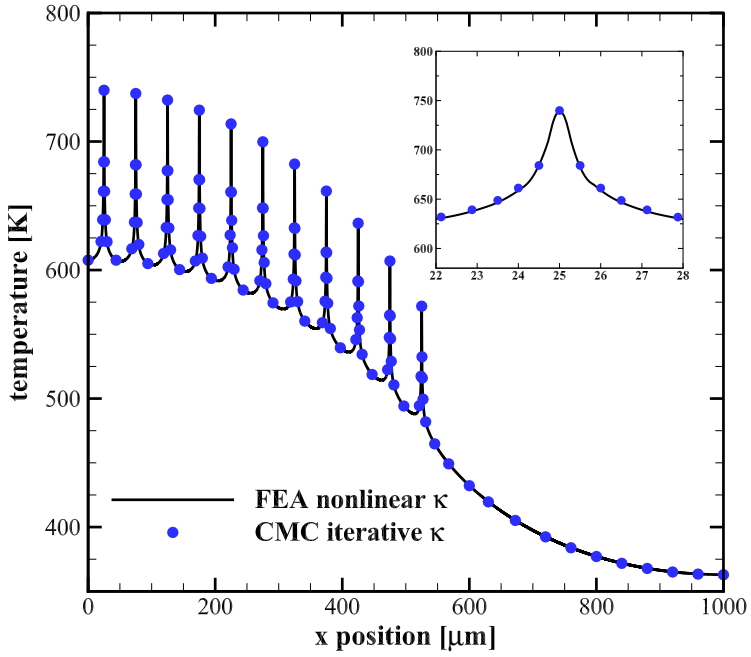
It should be noted again that when using the Kirchhoff transformation, the exponent must be identical for each material in the simulation, while this is true neither in the fully nonlinear FEA nor the iterative approach proposed here. Due to this, if using the Kirchhoff transformation, it is necessary to adjust the exponent for GaN to $\alpha = -1.4$ to match the value for SiC, as in [21].

As seen in Fig. 4.10a, The linearized version of the Kirchhoff transformation still gives a reasonably good result. In fact, the discrepancy in the peak temperature is only approximately 7K, with $T_{peak} \approx 746K$ for the Kirchhoff transformation vs. $T_{peak} = 739K$ for the nonlinear FEA solution. However, in order for the Kirchhoff transformation to obtain this level of accuracy, a simulation will always have to be performed twice. Firstly, with a constant thermal conductivity to obtain the appropriate reference temperature to be used, and secondly, using the Kirchhoff transformation to capture temperature-dependent effects.

On the other hand, the iterative approach follows the nonlinear FEA nearly exactly. The difference in the peak temperature between the two is 0.7K, which is likely due to using a fairly coarse grid with the iterative approach. The finite element mesh generated by the MATLAB® PDE toolbox [74] has 236,000 elements while the finite difference grid used for the iterative approach has dimensions of 383 cells x 205 cells for a total of 78,515 cells.



(a)



(b)

Figure 4.10: 2-D heat transfer solution for an 11-finger GaN RF power amplifier modeled using constant heat sources. a) The solution using the Kirchhoff Transformation. b) The solution using the proposed iterative method [22] ©2020 ASME.

Conclusion

In conclusion, the Kirchhoff transformation is able to provide reasonable results in the simulation of heat transfer with convective boundary conditions, however, it requires *a priori* knowledge of the solution. In contrast, the iterative method shows extremely good agreement with a nonlinear finite element method solution in the wide range of tests performed.

Chapter 5

COUPLED ELECTROTHERMAL SIMULATION: FORCING FUNCTIONS

5.1 Introduction

It is well known that electronic devices suffer from reliability issues and performance degradation due to self-heating effects [3, 4]. An excellent and comprehensive thermodynamic treatment of self-heating and heat transport in devices is given by Wachutka [9]. The simplest approach for computing the heat generation rates in a semiconductor is to use the dot product of the current density and electric field, $\mathbf{J} \cdot \mathbf{E}$ [W/m^3], which represents the projection of the current density along the electric field. Fushinobu *et al.* [10] developed one approach to include self-heating through a hydrodynamic model by assuming an energy decay path and using a relaxation time approximation, which has also been used elsewhere [13, 14]. On the other hand, Pilgrim *et al.* modeled heat generation through a net phonon emission approach and solved a thermal resistance matrix in order to obtain temperature maps [16–19]. While these approaches incorporate temperature-dependent scattering rates, the temperature dependence of material properties, like thermal conductivity, is often neglected.

Bonani and Ghione [15] took the temperature-dependent thermal conductivity into account in the context of device simulation by using the Kirchhoff transformation [60], which has also been used in the study of larger scale thermal spreading problems in semiconductor packages [21]. However, as discussed previously, convective boundary conditions, which physically represent an imperfect heat sink, become nonlinear under the Kirchhoff transformation.

A rather comprehensive model of electrothermal GaN device simulation is the one offered by Hao *et al.* [23, 24], where electron and phonon MC simulations are coupled along the 2DEG channel where heat generation is most significant, while Fourier's law is used in the rest of the domain. In [23, 24], as well as [10, 13, 14], parameters such as electron density, energy, and drift velocity are extracted from MC simulations in order to compute the energy exchanged between electrons and phonons, whereas here the heat generation rate is computed in real time by tracking the individual scattering events. However, in this work we do not carry out full phonon MC simulations solving the phonon BTE for heat transport, but instead we solve a steady-state EBE based on Fourier's law. In addition, a full electronic bandstructure and phonon dispersion is used, including all optical phonon modes without assuming them to be dispersionless. In other words, the optical modes are allowed to contribute to the thermal conductivity based on the material-specific phonon dispersion.

The computation of the heat generation rate is discussed in Sec. 5.2, while the Kirchhoff transformation and the iterative method for incorporating temperature-dependent thermal conductivities are applied to the electrothermal simulation of an experimentally characterized HEMT in Secs. 6.3 and 6.4, respectively.

5.2 Heat Generation Rate

The EBE from Sec. 3.2 is written as

$$\nabla \cdot \left(\kappa_{\mu}\left(\mathbf{r}, \mathrm{T}\right) \nabla \mathrm{T}\right) = -\left(\left.\frac{\partial W_{\mu}}{\delta t}\right|_{e-p} + \left.\frac{\partial W_{\mu}}{\delta t}\right|_{p-p}\right) = -P_{\mu}, \tag{5.1}$$

where the heat generation rate P_{μ} represents the forcing function for the EBE. The simplest choice to model the heat generation is given by $P_{\mu} = \mathbf{J} \cdot \mathbf{E}$ [75–77]. However, when this approach is used, the maximum value of P_{μ} is likely to occur at the peak of the electric field, E_{peak}. This is because the electric field also affects the carrier

velocity and concentration, which is taken into account as $\mathbf{J} = qn\mathbf{v}$.

One particular hydrodynamic model [10, 13] solves the system of equations given by

$$C_{LO}\frac{\partial T_{LO}}{\partial t} = \frac{3}{2}nk_B\left(\frac{T_e - T_{LO}}{\tau_{e-LO}}\right) + \frac{nm*v_d^2}{2\tau_{e-LO}} - C_{LO}\left(\frac{T_{LO} - T_A}{\tau_{LO-A}}\right) \tag{5.2}$$

$$C_A\frac{\partial T_A}{\partial t} = \nabla^2 T_A + C_{LO}\left(\frac{T_{LO} - T_A}{\tau_{LO-A}}\right) \tag{5.3}$$

where the LO subscript indicates the longitudinal optical mode value in particular for the temperature, specific heat, and relaxation times. The first term on the RHS of Eq. (5.2) represents the rate at which energy moves from electrons to LO phonons, and the final term that from LO phonons to acoustic phonons. Meanwhile, the second term is the energy relaxation rate from electrons to LO phonons, computed using parameters extracted from the MC.

The capabilities of the CMC simulation framework have been expanded to include a coupled flux-based electrothermal device solver based on the phonon scattering events themselves. This allows us to retain the speed advantages of the CMC, while obtaining a physically accurate picture of steady-state heat generation within a device. Here, rather than by using the macroscopic Joule Heating term, $\mathbf{J}\cdot\mathbf{E}$, the generation of heat is computed in real-time directly from the principles of energy conservation. As electron-phonon scattering events occur, the energy lost by electrons is used to compute a heat generation rate in each computational cell in the simulation. This offers a physically intuitive picture of heat generation in the device which is not dependent on macroscopic parameters such as the effective mass or the carrier mobility.

This heat generation rate is then used as the forcing function in the EBE for the optical and acoustic phonon modes. Meanwhile, the relaxation time approximation

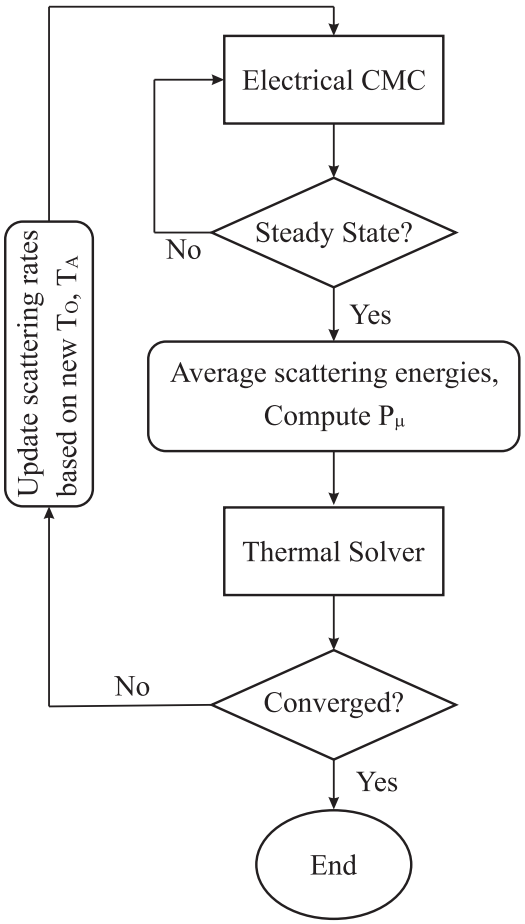


Figure 5.1: Flowchart of the electrothermal simulation process

is retained only for the net energy exchange between the optical and acoustic phonon modes, with a 10 ps relaxation time for GaAs [10]. The term for the phonon relaxation rate is written similarly to Eqs. (5.2) and (5.3) as

$$\left.\frac{\partial W_\mu}{\delta t}\right|_{p-p} = C_{op}\frac{T_O - T_A}{\tau_{OA}}. \tag{5.4}$$

In particular, C_{op} is approximated by the Einstein model [78, 79] where each optical mode has a contribution of

$$C_{\mu,op} = nk_B\frac{(\hbar\omega_{op}/k_BT)^2\exp(\hbar\omega_{op}/k_BT))}{(\exp(\hbar\omega_{op}/k_BT)-1)^2}, \tag{5.5}$$

where n is the number of primitive cells per unit volume and ω_{op} is the assumed single

frequency of the optical phonons. This must then be multiplied by the number of optical modes, dependent on the number of atoms in the unit cell.

Using the rejection algorithm, the scattering rates in each cell are subsequently updated according to the new temperature (as scattering increases with temperature), a new heat generation rate is obtained, and the process is repeated until the heat generation rate stabilizes and the temperature maps reach convergence in the outer iterative loop.

The final result is a temperature map for both acoustic and optical modes and the full electrothermal operating characteristics, *e.g.*, an I-V curve accounting for the effects of self-heating.

Although carriers will gain the bulk of their energy near the location of E_{peak}, they do not scatter instantaneously but instead travel up to a few mean free paths (mfps) before scattering away energy [77]. To illustrate this, and compare to the heat generation rate computed directly from the scattering events, simulations were performed on a rather simple 0.9μm long GaAs n^+-n-n^+ diode, seen schematically in Fig. 5.2. Here, the current density is computed as $\mathbf{J} = qn\mathbf{v}$, where n and $\mathbf{v}$ represent the average carrier density and carrier velocity obtained in each position-space cell in the Monte Carlo computational domain.

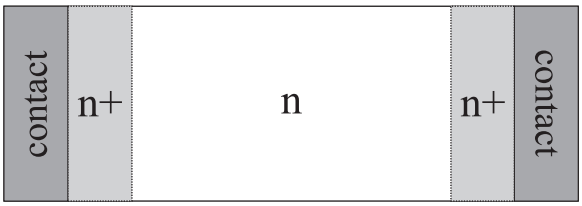
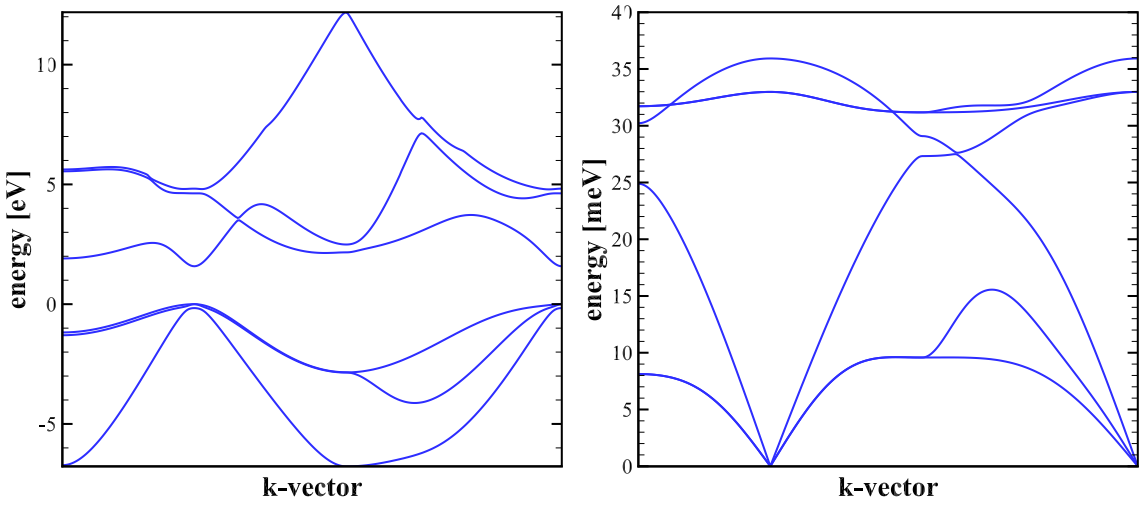


Figure 5.2: Schematic of 0.9μm x 0.2μm GaAs n^+-n-n^+ diode

Even in this simple device simulation, the CMC uses a full electronic bandstructure computed from the non-local Empirical Pseudo-Potential (EPM) method [80–82], and a full phonon dispersion computed from the 14 parameter valence shell model [83].

The bandstructure and phonon dispersion for GaAs are shown in Fig. 5.3.



(a) (b)

Figure 5.3: The full electronic bandstructure (a) and phonon dispersion (b) used in the CMC for GaAs.

The bias point being considered is that of $V = 3$ V applied to the right contact with respect to the left contact, while in the first case the heat generation rate was considered to be $\mathbf{J} \cdot \mathbf{E}$ and in the second case computed directly from the electron-phonon scattering events. The scattering rates were computed and tabulated for GaAs corresponding to a maximum temperature of 500K and this scattering rate ,as modeled by using the rejection algorithm, is shown in Fig. 5.4.

The diode is quasi-1-D as far as electrostatics are concerned because the contacts cover the entire device in the y-direction. Due to this, slices along the x-direction of the device can be taken, and their values averaged to smooth curves and reduce noise. This has been done in Fig 5.5, with the field and heat generation rates averaged over 10 horizontal slices for illustrative purposes. The heat generation rate given by $\mathbf{J} \cdot \mathbf{E}$ very closely follows the electric field seen in the device, and the peaks align exactly, as expected. However, since the electrons are expected to travel a few mfps before dissipating energy, the heat generation rate computed directly from

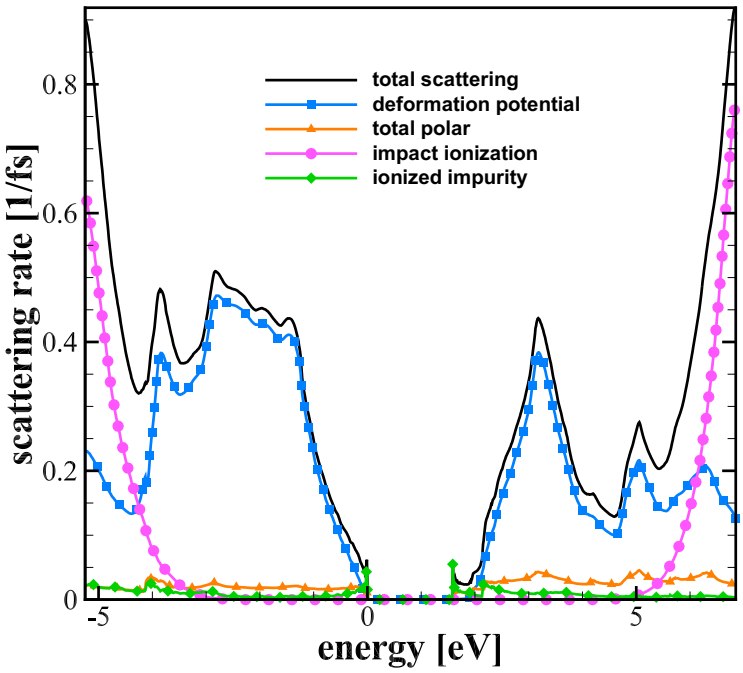


Figure 5.4: GaAs scattering rates for $T = 500$K used in the rejection algorithm.

the electron-phonon scattering events is shifted to the right approximately by 25nm. Since the saturation drift velocity of GaAs is $\approx 1.2 \cdot 10^7$cm/s, and the electron-phonon energy relaxation time is approximately 0.1ps, this corresponds to a mfp of $\approx$ 12nm, indicating that the shift seen if Fig. 5.5 represents approximately 2 mfps.

Moreover, the same behavior is seen in the temperature profiles obtained from the CMC simulations, seen in Fig. 5.6 even for very small temperature differences. This is because the difference in the location of the hot spot is simply due to the difference in the location of the peak heat generation.

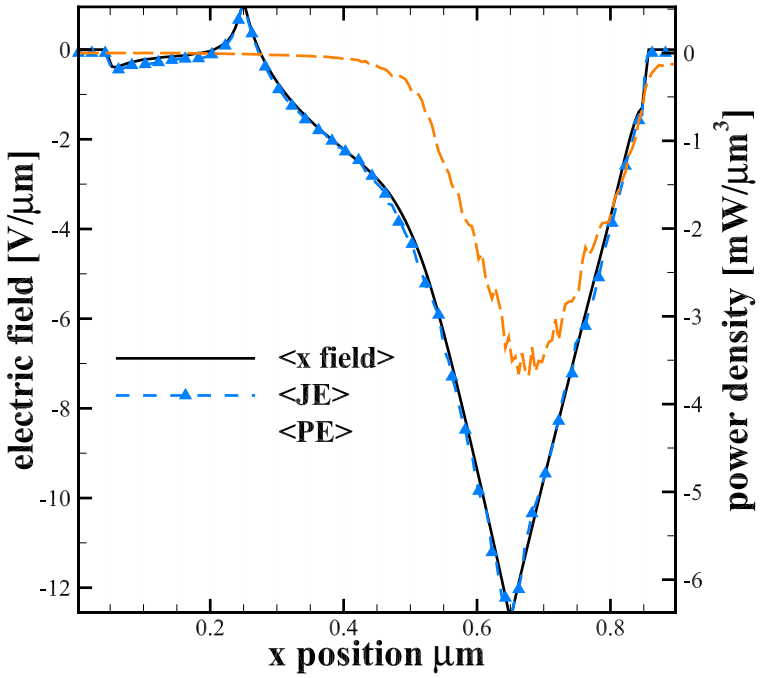


Figure 5.5: Electric field vs heat generation rates of $\mathbf{J} \cdot \mathbf{E}$ and that computed from electron-phonon scattering events (PE), respectively.

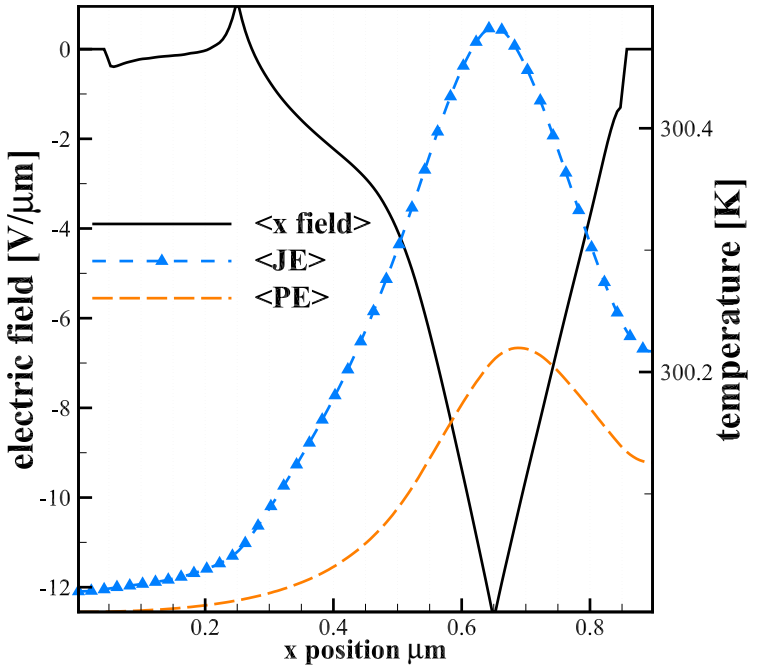


Figure 5.6: Electric field vs. lattice temperature obtained from the $\mathbf{J} \cdot \mathbf{E}$ heat generation rate, and the acoustic temperature from the electron-phonon heat generation rate, respectively.

Conclusion

The simplest choice for the heat generation rate, $\mathbf{J} \cdot \mathbf{E}$, yields inaccurate results for the location at which energy is dissipated by electrons through scattering events with the lattice. In contrast, a typical hydrodynamic model improves the heat generation rate by better accounting for the electron dynamics. However, these models generally use parameters which are often themselves obtained from Monte Carlo simulations.

In contrast, computing the heat generation rate directly from the CMC, by using the scattering rejection algorithm to find the energy transferred to the lattice through electron-phonon interactions, shows the expected shift of the hot spot of up to a few mean free paths from the location of the peak electric field.

Chapter 6

ELECTROTHERMAL GAN HEMT SIMULATION

6.1 Introduction

As wireless communications are ever-increasingly popular, radio-frequency communication systems are broadly encountered in everyday life. In consumer applications, fifth generation (5G) wireless networks have recently started rolling out and the so-called Internet of Things (IoT) has been a particular area of attention in the consumer product industry. However, other large markets in terms of applications are the defense sector as well as satellite and space exploration. Both high performance and high reliability are an emphasis in both fields of operation, however these two aims generally present a trade-off relationship, *i.e.*, an increase in one is often at the expense of the other. Additionally, many devices are generally battery powered, as in most cases a power source is not always locally available. Hence, power efficiency and portability are also of concern. All of these requirements need to be addressed in modern radio frequency (RF) systems.

Large bandgap materials improve reliability in two aspects: 1) higher temperatures are required for the doped material to become intrinsic, and 2) they possess higher breakdown voltages. Electrical breakdown is often driven by impact ionization scattering events which occur at high electric fields, as seen in the scattering rates of Fig. 6.1 where impact ionization becomes significant at energies above 7eV. Although the rate is fairly low at these energies, it does not take a large number of impact ionization events to set off an "avalanche effect". The GaN impact ionization rate can be compared with the rates shown for GaAs in Fig. 5.4, where it becomes

significant under 5eV and dominates the total scattering rate above 6eV.

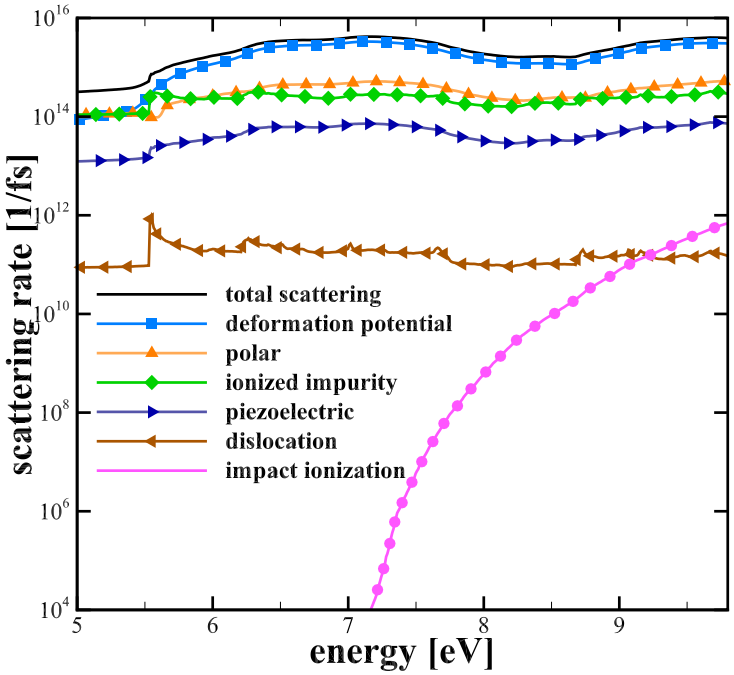


Figure 6.1: GaN scattering rates used in the rejection algorithm, where it is seen impact ionization begins to become significant above 7eV.

Another advantage for GaN is its high thermal conductivity as compared to traditional semiconductors such as Si or GaAs. This, in principle, allows GaN to better dissipate the heat generated inside the device. However, as seen in Sec. 6.4, convective conditions at material interfaces can be a bottleneck for heat dissipation. Additionally, AlGaN/GaN heterostructures have the advantage of the spontaneous formation of a 2DEG at the junction due to polarization effects, whereas in AlGaAs/GaAs systems the 2DEG must be formed with a modulation doping layer.

6.2 Polarization Effects

In III-V compound materials such as GaN, AlN, and InN, the crystal structure is typically of the Wurtzite variety, in contrast to GaAs which has a zinc-blende crystal structure [84]. This hexagonal crystal structure is characterized by the dimensions

a, the sides of the hexagon, c, the height, and u, the distance from the Ga and N neighbors, as depicted in Fig. 6.2.

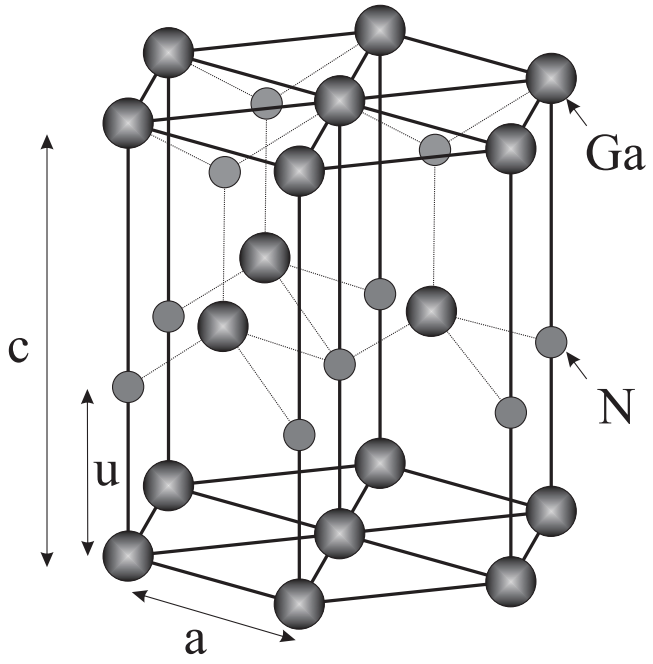


Figure 6.2: Illustration of the (Ga face) Wurtzite GaN crystal structure. In the N face configuration, the positions of the Ga and N atoms are interchanged.

An electric polarization results from the fact that N is more electronegative than Ga. At the face of the material, this spontaneous polarization is not canceled out as it is inside the material and hence there exists a polarization along the c-axis. For InN, GaN, and AlN spontaneous polarization values of $P_{SP} = -0.032, -0.029, and -0.081$ C/m^2 have been reported, respectively [84, 85]. Additionally, AlGaN/GaN heterostructures display a strong piezoelectric polarization at the interface due to the differing lattice constants of the two materials resulting in a strain in the lattice. Ambacher *et al.* [86] obtained the piezoelectric polarization as

$$P_{PE} = 2\frac{a_1 - a}{a}\left(e_{31} - e_{33}\frac{c_{13}}{c_{33}}\right), \tag{6.1}$$

where a is the hexagonal edge length of the relaxed crystal, as in Fig. 6.2, a_1 is the

lattice constant of the strained layer, e_{31} and e_{33} the piezoelectric coefficients, and c_{13} and c_{33} the elastic constants.

The total polarization is given by the sum of the spontaneous and piezoelectric coefficients as

$$P = P_{SP} + P_{PE}, \tag{6.2}$$

which, in the proximity of an interface with a high polarization discontinuity is related to the charge density [C/m^2] by the following relation:

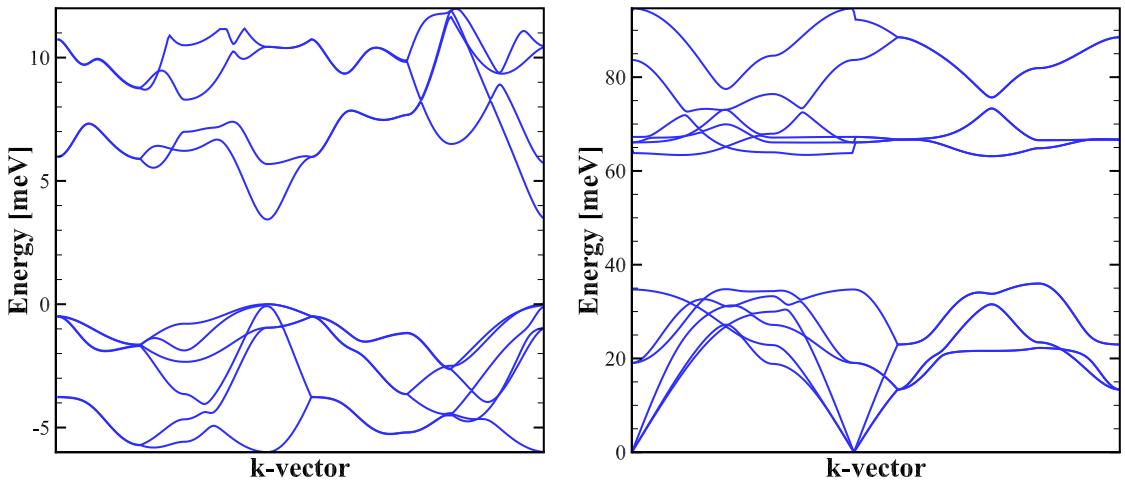
$$\sigma_P = -\nabla P. \tag{6.3}$$

6.3 GaN HEMT: Kirchhoff Transformation

An area of particular interest in electrothermal simulation is that of device reliability. Self-heating effects are well known to cause a decrease in operating current, as seen in phenomena such as the well known current collapse in HEMTs. The ability to design devices accounting also for their thermal characteristics, rather than only the electrical ones, would help in addressing reliability concerns prior to fabrication.

In this section, the CMC framework used for charge transport simulations takes into account scattering due to both optical and acoustic phonons (both polar and deformation potential), ionized impurities, and impact ionization. The material properties used for GaN are the same as those in Yamakawa *et al.* [87]. The electronic bandstructure and phonon dispersion used in the simulation are shown in Fig. 6.3.

To perform the coupled electrothermal simulation, first the electrical CMC is allowed to reach steady-state, just as in an isothermal simulation. Then, we average the power density transferred from the charge carriers to the respective phonon modes over a sufficiently large window to reduce statistical noise. That is, we obtain the



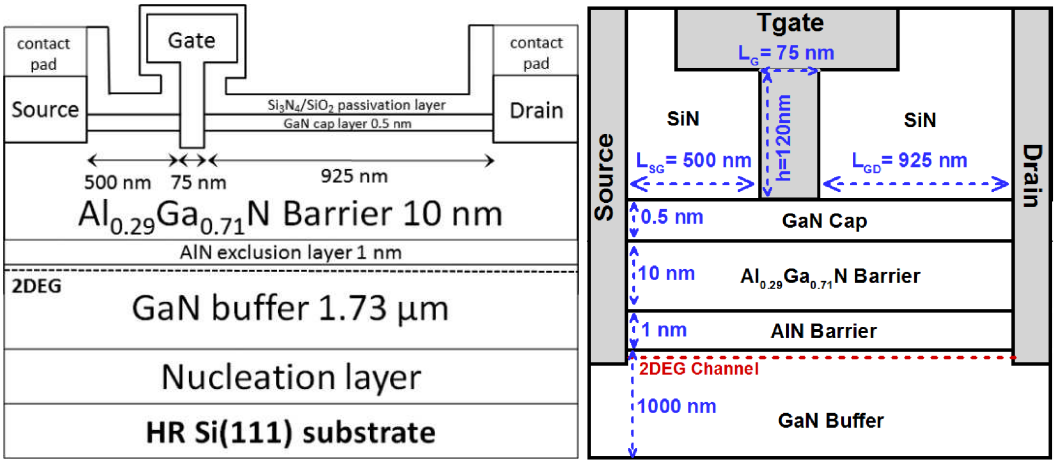
(a) (b)

Figure 6.3: The full electronic bandstructure (a) and phonon dispersion (b) used in the CMC for Wurtzite GaN.

value of $\partial W_\mu/\delta t|_{e-p}$ in Eq. (3.10) directly from the CMC simulation. Nevertheless, the energy exchange between the acoustic and optical phonon modes is still treated via the relaxation time approximation, and for GaN assumes values between 2 and 5 ps [88].

Here we have chosen the GaN/AlGaN structure characterized experimentally by Altuntas *et al.* [89]. The experimental layout of this device is shown in Fig. 6.4a, and the simulated layout consisting of a T-gate structure with a relatively large 1 μm buffer layer is shown if Fig. 6.4b [90]. A thin 1 nm AlN barrier is used to create a 2DEG conduction channel within the device due to a triangular well that is formed because of the conduction band mismatch at the heterojunction. In addition, thread dislocations are included in the simulations at a 2-D density of $T_{DD} = 5 \cdot 10^9$ cm^{-2}. Meanwhile, in addition to the triangular well, in the simulation the 2DEG channel is created by placing a polarization charge sheet at the heterointerface according to Ambacher's formalism [86]. A Dirichlet boundary condition is then imposed on the bottom of the simulated device representing a perfect heat sink held constant at 300K,

and so there is no complication due to nonlinearities introduced under the Kirchhoff transformation.

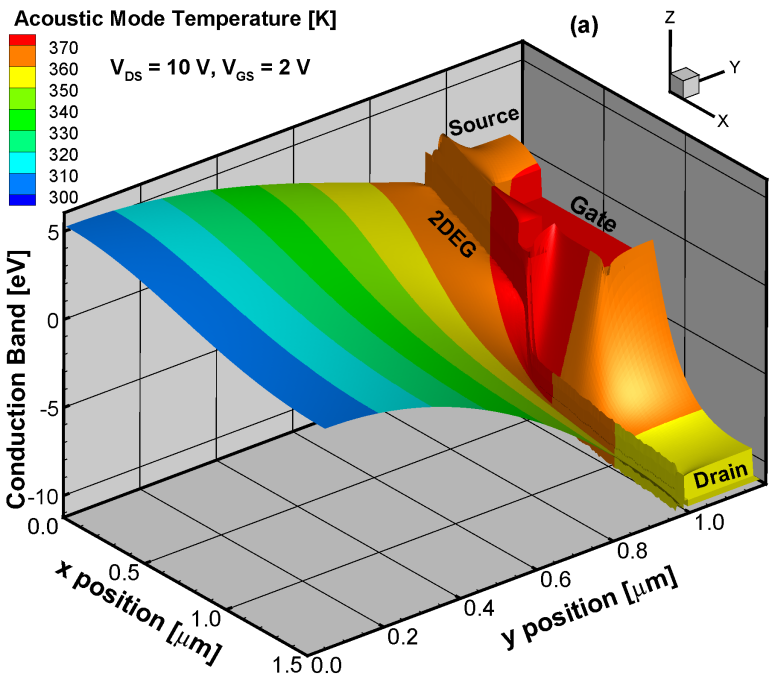


(a) (b)

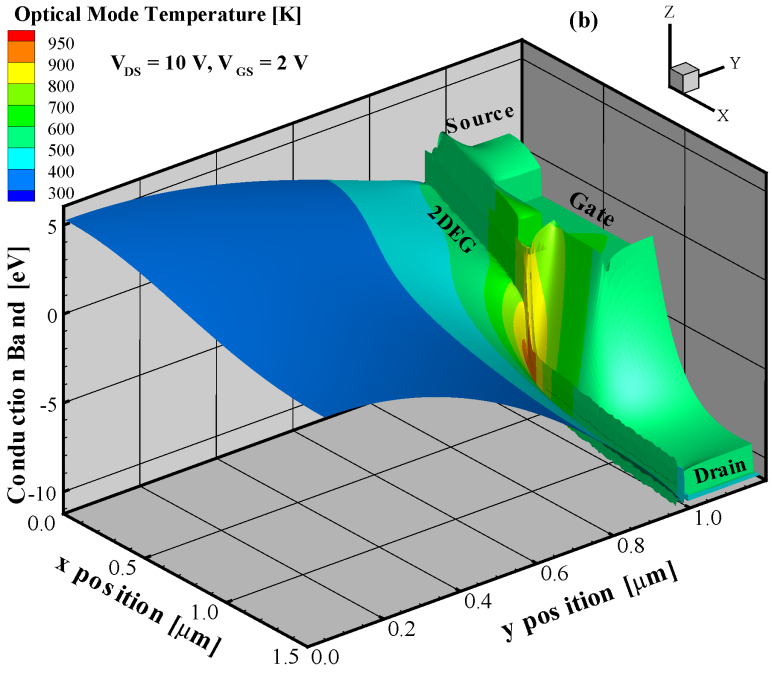
Figure 6.4: (a) Layout of the GaN/AlGaN HEMT characterized experimentally by Altuntas [89] and (b) the layout used in the coupled electrothermal simulation [90] ©2019 IEEE.

The resulting acoustic mode and optical mode temperatures at the bias point $V_{DS} = 2\text{V}$, $V_{GS} = 10\text{V}$, where both high currents and high electric fields are present, are shown in Figs. 6.5a and 6.5b, respectively. Here the temperatures are shown as a color map on top of the 3-D surface representing the conduction band to illustrate where the gate, source, drain, and channel regions are located in the device.

The optical mode temperature peaks at near 950K and is highly localized to the drain edge of the gate as should be expected. This is because the optical phonons have very low group velocity, and hence a low thermal conductivity. Hence, the dominant decay path available to them is to decay into acoustic phonons which generally happens at a rate slower than new optical phonons are generated. On the other hand, the acoustic temperature map is much more spread out because of the acoustic phonon modes' high thermal conductivity, and the peak acoustic temperature is seen to be roughly 375K, also to the drain side of the gate.



(a)



(b)

Figure 6.5: Acoustic (a) and optical (b) mode temperature maps obtained for the T-gate device at $V_{DS} = 10$V and $V_{GS} = 2$V [90] ©2019 IEEE.

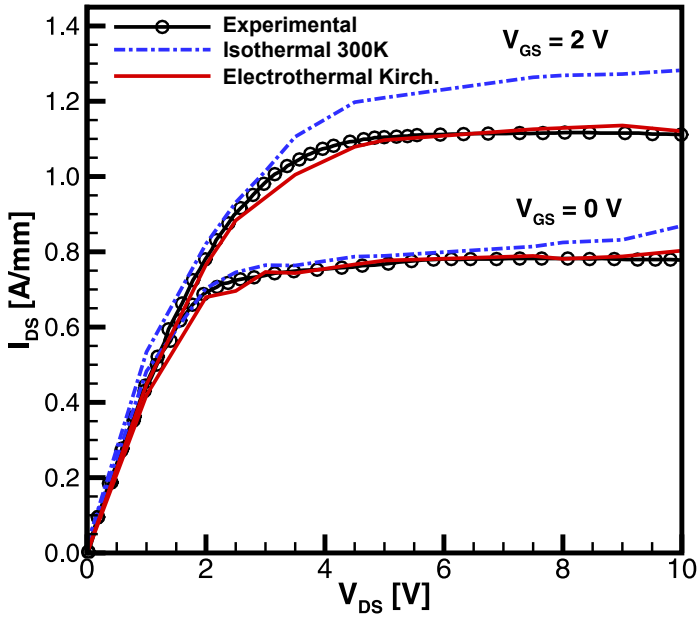


Figure 6.6: Full electrothermal I-V curve for V_{GS} = 2V and V_{GS} = 0V for the GaN/AlGaN HEMT using the Kirchhoff transformation with a Dirichlet boundary condition. The isothermal 300K simulation result is shown as a reference.

Finally, the I_{DS}-V_{DS} curves under DC operation from Altuntas [89] along with the those of the electrothermal model presented in this section are shown in Fig. 6.6 [90]. In this case, capturing thermal effects through the Kirchhoff transformation is seen to accurately reproduce the operating currents throughout the entire curve. In comparison, the isothermal simulations at 300K overestimate the current at higher drain voltages.

6.4 GaN HEMT: Iterative Method

In this section, first the electrothermal simulation on the same GaN HEMT from Sec. 6.3 is performed with a simple Dirichlet boundary condition and the result compared to that of the Kirchhoff transformation. Then, the effects of introducing convective boundary conditions are studied using the iterative approach, as it this a

principle advantage of the iterative method compared to the Kirchhoff transformation.

In the Dirichlet case, the boundary condition at the bottom edge of the device is again taken as a constant 300K, while for the convective condition the boundary condition is that of Eq. (3.18), *i.e.*, $-\kappa\nabla T_{b,C} = H(T_{b,C} - T_\infty)$ with the ambient temperature, T_∞ assumed to be a constant 300K, denoted in the results of Figs. 6.9 and 6.10. All other boundary conditions are assumed to be perfectly insulating Neumann conditions, *i.e.*, a zero flux exiting in the direction normal to the boundary. In other words, we iteratively solve:

$$\nabla^2 \mathrm{T}^i_{\mu,C} = -\frac{P_\mu(\mathbf{r})}{\kappa_{\mu,C}\left(\mathrm{T}^{i-1}\right)}, \tag{6.4}$$

subject to either the Dirichlet or convective boundary condition

$$\mathrm{T}_b = 300\mathrm{K} \tag{6.5}$$

$$-\kappa\nabla \mathrm{T}_b = H\left(\mathrm{T}_{b,C} - 300\mathrm{K}\right), \tag{6.6}$$

where $\mathrm{T}_{b,C}$ denotes the temperature in each respective boundary cell C along the bottom of the device.

Once the iteration of Eq. (6.4) has converged, in this case using the SOR (successive over-relaxation) method, the scattering rates are then updated in every cell of the simulation domain according to the new local temperature, and the process is repeated until both the optical and acoustic temperatures have converged to a steady-state value everywhere. The described simulation procedure is depicted in the flowchart of Fig. 6.7. Note that the process depicted in the thermal solver is solved self-consistently in an inner loop, and then convergence is tested based on the change between two successive temperature solutions.

Using this iterative approach, the final temperature maps obtained at the bias

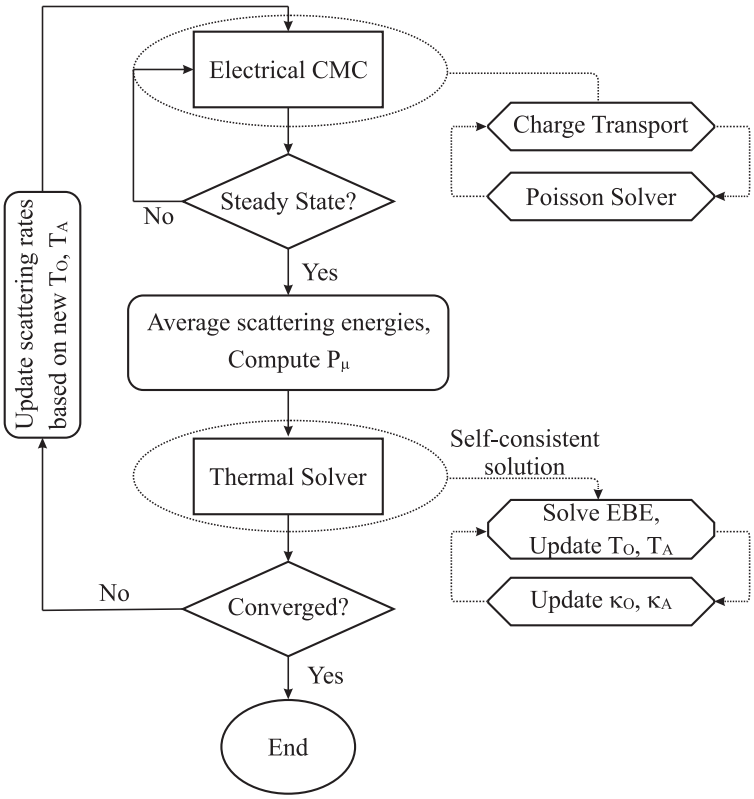


Figure 6.7: Flowchart of the electrothermal simulation process using the iterative thermal solver [22] ©2020 ASME.

point $V_{DS} = 2\text{V}$, $V_{GS} = 10\text{V}$, where both high currents and high electric fields are present, are shown in Fig. 6.8a and Fig. 6.8b, respectively. Here the temperatures are again shown as a color map on top of the 3-D surface representing the conduction band to illustrate where the gate, source, drain, and channel regions are located in the device. These can be directly compared to the previous result obtained using the Kirchhoff transformation in Sec. 6.3, and we see good agreement between the two sets of results.

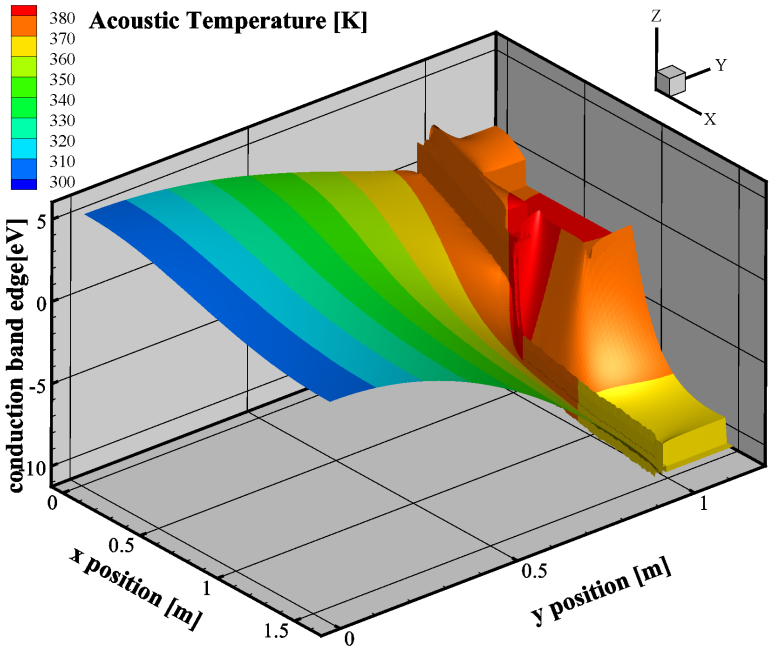
Namely, the optical mode temperature peaks at near 950K and is again highly localized to the drain edge of the gate. While on the other hand, the acoustic temperature map is much more dispersive because of their high thermal conductivity, and the peak acoustic temperature is seen to be roughly 380K, also to the drain side of

the gate.

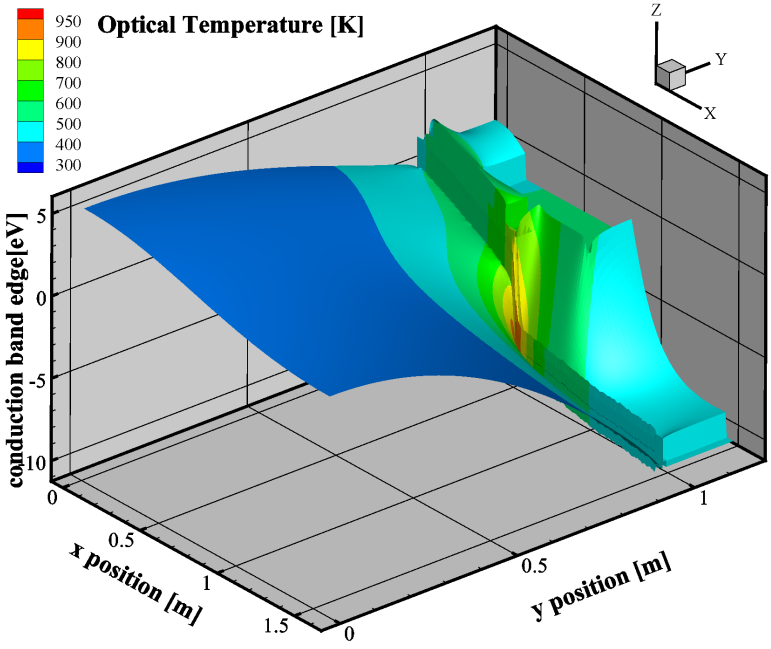
Next, the effects of using of using both convective boundary conditions and convective interface conditions are illustrated, as this is one of the main features of the iterative method.

In Fig 6.9, the I-V curves for three separate cases of modeling the heat sink are shown as follows: 1) using a simple 300K Dirichlet boundary condition, 2) using a convective boundary condition with $H = 10^8$ W/m^{2}K, and 3) using a convective boundary condition with $H = 10^7$ W/m^{2}K. Note that the perfect heat sink represented by the Dirichlet boundary condition would be represented by $H \to \infty$, and also that the most extreme condition of $H = 10^7$ W/m^{2}K is an unrealistically poor heat sink, as evidenced by the extremely large temperatures seen in Fig 6.10. However, the heat transport across semiconductor/substrate material interfaces is often poor, resulting in lower values of H than considered here.

Intuitively, as the value of the heat transport coefficient, H, of the heat sink decreases, its ability to dissipate the heat generated in the device is deteriorated, and an associated large decrease in current is seen due to the increase in electron-phonon scattering rates with temperature. Moreover, the current degradation is much more pronounced going from $H = 10^8$ W/m^{2}K to $H = 10^7$ W/m^{2}K than is seen between the Dirichlet and $H = 10^8$ W/m^{2}K boundary conditions. This is due to the large temperature increase when using the $H = 10^7$ W/m^{2}K heat sink, which in turn greatly increases the scattering rates seen in the device leading to the reduced current. Therefore, the ability to simulate non-ideal thermal contacts is highly valuable, as their effect on electrical characteristics are significant even for relatively large values of H, as seen in Fig. 6.9.



(a)



(b)

Figure 6.8: Acoustic (a) and optical (b) mode temperature maps obtained for the T-gate device at V_{DS} = 10V and V_{GS} = 2V using the iterative method [22] ©2020 ASME.

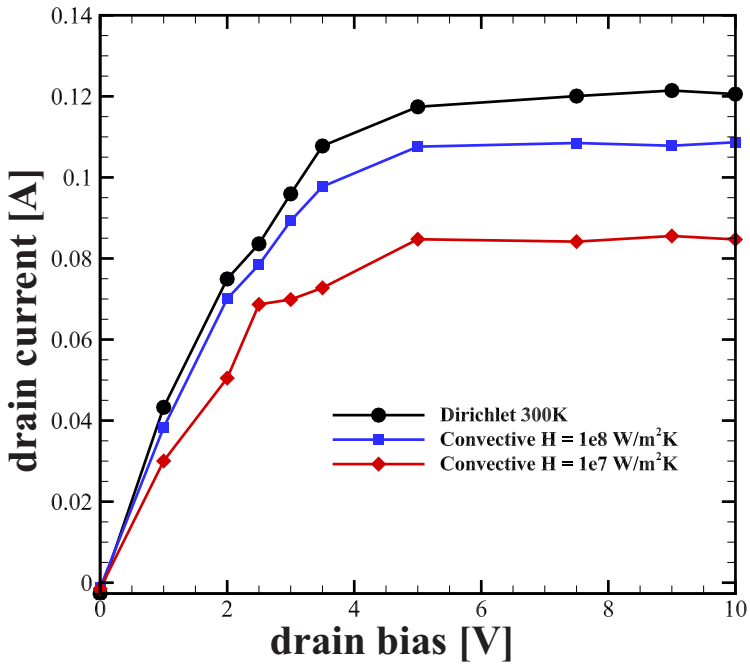
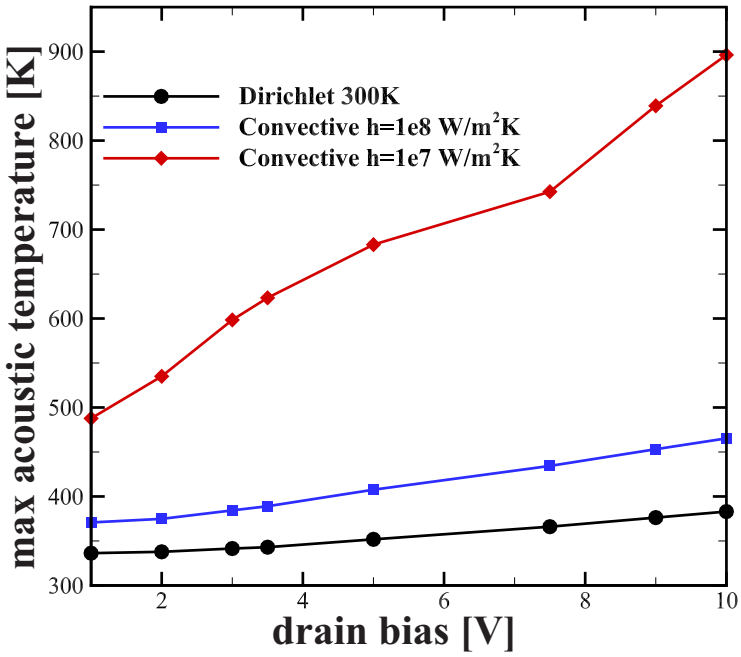
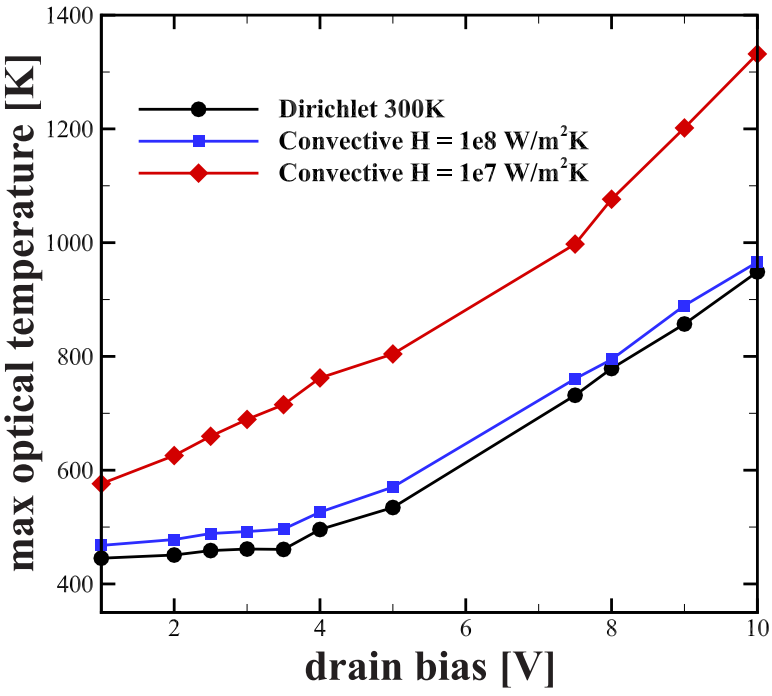


Figure 6.9: Full electrothermal I-V curve for $V_{GS} = 2$V for the GaN/AlGaN HEMT using both Dirichlet and convective boundary conditions [22] ©2020 ASME.



(a)



(b) [22] ©2020 ASME

Figure 6.10: Maximum optical and acoustic temperatures seen in the device at each simulated bias point.

Conclusion

The presence of polarization effects makes GaN heterostructures an attractive candidate for RF applications. The spontaneous polarization at the interface due to the differing electronegativity of Gallium and Nitrogen, in addition to the piezoelectric polarization due to strain between epitaxial layers create a 2DEG without the need for a modulation doping layer.

The Kirchhoff transformation has been used in the electrothermal characterization of an experimental AlGaN/GaN HEMT, and accurately reproduced the effects of self-heating in the DC operating characteristics.

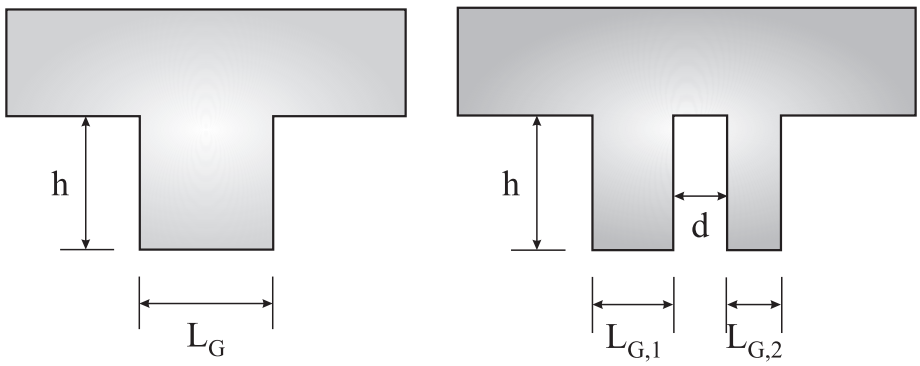
Additionally, the results of the iterative method agree well with those from the Kirchhoff transformation. Furthermore, the iterative method is also able to account for the study of imperfect interfaces. The ability to more accurately study such heat transport characteristics in electrothermal simulations can alleviate issues prior to fabrication of a device, and potentially lead to device layouts and architectures which more effectively dissipate heat.

Chapter 7

NOVEL GATE STRUCTURES

7.1 Introduction

It has been suggested [91] that rather than using the T-gate structure of Fig. 6.4, hot electron generation can be reduced by splitting the gate into two stems, creating a Π-gate. This layout has the effect of increasing reliability while retaining performance. In essence, this effect is achieved by increasing the electron-phonon scattering rates, which is due to the Π-gate layout producing a higher carrier concentration under the gate as well as a longer transit time due to a spacer region between the two stems [91]. The schematics of these layouts are shown in Fig. 7.1. In particular, in [91] they found that the Π-gate layout with $L_{G,1} > L_{G,2}$ reduced hot electron generation up to 40% under DC operation at bias points corresponding to peak DC power in the dynamic load lines.



(a) (b)

Figure 7.1: The T-gate and Π-gate structures. Design rules require that $L_G = L_{G,1} + L_{G,2}$ [91].

This idea has been applied to the T-gate layout of Fig. 6.4b to produce the

Π-gate layout shown in 7.1b, where $L_{G,1} = 50$nm, $L_{G,2} = 25$nm, $d = 25$nm, and $h = 120$nm. The source-to-gate and gate-to-drain lengths have been kept constant at $L_{SG} = 500$nm and $L_{DG} = 925$nm, respectively. All other device and simulation parameters are kept the same as in Sec. 6.4. As this is expected to decrease hot electron generation, it should correspond to a decrease in peak temperatures obtained from the iterative electrothermal model.

The temperature maps obtained for the bias point $V_{DS} = 10$V and $V_{GS} = 2$V are shown in Fig. 7.2, with the T-gate results shown in 7.2a and 7.2b and those for the Π-gate in 7.2c and 7.2d.

In these 3-D plots, the temperature is again plotted as a contour map on top of the 3-D conduction band edge. From this, the acoustic temperature is seen to be uniformly lower throughout the active region of the device near the AlN/GaN barrier at approximately $y = 1\mu$m, while the difference in the optical temperature distributions are more difficult to discern. To further investigate the temperature maps, the maximum acoustic and optical temperature at each bias point were extracted from the simulation results, and plotted in Fig. 7.3. A decrease of 4K is seen in the peak acoustic temperature at higher voltages while for the optical temperatures the peak decreases 49K in the asymmetric Π-gate as compared to the conventional T-gate structure. The decrease in the acoustic temperature corresponds to a reduction of 4.8% in the temperature rise above the heat sink temperature of 300K, and for the optical case a 7.5% decrease.

Moreover, the drain current for the two layouts was studied for a $V_{GS} = 2V$ and $V_{DS} = 0 - 10$V. It was found that the drain current for the two layouts in these simulations are nearly identical, and therefore the simulated decrease in temperatures came with no cost in terms of DC performance. Hence, novel gate architectures such as the Π-gate may be a viable path to mitigating undesirable hot electron effects,

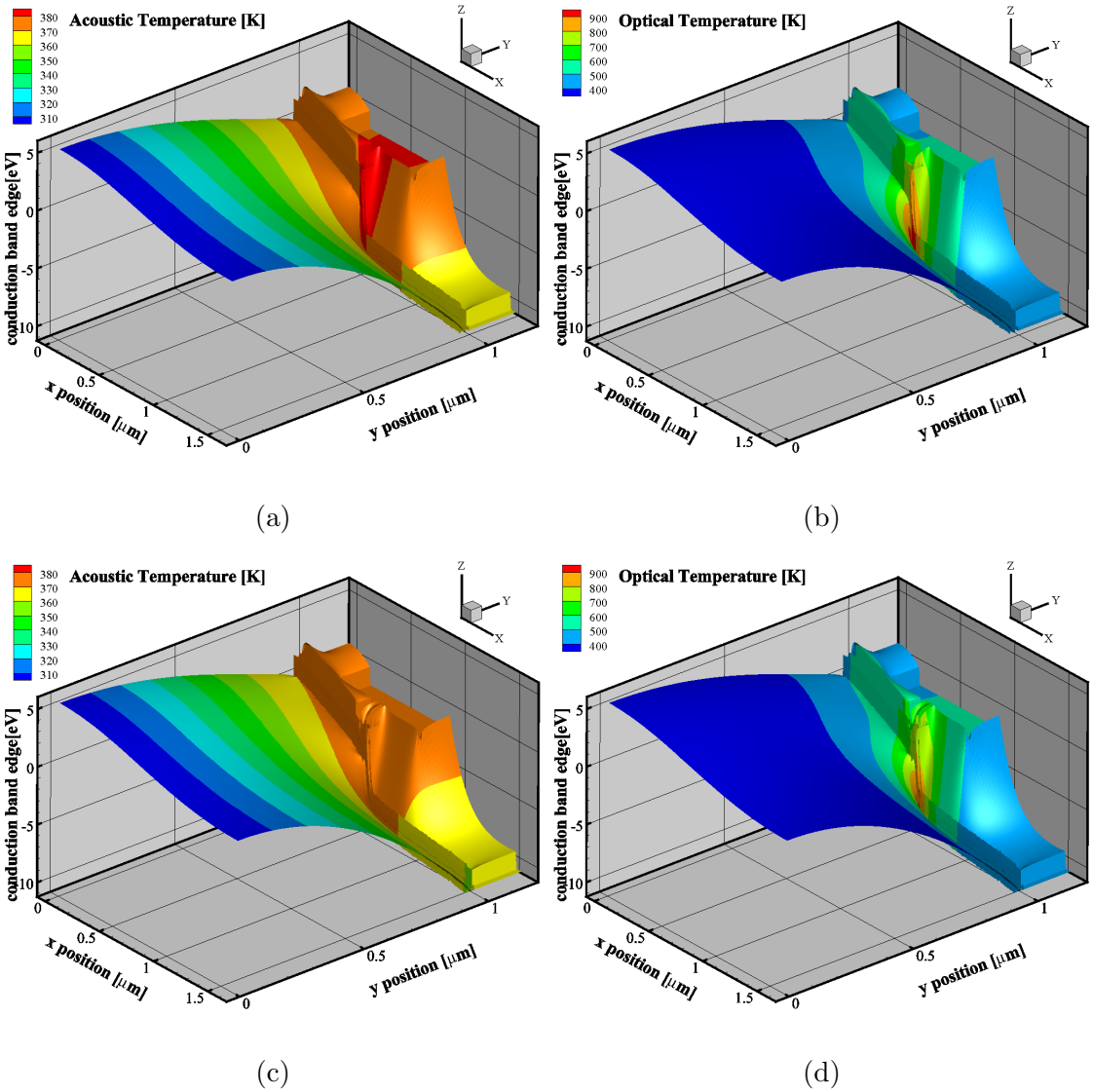
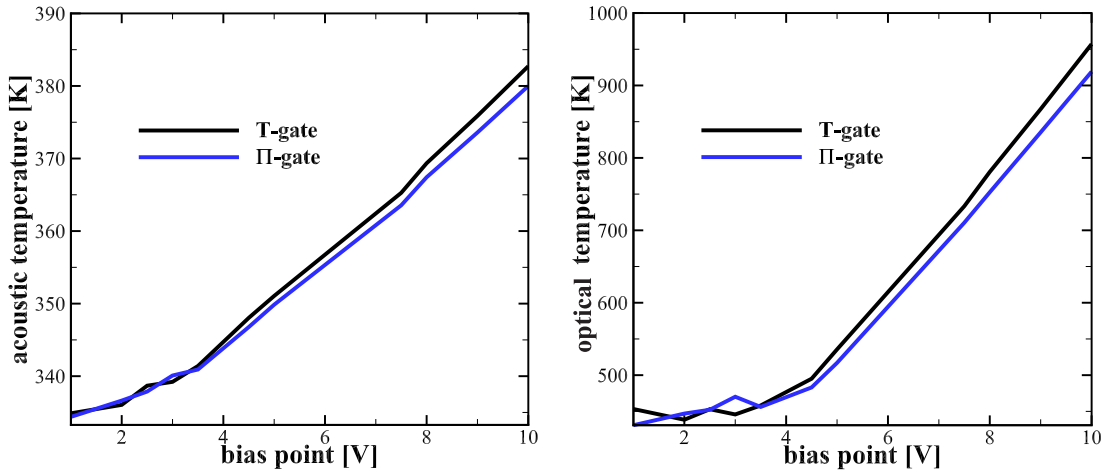


Figure 7.2: Acoustic and Optical temperature obtained for the T-gate (a,b) and the Π-gate (c,d) device structures.

without drastically reducing device performance.



(a) (b)

Figure 7.3: Maximum optical and acoustic temperature seen in the T-gate and the Π-gate structures.

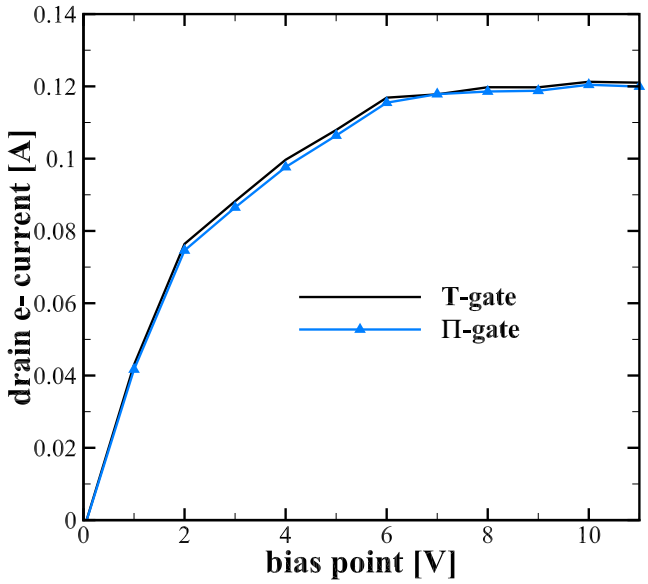


Figure 7.4: Drain current vs drain bias for the T-gate and the Π gate layouts. Results are for a gate voltage of 2V, as it was studied in Ch. 6.

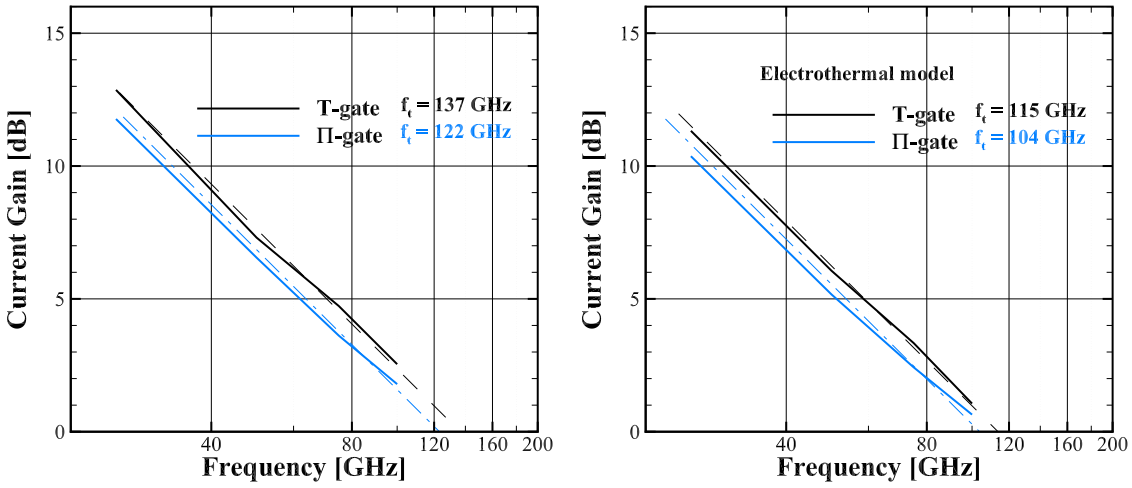
7.2 Electrothermal Small-Signal RF Simulation

The iterative method developed for electrothermal device simulation has also been extended to the domain of small signal AC simulations. In the case of small-signal simulation, it can be assumed that the applied AC signal doesn't have a significant

effect on the temperature maps of the device. Hence, the electrothermal simulations from the previous section can be used to generate small signal RF device characteristics as in [92, 93].

These simulations have been performed as follows: 1) perform a full DC electrothermal characterization of the device, 2) superimpose a small-signal AC sine wave on top of the DC operating point of interest, and 3) extract the relevant RF parameters to determine the cutoff frequency.

In Fig. 7.5a, the results of an isothermal sinusoidal small-signal AC simulation are shown. Using a 20 dB/decade line, for the T-gate layout from [89] a cutoff frequency of 135GHz is found compared to $f_t = 122$ GHz in the Π-gate. However, in [89] the cutoff frequency was experimentally determined to be 116 GHz. It is hypothesized that this discrepancy is due to thermal effects not being taken into consideration.



(a) (b)

Figure 7.5: Small signal simulation using sinusoidal inputs to determine the cutoff frequency in both (a) isothermal and (b) electrothermal simulations.

To test this hypothesis, the iterative electrothermal model has been applied to both layouts, and the results shown in Fig. 7.5b. Here, a cutoff frequency of $f_t = 115$ GHz is found for the Altuntas device, agreeing very closely with the value of $f_t = 116$

GHz, while for the Π-gate structure a cutoff frequency of 104 GHz is obtained. Hence, the cutoff frequency in the Π-gate structure is slightly less degraded by the inclusion of thermal effects in the model.

However, it is of note that the electrothermal simulation methods studied here are able to obtain much improved results for small-signal AC simulations in addition to those seen for DC operating currents. Therefore, this new model represents an obvious improvement on isothermal simulation

7.3 Conclusion

A novel gate architecture based on splitting the conventional T-gate in order to mitigate the generation of hot electrons also shows a decrease in the peak acoustic and optical temperatures obtained through electrothermal simulations. In addition, small-signal AC simulations using the electrothermal model are able to accurately reproduce the effects of self-heating in the determination of cutoff frequencies. This is evidenced by the electrothermal simulation for the T-gate layout recovering the experimental cutoff frequency of 115 GHz.

Chapter 8

CONCLUSION AND FUTURE WORK

8.1 Conclusion

Thermal management is nowadays perhaps the largest problem facing semiconductor devices, and is still an area of active research. Methods based on low order moments of the BTE can provide reasonable results in many cases, and lend insight to device characteristics in a relatively short simulation time. However, they necessarily neglect whatever information about the carrier ensemble is contained in the higher order moments that are not taken into account. In addition, assumptions made in order to derive these models in closed form are often only valid near-equilibrium.

Monte Carlo methods, on the other hand, are more computationally expensive but yield a full stochastic solution to the BTE. Furthermore, Monte Carlo methods are invaluable at high electric fields where far-from-equilibrium conditions apply. The cellular Monte Carlo (CMC) approach shifts some of the computational cost to hardware cost by pre-tabulating the scattering rates and loading them into memory at the start of a simulation. This provides a significant decrease of the simulation time as compared to the traditional Ensemble Monte Carlo (EMC), where final states are obtained at run-time for each scattering event. Moreover, the CMC is able to be adapted to local runtime conditions, akin the the EMC algorithm, by employing a rejection method to decide if a "proposed" scattering event should occur or not based on the local conditions.

Still, the use of MC methods in the study of thermal transport along with electrical characteristics has been less widespread until recently. While thermal models

have been developed and implemented in MC simulation, they often neglect the temperature dependence of material properties, namely the thermal conductivity.

The iterative approach for solving an energy balance equation, based on the $n = 2$ moment of the phonon BTE, and including temperature-dependent properties within a Monte Carlo simulation has been detailed here. This approach has been developed to overcome the drawbacks of using the Kirchhoff transformation while still allowing us to solve a linear Poisson-like equation. In particular, it tackles problems including convective boundary conditions, such as material interfaces or non-ideal heat sinks, in a straightforward way which is relatively easy to implement. The importance of solving a Poisson-like equation is that often an efficient finite-difference solver for the Poisson equation is present in any MC electrical device simulator.

However, one potential drawback of the iterative approach is that, since the thermal conductivity is likely to have a different value at each computational gridpoint, multigrid methods are likely to become very inefficient. Fortunately, in an electrothermal simulation, the solution to Poisson's equation is required much more often (orders of magnitude more times), than that of the phonon EBE.

Furthermore, a novel method for computing the heat generation rate is used in the CMC based directly on the electron-phonon scattering events using energy conservation principles. This provides a physically sound model of heat generation, and is not subject to macroscopic parameters such as the effective mass or carrier mobility.

The iterative method has been shown in Chapter 4 to work extremely well for thermal large-scale thermal spreading problems, as well as coupled to the CMC framework for full non-linear electrothermal device simulation, *e.g.*, as applied to the GaN HEMT devices as shown in Chapter 6. Moreover, it can be used to study the electrothermal characteristics of novel device layouts with the aim of mitigating thermal effects and improving device performance and reliability. Small-signal AC simulations using the

iterative electrothermal model are able to reproduce experimentally determined cutoff frequencies in devices.

8.2 Future Work

Recommendations for future work are as follows: 1) further validation of the electrothermal CMC model with experimental results 2) comparison with results for full phonon MC simulations, 3) coupling of the electrothermal model developed here to a full phonon MC, and 5) Expanding the work to other high power materials.

Any model should always be further validated, especially one which is fully numerical and whose results are not easily seen experimentally. For example, direct measurement of the internal temperature distribution of a device under operation is difficult, and most often surface measurements are used to deduce the internal temperature profile. Therefore, efforts have been made to reproduce macroscopic experimental results such as I-V curves and cutoff frequencies. The most significant use of the model presented here would be in the development of further device architectures to mitigate the effects of self heating by offering improved thermal management, either through decreased heat generation or through improved heat dissipation, or both.

As mentioned in [23, 24], a full phonon MC simulator has been coupled to an electron MC, in specific regions of interest, to obtain electrothermal device characteristics using macroscopic parameters extracted from the electron MC. In fact, that is also a goal of the electrothermal model presented here. While it provides many useful results of its own, as the hydrodynamic models for electron transport do, one aim is to use the temperature solutions obtained by the present simulation in order to initialize a full phonon MC simulation as presented in [57]. This could decrease the cost of the phonon MC by providing an improved initial guess of the temperature map in the device in order to start the phonon population closer to its steady state

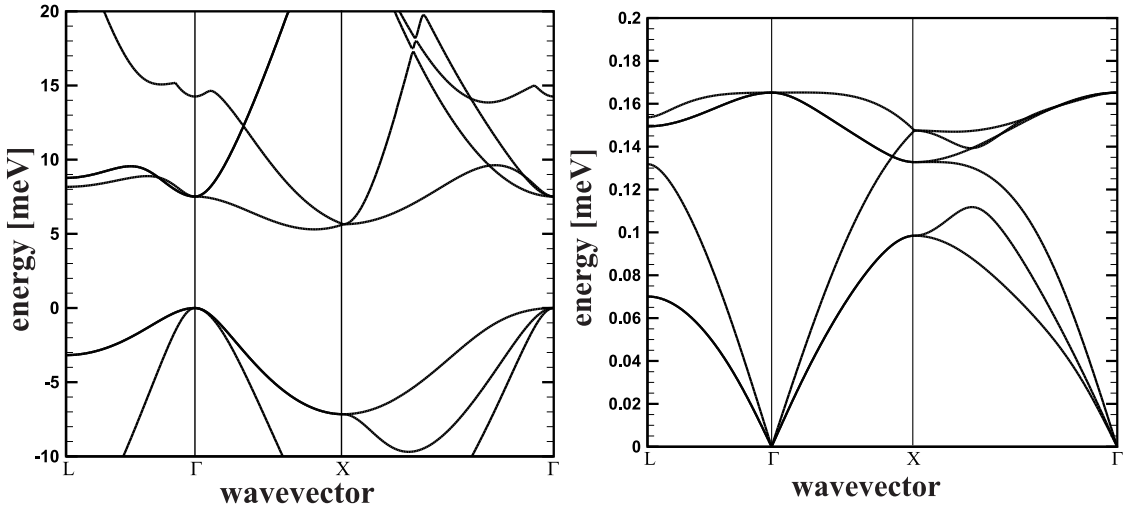
distribution. Additionally, this will provide evidence on the effects of the approximation made by using Fourier's law in the phonon EBE, as it has been suggested that Fourier's Law overestimates the heat flux [94, 95]. Another open question is what useful information may be obtained by using higher order moments of the phonon BTE as is sometimes done in the similar models for the electron BTE.

Additionally, since Fourier's law is diffusive in nature, it is believed that ballistic heat transport cannot be included in the model presented. However, it has been suggested in [96] that Fourier's law may contain information on ballistic heat transfer in some cases, and it can be modeled with a boundary condition precisely the form of the convective boundary conditions used here.

Finally, the electrothermal CMC model has principally been applied to Wurtzite GaN. However, other materials such as SiC and diamond are potentially attractive alternatives for high power devices due to their high thermal conductivity, in principle allowing for better thermal management. The potential bottleneck for GaN and SiC is that these materials are generally grown on top of another material and hence present an inherent interface between the active region of the device and the heat sink. Diamond on the other hand, if it can be efficiently doped with impurities, provides a very attractive alternative as its thermal conductivity is roughly one order of magnitude larger than other materials (approximately 2200 W/mK for diamond compared to approximately 220 W/mK for GaN).

In fact, it has been predicted that diamond is potentially the principle next generation material allowing for significant gains in performance and reliability [97]. Many advances have been made regarding using diamond in semiconductor devices [98], and an RF transistor has been fabricated showing a cutoff frequency of $f_t = 41$ GHz [99]. Hence, the characterization of such devices in the electrothermal CMC is a logical next step. In fact, the initial work for adding diamond as a material has been un-

dertaken within the CMC. The full electronic bandstructure and phonon dispersion computed in our CMC is shown in Fig. 8.1, which was independently verified against Watanabe, *et al.* [100].



(a) (b)

Figure 8.1: Full electronic bandstructure (a) and full phonon dispersion (b) for diamond as modeled in the CMC.

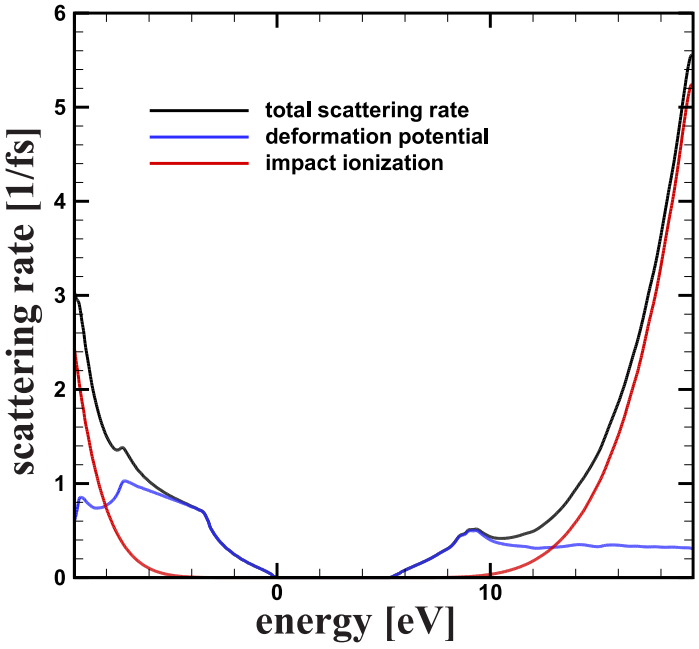
Additionally, a simple scattering model for diamond accounting for impact ionization and deformation potential has been implemented and shows good agreement with the velocity vs. field curves found experimentally by Nava *et al* [101–103], as shown in Fig. 8.2.

Regarding the thermal properties of diamond, the power law model used here show agrees well with temperature dependent values reported in [104, 105] with

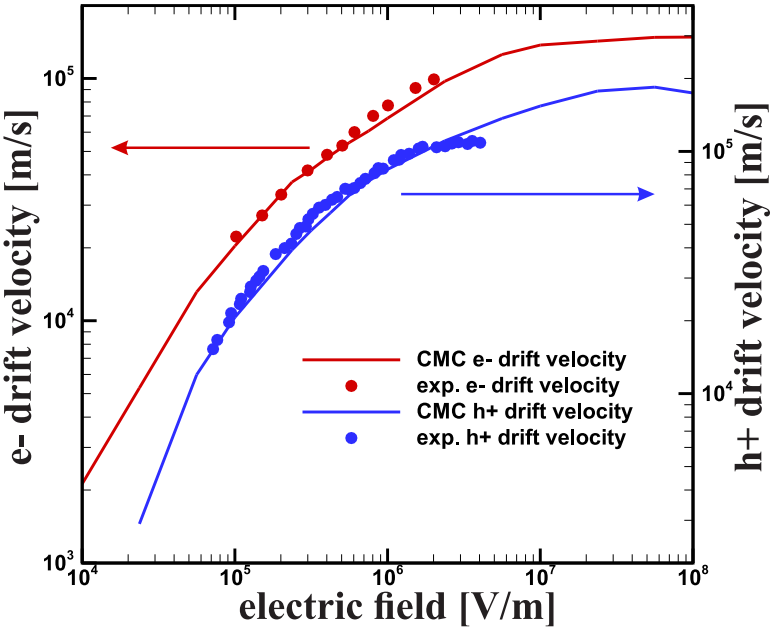
$$\kappa(\mathrm{T}) = 2450 \left(\frac{\mathrm{T}}{300\mathrm{K}} \right)^{-1.3}, \tag{8.1}$$

where $\kappa_{300} = 2450$ W/mK and $\alpha = -1.3$.

Subsequently, the multifinger IC benchmark simulation of Fig. 4.10b has been reproduced using diamond instead of GaN on SiC. The thermal conductivity relationship is that of Eq. (8.1), and the resulting temperature map along the top plane



(a)

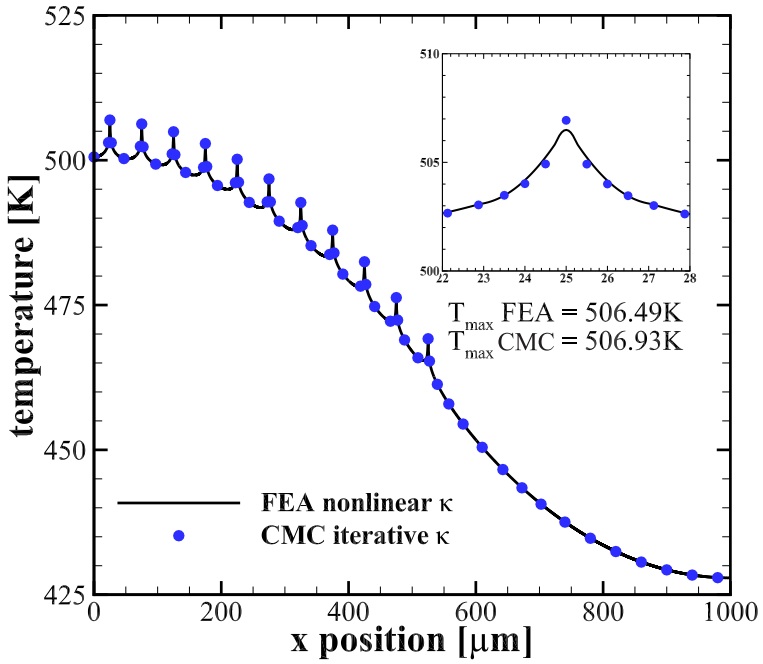


(b)

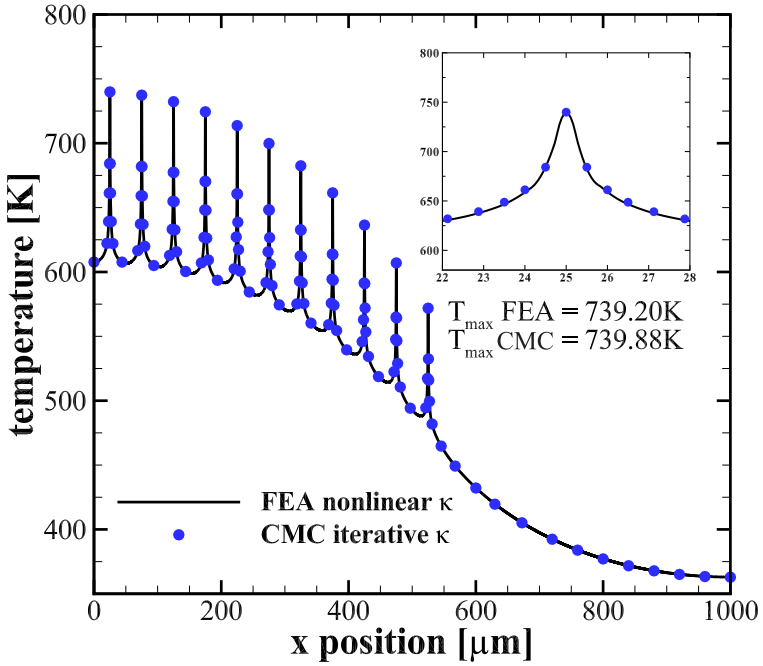
Figure 8.2: Scattering rates including deformation potentials and impact ionization (a), and field-velocity curves (b). Experimental data from [101–103].

of the layout is shown in Fig. 8.3 where it is shown that diamonds superior thermal properties, as modeled here, results in a peak temperature reduction of $\approx$ 233K in the 2-D simulation. Therefore, diamond is a promising material for the potential future of devices.

Each of these represent viable potential paths for the extension of, or applications for, the current state the electrothermal CMC package as detailed here.



(a)



(b)

Figure 8.3: 2-D heat transfer solution for an 11-finger RF power amplifier modeled using constant heat sources (a) composed of diamond and (b) composed of GaN-on-SiC [22] ©2020 ASME.

REFERENCES

[1] S. Sze and K. Ng, *Physics of Semiconductor Devices*, 3rd ed. Hoboken, NJ: Wiley, 2007.

[2] C. Sah, *Fundamentals of Solid-State Electronics*. Singapore: World Scientific, 1992.

[3] M. Artaki and P. J. Price, "Hot phonon effects in silicon field-effect transistors," *Journal of Applied Physics*, vol. 65, no. 3, pp. 1317–1320, Feb 1989. [Online]. Available: http://aip.scitation.org/doi/10.1063/1.343027

[4] O. Semenov, A. Vassighi, and M. Sachdev, "Impact of self-heating effect on long-term reliability and performance degradation in CMOS circuits," *IEEE Transactions on Device and Materials Reliability*, vol. 6, no. 1, pp. 17–27, Mar 2006. [Online]. Available: http://ieeexplore.ieee.org/document/1618650/

[5] M. V. Fischetti and S. E. Laux, "Monte Carlo analysis of electron transport in small semiconductor devices including band-structure and space-charge effects," *Physical Review B*, vol. 38, no. 14, pp. 9721–9745, Nov 1988. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevB.38.9721

[6] K. Hess, *Monte Carlo Device Simulation: Full Band and Beyond*. New York, NY: Springer, 1991.

[7] C. Jacoboni and L. Reggiani, "The Monte Carlo method for the solution of charge transport in semiconductors with applications to covalent materials," *Reviews of Modern Physics*, vol. 55, no. 3, pp. 645–705, Jul 1983. [Online]. Available: https://link.aps.org/doi/10.1103/RevModPhys.55.645

[8] C. Jacoboni and P. Lugli, *The Monte Carlo Method for Semiconductor Device Simulation*. Vienna: Springer, 1989. [Online]. Available: http://link.springer.com/10.1007/978-3-7091-6963-6

[9] G. Wachutka, "Rigorous thermodynamic treatment of heat generation and conduction in semiconductor device modeling," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 9, no. 11, pp. 1141–1149, 1990. [Online]. Available: http://ieeexplore.ieee.org/document/62751/

[10] K. Fushinobu, A. Majumdar, and K. Hijikata, "Heat generation and transport in submicron semiconductor devices," *Journal of Heat Transfer*, vol. 117, no. 1, p. 25, Feb 1995. [Online]. Available: http://heattransfer.asmedigitalcollection.asme.org/article.aspx?articleid=1441917

[11] J. Lai and A. Majumdar, "Concurrent thermal and electrical modeling of sub-micrometer silicon devices," *Journal of Applied Physics*, vol. 79, p. 11305, 1996. [Online]. Available: https://doi.org/10.1063/1.361424

[12] C. L. Tien, A. Majumdar, and F. M. Gerner, *Microscale Energy Transport*. CRC Press, 1998, no. 310.

[13] K. Raleva, D. Vasileska, S. M. Goodnick, and M. Nedjalkov, "Modeling thermal effects in nanodevices," *IEEE Transactions on Electron Devices*, vol. 55, no. 6, pp. 1306–1316, Jun 2008. [Online]. Available: http://ieeexplore.ieee.org/document/4511411/

[14] B. Padmanabhan, D. Vasileska, and S. M. Goodnick, "Is self-heating responsible for the current collapse in GaN HEMTs?" *Journal of Computational Electronics*, vol. 11, no. 1, pp. 129–136, Mar 2012.

[15] 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 inhomogenous thermal conductivity," *Solid-State Electronics*, vol. 38, no. 7, pp. 1409–1412, September 1994.

[16] T. Sadi, R. Kelsall, and N. Pilgrim, "Simulation of electron transport in InGaAs/AlGaAs HEMTs using an electrothermal Monte Carlo method," *IEEE Transactions on Electron Devices*, vol. 53, no. 8, pp. 1768–1774, 2006.

[17] T. Sadi, R. W. Kelsall, and N. J. Pilgrim, "Investigation of self-heating effects in submicrometer GaN/AlGaN HEMTs using an electrothermal Monte Carlo method," *IEEE Transactions on Electron Devices*, vol. 53, no. 12, pp. 2892–2900, 2006.

[18] T. Sadi, R. Kelsall, and N. J. Pilgrim, "Electrothermal monte carlo simulation of submicrometer Si/SiGe MODFETs," *IEEE Transactions on Electron Devices*, vol. 54, no. 2, p. 332–339, 2007.

[19] T. Sadi, R. W. Kelsall, and N. J. Pilgrim, "Electrothermal Monte Carlo simulation of submicron wurtzite GaN/AlGaN HEMTs," *Journal of Computational Electronics*, vol. 6, no. 1-3, pp. 35–39, 2007.

[20] W. Batty, C. E. Christoffersen, A. J. Panks, S. David, C. M. Snowden, and M. B. Steer, "Electrothermal CAD of power devices and circuits with fully physical time-dependent compact thermal modeling of complex nonlinear 3-D systems," *IEEE Transactions on Components and Packaging Technologies*, vol. 24, no. 4, pp. 566–590, dec 2001.

[21] K. R. Bagnall, Y. S. Muzychka, and E. N. Wang, "Application of the Kirchhoff transform to thermal spreading problems with convection boundary conditions," *IEEE Transactions on Components, Packaging and Manufacturing Technology*, vol. 4, no. 3, pp. 408–420, Mar 2014.

[22] K. Merrill and M. Saraniti, "Nonlinear electrothermal Monte Carlo device simulation," *Journal of Heat Transfer*, vol. 142, no. 2, Feb 2020.

[23] Q. Hao, H. Zhao, and Y. Xiao, "A hybrid Simulation technique for electrothermal studies of two-dimensional GaN-on-SiC high electron mobility transistors," *Journal of Applied Physics*, vol. 121, no. 20, p. 204501, May 2017.

[24] Q. Hao, H. Zhao, Y. Xiao, and M. B. Kronenfeld, "Electrothermal studies of GaN-based high electron mobility transistors with improved thermal designs," *International Journal of Heat and Mass Transfer*, vol. 116, pp. 496–506, Jan 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S001793101732015X

[25] M. Saraniti and S. Goodnick, "Hybrid fullband cellular automaton/Monte Carlo approach for fast simulation of charge transport in semiconductors," *IEEE Transactions on Electron Devices*, vol. 47, no. 10, pp. 1909–1916, 2000. [Online]. Available: http://ieeexplore.ieee.org/document/870571/

[26] N. Ashcroft and D. Mermin, *Solid State Physics*. Belmont, CA: Brooks/Cole, 1976.

[27] M. Lundstrom, *Fundamentals of Carrier Transport*. Reading, MA: Addison-Wesley, 1990.

[28] S. Selberherr, *Analysis and Simulation of Semiconductor Devices*. New York, NY: Springer, 1984.

[29] R. A. Smith, *Wave Mechanics of Crystalline Solids*. London: Chapman and Hall, 1969.

[30] C. Hamaguchi, *Basic Semiconductor Physics*, 3rd ed. Cham, Switzerland: Springer, 2017.

[31] S. Ross, *Introduction to Probability Models*, 9th ed. Burlington, MA: Elsevier, 2007.

[32] F. Beichelt, *Applied Probability and Stochastic Processes*. Chapman and Hall/CRC, 2018.

[33] C. R. Heathcote, *Probability: elements of the mathematical theory*. Dover, 2000.

[34] T. Grasser, H. Kosina, M. Gritsch, and S. Selberherr, "Using six moments of Boltzmann's transport equation for device simulation," *Journal of Applied Physics*, vol. 90, no. 5, p. 2389–2396, 2001.

[35] L. T. Decarlo, "On the meaning and use of kurtosis," *Psychological Methods*, vol. 2, no. 3, p. 292–307, 1997.

[36] D. D. Wackerly, W. Mendenhall, and R. L. Scheaffer, *Mathematical statistics with applications*. Duxbury, 2002.

[37] D. Vasileska, S. M. Goodnick, and G. Klimeck, *Computational electronics: Semiclassical and Quantum Device Modeling and Simulation*. CRC Press, 2017.

[38] D. Woolard, H. Tian, M. Littlejohn, R. Trew, and K. Kim, "The application of Monte Carlo techniques in advanced hydrodynamic transport models," in *Monte Carlo Device Simulation: Full Band and Beyond*, K. Hess, Ed. New York: Springer, 1991, ch. 8, pp. 219–266.

[39] M. Bina, "Charge transport models for reliability engineering of semiconductor devices," Ph.D. dissertation, Vienna University of Technology, 2014.

[40] J. Zhou, K. A. M., and D. Ferry, "The conditions of device simulation using full hydrodynamic equations," in *Computational Electronics: Semiconductor Transport and Device Simulation*, K. Hess, J. P. Leburton, and U. Ravaioli, Eds. Norwell, MA: Kluwer Academic Publishers, 1991, ch. 12, pp. 63–66.

[41] K. Blotekjaer, "Transport equations for electrons in two-valley semiconductors," *IEEE Transactions on Electron Devices*, vol. ED-17, no. 1, pp. 38–47, 1970.

[42] E. M. Azoff, "Semiclassical high-field transport equations for nonparabolic heterostructure degenerate semiconductors," *Journal of Applied Physics*, vol. 64, no. 5, p. 2439–2446, 1988.

[43] M. Gritsch, "Numerical modeling of Silicon-on-Insulator MOSFETs," Ph.D. dissertation, Vienna University of Technology, 2002.

[44] B. R. Nag, *Theory of electrical transport in semiconductors.* Pergamon Pr., 1972.

[45] T. Grasser, T.-W. Tang, H. Kosina, and S. Selberherr, "A review of hydrodynamic and energy-transport models for semiconductor device simulation," *Proceedings of the IEEE*, vol. 91, no. 2, p. 251–274, 2003.

[46] R. Stratton, "Diffusion of hot and cold electrons in semiconductor barriers," *Physical Review*, vol. 126, no. 6, pp. 2002–2014, jun 1962.

[47] C. Gardner, "Semiconductor device simulation: The hydrodynamic model," *IEEE Potentials*, vol. 23, no. 5, p. 17–19, 2004.

[48] N. Metropolis and S. Ulam, "The Monte Carlo method," *Journal of the American Statistical Association*, vol. 44, no. 247, pp. 335–341, 1949.

[49] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, "Equation of state calculations by fast computing machines," *Citation: J. Chem. Phys*, vol. 21, no. 6, p. 1087, 1953.

[50] N. T. Thomopoulos, *Essentials of monte carlo simulation.* New York: Springer, 2015.

[51] P. Glasserman, *Monte Carlo methods in financial engineering.* Springer, 2010.

[52] T. M. Carsey and J. J. Harden, *Monte Carlo Simulation and Resampling Methods for Social Science.* Sage Publications, 2013.

[53] C. Z. Mooney, *Monte Carlo simulation*. Sage Publications, 2005.

[54] R. W. Hockney and J. W. Eastwood, *Computer simulation using particles*. Institute of Physics Publ., 1999.

[55] H. Rees, "Calculation of distribution functions by exploiting the stability of the steady state," *Journal of Physics and Chemistry of Solids*, vol. 30, no. 3, p. 643–655, 1969.

[56] W. H. Press, *Numerical recipes in C : the art of scientific computing*. Cambridge Univ. Press, 2007.

[57] F. Sabatti, "Cellular Monte Carlo simulation of coupled electron and phonon dynamics," Ph.D. dissertation, Arizona State University, 2018.

[58] D. K. Ferry, *Semiconductors: Bonds and Bands*. IOP Publishing, 2020.

[59] P. Devolder, J. Janssen, and R. Manca, *Basic stochastic processes*. Wiley ISTE, 2015.

[60] M. N. Ozisik, *Boundary Value Problems of Heat Conduction*. Mineola, NY: Dover Publications, 2013.

[61] D. Bleecker and G. Csordas, *Basic Partial Differenctial Equations*. Cambridge, MA: International Press of Boston, 2003.

[62] R. E. Peierls, *Quantum theory of solids*. Clarendon Pr., 1956.

[63] A. F. J Fourier, *The Analytical Theory of Heat*. London: Cambridge Press, 1878.

[64] J. Maassen and M. Lundstrom, "Steady-state heat transport: Ballistic-to-diffusive with Fourier's law," *Journal of Applied Physics*, vol. 117, no. 3, Jan 2015.

[65] P. Maycock, "Thermal conductivity of Silicon, Germanium, III-V compounds and III-V alloys," *Solid-State Electron*, vol. 10, pp. 161–168, 1967.

[66] M. Landolt and J. Bornstein, *Numerical Data and Functional Relationships in Science and Technology*. Berlin: Springer, 1987.

[67] A. Katz, *Indium Phosphide and Related Materials*. Boston: Artech House, 1992.

[68] S. Adachi, *Physical Properties of III-V Semiconductor Compounds*. Wiley, 1992.

[69] V. Palankovski and S. Selberherr, "Thermal models for semiconductor device simulation," in *3rd European Conference High Temperature Electronics*. IEEE, 1999, pp. 25–28.

[70] E. K. de Rivas, "On the use of nonuniform grids in finite-difference equations," *Journal of Computational Physics*, vol. 10, pp. 202–210, 1972.

[71] J. Thomas, *Numerical Partial Differential Equations: Conservation Laws and Elliptic Equations*, ser. Texts in Applied Mathematics. New York, NY: Springer, 1999, no. 33.

[72] R. Hickson, S. Barry, G. Mercer, and H. Sidhu, "Finite difference schemes for multilayer diffusion," *Mathematical and Computer Modelling*, vol. 54, no. 1-2, pp. 210–220, Jul 2011.

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

[74] "MATLAB R2018b," Natick, Massachusetts, United States, 2018.

[75] S. Gaur and D. Navon, "Two-dimensional carrier flow in a transistor structure under nonisothermal conditions," *IEEE Transactions on Electron Devices*, vol. 23, no. 1, p. 50–57, 1976.

[76] D. Sharma and K. Ramanathan, "Modeling thermal effects on MOS I-V characteristics," *IEEE Electron Device Letters*, vol. 4, no. 10, p. 362–364, 1983.

[77] E. Pop, S. Sinha, and K. Goodson, "Heat generation and transport in nanometer-scale transistors," *Proceedings of the IEEE*, vol. 94, no. 8, p. 1587–1601, 2006.

[78] C. Kittel, *Introduction to Solid State Physics*, 8th ed. Wiley, 2004.

[79] T. S. Fisher, *Thermal Energy at the Nanoscale*. World Scientific, 2014.

[80] J. C. Phillips, "Energy-band interpolation scheme based on a pseudopotential," *Physical Review*, vol. 112, no. 3, p. 685–695, Jan 1958.

[81] J. C. Phillips and L. Kleinman, "New method for calculating wave functions in crystals and molecules," *Physical Review*, vol. 116, no. 2, p. 287–294, 1959.

[82] J. R. Chelikowsky and M. L. Cohen, "Nonlocal pseudopotential calculations for the electronic structure of eleven diamond and zinc-blende semiconductors," *Physical Review B*, vol. 14, no. 2, p. 556–582, 1976.

[83] K. Kunc and O. Nielsen, "Lattice dynamics of zincblende structure compounds ii. shell model," *Computer Physics Communications*, vol. 17, no. 4, p. 413–422, 1979.

[84] M. Meneghini, G. Meneghesso, and E. Zanoni, *Power GaN Devices*. Springer International Publishing AG, 2016.

[85] J. Zúñiga-Pérez, V. Consonni, L. Lymperakis *et al.*, "Polarity in GaN and ZnO: Theory, measurement, growth, and devices," *Applied Physics Reviews*, vol. 3, no. 4, p. 041303, 2016.

[86] O. Ambacher, J. Majewski, C. Miskys *et al.*, "Pyroelectric properties of Al(In)GaN/GaN hetero- and quantum well structures," *Journal of Physics: Condensed Matter*, vol. 14, no. 13, pp. 3399–3434, Apr 2002.

[87] S. Yamakawa, S. Aboud, M. Saraniti, and S. M. Goodnick, "Fast full-band device simulator for wurtzite and zincblende GaN MESFET using a cellular Monte Carlo method," *Journal of Computational Electronics*, vol. 2, no. 2-4, pp. 481–485, 2004.

[88] K. T. Tsen, D. K. Ferry, A. Botchkarev, B. Sverdlov, A. Salvador, and H. Morkoç, "Direct measurements of electron-longitudinal optical phonon scattering rates in wurtzite GaN," *Applied Physics Letters*, vol. 71, no. 13, p. 1852, Jun 1998. [Online]. Available: https://aip.scitation.org/doi/10.1063/1.119420

[89] P. Altuntas, F. Lecourt, A. Cutivet *et al.*, "Power performance at 40 GHz of AlGaN/GaN high-electron mobility transistors grown by molecular beam epitaxy on Si(111) substrate," *IEEE Electron Device Letters*, vol. 36, no. 4, pp. 303–305, Apr 2015. [Online]. Available: http://ieeexplore.ieee.org/document/7042823/

[90] A. D. Latorre-Rey, K. Merrill, J. D. Albrecht, and M. Saraniti, "Assessment of self-heating effects under lateral scaling of GaN HEMTs," *IEEE Transactions on Electron Devices*, vol. 66, pp. 908–916, Feb 2019. [Online]. Available: https://ieeexplore.ieee.org/document/8601308/

[91] A. D. L. Rey, J. D. Albrecht, and M. Saraniti, "A π -shaped gate design for reducing hot-electron generation in gan hemts," *IEEE Transactions on Electron Devices*, vol. 65, no. 10, p. 4263–4270, 2018.

[92] J. Branlard, S. Aboud, P. Osuch, S. Goodnick, and M. Saraniti, "Frequency analysis of semiconductor devices using full-band cellular monte carlo simulations," *Monte Carlo Methods and Applications*, vol. 10, no. 3-4, 2004.

[93] D. Guerra, F. A. Marino, S. Goodnick, D. Ferry, and M. Saraniti, "Extraction of gate capacitance of high-frequency and high-power gan hemts by means of cellular monte carlo simulations," *International Journal of High Speed Electronics and Systems*, vol. 20, no. 03, p. 423–430, 2011.

[94] T. T. Nghiem, N. Trannoy, and J. Randrianalisoa, "Monte Carlo prediction of ballistic effect on phonon transport in silicon in the presence of small localized heat source," *Nanotechnology*, vol. 30, no. 41, p. 415403, 2019.

[95] K. Kukita, I. N. Adisusilo, and Y. Kamakura, "Monte carlo simulation of diffusive-to-ballistic transition in phonon transport," *Journal of Computational Electronics*, vol. 13, no. 1, p. 264–270, 2013.

[96] J. Kaiser, T. Feng, J. Maassen, X. Wang, X. Ruan, and M. Lundstrom, "Thermal transport at the nanoscale: A fouriers law vs. phonon boltzmann equation study," *Journal of Applied Physics*, vol. 121, no. 4, p. 044302, 2017.

[97] "Elusive diamond transistors could spark multifaceted electronics performance gains," Jan 2019. [Online]. Available: https://www.engr.wisc.edu/elusive-diamond-transistors-spark-multifaceted-electronics-performance-gains/

[98] H. Umezawa, "Recent advances in diamond power semiconductor devices," *Materials Science in Semiconductor Processing*, vol. 78, p. 147–156, 2018.

[99] T. G. Ivanov, J. Weil, P. B. Shah, A. G. Birdwell, K. Kingkeo, and E. A. Viveiros, "Diamond rf transistor technology with ft=41 ghz and fmax=44 ghz," *2018 IEEE/MTT-S International Microwave Symposium - IMS*, 2018.

[100] T. Watanabe, T. Teraji, T. Ito, Y. Kamakura, and K. Taniguchi, "Monte Carlo simulations of electron transport properties of diamond in high electric fields using full band structure," *Journal of Applied Physics*, vol. 95, no. 9, p. 4866–4874, 2004.

[101] F. Nava, C. Canali, M. Artuso, E. Gatti, P. F. Manfredi, and S. F. Kozlov, "Transport properties of natural diamond used as nuclear particle detector for a wide temperatue range," *IEEE Transactions on Nuclear Science*, vol. 26, no. 1, p. 308–315, 1979.

[102] F. Nava, C. Canali, C. Jacoboni, L. Reggiani, and S. Kozlov, "Electron effective masses and lattice scattering in natural diamond," *Solid State Communications*, vol. 33, no. 4, p. 475–477, 1980.

[103] L. Reggiani, S. Bosi, C. Canali, F. Nava, and S. F. Kozlov, "Hole-drift velocity in natural diamond," *Physical Review B*, vol. 23, no. 6, p. 3050–3057, 1981.

[104] D. G. Onn, A. Witek, Y. Z. Qiu, T. R. Anthony, and W. F. Banholzer, "Some aspects of the thermal conductivity of isotopically enriched diamond single crystals," *Physical Review Letters*, vol. 68, no. 18, p. 2806–2809, Apr 1992.

[105] S. Barman and G. P. Srivastava, "Temperature dependence of the thermal conductivity of different forms of diamond," *Journal of Applied Physics*, vol. 101, no. 12, p. 123507, 2007.

[106] W. Dai and R. Nassar, "A finite difference scheme for solving the heat transport equation at the microscale," *Numerical Methods for Partial Differential Equations*, vol. 15, no. 6, pp. 697–708, 1999.

[107] O. Ambacher, "Polarization induced effects in]AlGaN/GaN heterostructures," *Acta Physica Polonica A*, vol. 98, no. 3, p. 195–201, 2000.

[108] T. Luo *et al.*, "Gallium Arsenide thermal conductivity and optical phonon relaxation times from first principles calculations," *Europhysics Letters*, vol. 101, no. 1, 2013.

[109] K. Kramer and W. N. Hitchon, *Semiconductor Devices: A Simulation Approach*. Upper Saddle River, NJ: Prentice-Hall, 1997.

[110] Jacques Ruch, "Electron dynamics in short channel Field-Effect transistors," *IEEE Transactions on Electron Devices*, vol. 19, no. 5, pp. 652–654, 1972.

[111] Y. P. Zhao, J. R. Watling, S. Kaya, A. Asenov, and J. R. Barker, "Drift diffusion and hydrodynamic simulations of Si/SiGe p-MOSFETs," *Materials Science and Engineering B: Solid-State Materials for Advanced Technology*, vol. 72, no. 2, pp. 180–183, mar 2000.

[112] L. A. Hageman and D. M. Young, *Applied iterative methods*. Dover, 2004.

[113] M. Saraniti, A. Rein, G. Zandler, P. Vogl, and P. Lugli, "An efficient multigrid Poisson solver for device simulations," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 15, no. 2, p. 141–150, 1996.

[114] R. Y. Rubinstein and D. P. Kroese, *Simulation and the Monte Carlo method*. Hoboken, NJ: Wiley, 2017.

[115] U. Mishra, P. Parikh, and Y.-F. Wu, "Algan/gan hemts-an overview of device operation and applications," *Proceedings of the IEEE*, vol. 90, no. 6, p. 1022–1031, 2002.

[116] U. Mishra, S. Likun, T. Kazior, and Y.-F. Wu, "GaN-based RF power devices and amplifiers," *Proceedings of the IEEE*, vol. 96, no. 2, p. 287–305, 2008.

APPENDIX A

PUBLICATION LIST

Archival Journal Papers

[1] K. Merrill and M. Saraniti, "Electrothermal simulation of Π-gate devices," In Progress.

[2] K. Merrill and M. Saraniti, "Nonlinear Electro-Thermal Monte Carlo Device Simulation," J. Heat Transfer, vol. 142, no. 2, Feb. 2020.

[3] K. E. Merrill and G. T. Heydt, "The Calculation of Energy Dissipation in Metal Oxide Varistors for Power Distribution Applications," IEEE Trans. Power Syst., vol. 34, no. 5, pp. 3967–3969, Sep. 2019.

[4] A. D. Latorre-Rey, K. Merrill, J. D. Albrecht, and M. Saraniti, "Assessment of Self-Heating Effects Under Lateral Scaling of GaN HEMTs," IEEE Trans. Electron Devices, 2019.

Archival Conference Papers

[1] K. Merrill and M. Saraniti, "Multiscale Coupled Electro-Thermal Simulations of Electron Devices," in 2018 24rd International Workshop on Thermal Investigations of ICs and Systems (THERMINIC), 2018.

[2] A. D. Latorre-Rey, K. Merrill, J. D. Albrecht, and M. Saraniti, "Electro-thermal characterization of GaN HEMT on Si through self-consistent energy balance-cellular Monte Carlo device simulations," in 2017 IEEE Compound Semiconductor Integrated Circuit Symposium (CSICS), 2017.