Thermal Modeling of Wide and Ultra-wide Bandgap Materials and Devices Through

Cellular Monte Carlo

by

Joy Acharjee

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

Approved March 2024 by the
Graduate Supervisory Committee:

Marco Saraniti, Chair
Stephen M. Goodnick
Trevor Thornton
Robert Wang

ARIZONA STATE UNIVERSITY

May 2024

ABSTRACT

An efficient thermal solver is available in the Cellular Monte Carlo (CMC) that allows modeling self-heating in the electrical simulations, which treats phonons as flux and solves the energy balance equation to quantify thermal effects. Using this solver, thermal simulations were performed on GaN-HEMTs in order to test effect of gate architectures on the DC and RF performance of the device. A Π- gate geometry is found to suppress 19.75% more hot electrons corresponding to a DC power of 2.493 W/mm for V_{gs} = -0.6V (max transconductance) with respect to the initial T-gate. For the DC performance, the output current, I_{ds} is nearly same for each device configuration over the entire bias range. For the RF performance, the current gain was evaluated over a frequency range 20 GHz to 120 GHz in each device for both thermal (including self-heating) and isothermal (without self-heating). The evaluated cut off frequency is around 7% lower for the thermal case than the isothermal case. The work was extended to the study of ultra-wide bandgap material (Diamond), where isotope effect causes major deterioration in thermal conductivity. Simulations were performed for 0.001% (ultra pure), 0.1% and 1.07% isotope concentration (^{13}C) of diamond, showing good agreement with the experimental values. Further investigation was performed on the effect of isotope on the dynamics of individual phonon branches, thermal conductivity and the mean free path, to identify the dominant phonon branch. Acoustic phonons are found to be the principal contributors to thermal conductivity across all isotope concentrations with transverse acoustic (TA2) branch is the dominant branch with a contribution of 40% at room temperature and 37% at 500K. Mean free path computations show the lower bound of device dimensions in order to obtain maximum thermal conductivity. At 300K, the lowest mean free path (which is attributed to Longitudinal Optical phonon) reduces from 24nm to 8 nm for isotope concentration of 0.001% and 1.07% respectively. Furthermore, PETSc (Portable,

Extensible Toolkit for Scientific Computation) developed by Argonne National Lab, was included in the existing CMC device simulator as a Poisson solver to further extend the capability of the simulator.

TABLE OF CONTENTS

CHAPTER   Page

LIST OF FIGURES

Chapter 1

INTRODUCTION

Heat management and the associated dissipation techniques are crucial components of device engineering as trapped heat can affect not only the performance but also the reliability of the device [9]. As the semiconductor industry moves into more complex and dense circuits, the need for thermal management will become more pronounced.

For electrical modeling, Monte Carlo (MC) methods have been widely used [10]. MC methods can provide an accurate solution of the Boltzmann Transport Equation (BTE) compared to techniques based on lower order moments of BTE, like the drift-diffusion and hydrodynamic methods. These techniques are applicable for near-equilibrium conditions, while the MC method can provide full stochastic solution at the expense of simulation time or hardware requirements.

Omission of self-heating models can lead to inaccurate prediction of electrical characteristics. The inclusion of a thermal solver into MC algorithms has been performed in only a few studies [11, 12]. In these studies, both the interactions between electron and optical phonons as well as optical-acoustic phonons are modeled using a relaxation time approximation (RTA) [13]. In this study, the electron-phonon interaction is directly computed from the scattering rates and the phonon-phonon interactions are modeled using RTA. Furthermore, this study both the temperature-dependent scattering rates and the temperature dependent thermal conductivity are taken into account. Additionally, the thermal solver is considered self-consistently coupled, as the scattering rates get updated according to the local temperature obtained in each iteration inside the forcing function.

For the thermal modeling of GaN HEMT, this study uses a flux-based approach,

where phonons are modeled as flux. In this approach, energy balance equation (EBE) of phonon is solved with the help of Kirchhoff transformation similar as performed by Pilgrim *et al.* [14]. While Pilgrim *et al.* considers optical phonons as dispersionless, this study takes full phonon dispersion into account. The flux-based thermal solver is then used to obtain detailed acoustic and optical temperature maps of the HEMT. The CMC allows to compute energy distribution function of the electrons precisely, which can lead to study of hot electron suppression for the device geometries of interest.

In the thermal conductivity analysis of bulk diamond, phonons are modeled as particles, where the code is capable of modeling non-equilibrium phonon distributions. A precomputed scattering table is used during runtime to analyze the different states of scattering. Based on the local phonon distribution, a rejection algorithm is then employed to compute the scattering rates. In this approach, both Umklapp and Normal process [15] are treated equally. The particle-based thermal solver is used to determine the effect of isotope concentration in deterioration of the thermal conductivity of diamond. To investigate further, different thermal properties like mean free path, phonon modewise thermal conductivity were also obtained.

Chapter 2 provides a detailed overview of the Boltzmann Transport Equation along with its incorporation into cellular Monte Carlo (CMC) framework. In chapter 3, PETSc is presented as an external multi-purpose Poisson solver to be used in CMC. Chapter 4 discusses the effect of gate geometries in GaN HEMT through the solution of energy balance equation and coupling it to electron dynamics. Chapter 5 presents the thermal conductivity analysis of diamond performed through CMC. Finally, chapter 6 provides a summary and discusses future work planned.

Chapter 2

MODELING OF CARRIER TRANSPORT

2.1   Introduction

As the feature sizes reduces to sub-nanometer regime, device modeling and simula-
tion has become an integral part of semiconductor research & development for several
key reasons: 1) understanding the newer physical phenomenons introduced by shorter
dimensions, 2) optimization of device performance through trial-and-error in fabrica-
tion is nearly impossible due to time and cost and 3) showing insights about devices
yet to be manufactured along with physically immeasurable attributes [1, 16]. For
these reasons, arduous efforts have been taken to capture the physics of semiconductor
devices along with minimizing the computation time. In any modern device simulator,
the charge transport and scattering is coupled with the field equations solver along
with proper boundary conditions.
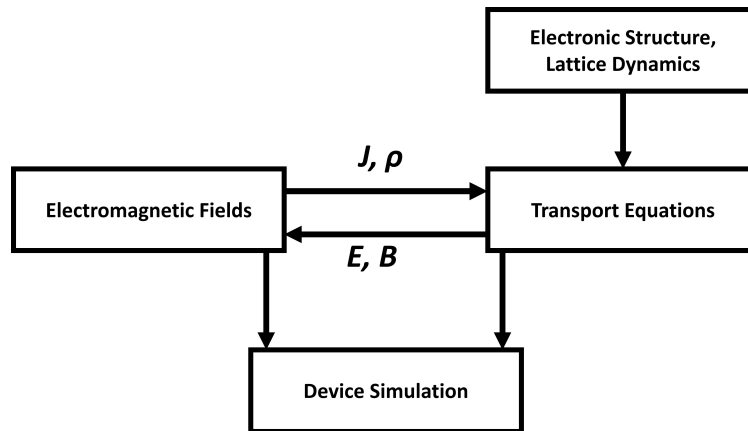
Figure 2.1: Generic diagram of a device simulation code [1].

Figure 2.1 shows the generic components of a device simulation tool [1] with two

self-consistently coupled components: the transport equation solver which models the charge dynamics and field solver which drives the flow of charge.

Initial efforts for modeling transport mostly involved the drift-diffusion (DD) equation:

$$\mathrm{J}_n = nq\mu_n E + qD_n \nabla n \tag{2.1}$$

where J_n, n,q,μ_n, E, D_n represent current density, charge, mobility, electric field and diffusion coefficient, respectively. The values of transport parameter μ_n and D_n were experimentally measured and it worked until the device dimensions started shrinking. For modern devices, these parameters not only depend on material and field, but also on the structure of the device and the microscopic physics [17]. Furthermore, quantum mechanical treatment of the scattering mechanisms is also necessary. Hence, most of the device simulation tools revolves around the solution of Boltzmann Transport Equation (BTE) [18], where the particle distribution function $f(\mathbf{r}, \mathbf{k}, t)$ completely defines the state of the system.

In this study two types of simulation models have been used: 1) modeling phonons as flux and 2) treating them as particles.

2.2   Boltzmann Transport Equation

The particle distribution function, $f(\mathbf{r}, \mathbf{k}, t)$, represents the probability of finding electron at a specific time t, position $\mathbf{r}$ and momentum $\mathbf{k}$. As the Fermi-Dirac distribution is a function of only the energy of particles not momentum, the distribution function becomes symmetric in momentum space (k-space) [19]. The change in distribution function due to scattering can be represented as:

$$\frac{\partial f(\mathbf{r}, \mathbf{k}, t)}{\partial t} = \left.\frac{\partial f}{\partial t}\right|_{\mathrm{coll}} \tag{2.2}$$

If the left hand side of the equation above is expanded, it becomes:

$$\frac{\partial f}{\partial t} = -\nabla_{\mathbf{r}} \cdot \left[\frac{d\mathbf{r}}{dt}f\right] - \nabla_{\mathbf{k}} \cdot \left[\frac{d\mathbf{k}}{dt}f\right] + \left.\frac{\partial f}{\partial t}\right|_{\text{coll}} \tag{2.3}$$

The momentum space flux is called "force term" which can be modified as $-\nabla_{\mathbf{r}} \cdot \left[\frac{d\mathbf{r}}{dt}f\right] = -\frac{1}{\hbar}\mathbf{F} \cdot \nabla_k f$, where $\mathbf{F} = \hbar\frac{d\mathbf{k}}{dt} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B})$, where $\mathbf{E}$, $\mathbf{B}$ and $\mathbf{v}$ stands for the electric field, magnetic flux and velocity of the particles, respectively. The real-space flux results from any change in concentration of the particles or temperature gradients, which leads to a diffusion process. The collision integral represents the change in distribution function due to collision or scattering events. This term is computed quantum mechanically accounting for the nature of the semi-classical approach. This term can be computed from the difference in in-scattering (particles moving inside the cell after scattering) and out-scattering (particle moving outside the cell after scattering) such as:

$$\left(\frac{\partial f}{\partial t}\right)_{\text{coll}} = \sum_{k'} \{S\left(\mathbf{k}', \mathbf{k}\right) f\left(\mathbf{k}'\right)\left[1 - f(\mathbf{k})\right] - S\left(\mathbf{k}, \mathbf{k}'\right) f(\mathbf{k})\left[1 - f\left(\mathbf{k}'\right)\right]\}$$

The analytical solution of BTE becomes much difficult due to presence of both $f(\mathbf{k})$ and $f(\mathbf{k}')$ at the same time, which makes it a complex integro-differential equation of the distribution function and requires a series of approximation to solve it [1].

The modified form of BTE stands as:

$$\frac{\partial f}{\partial t} = -\frac{1}{\hbar}\mathbf{F} \cdot \nabla_k f - \mathbf{v} \cdot \nabla_r f + \left.\frac{\partial f}{\partial t}\right|_{\text{coll}} \tag{2.4}$$

There are certain conditions need to be met for this equation [19]:

1) The distribution should evolve on the order of relaxation time of the carriers (mean free time between collisions), so the time variation of the distribution function should be really slow.

2) The force term should be small enough so that wave functions from different bands of the carrier don't mix and it can still be treated semiclassically.

3) The dimensions of the particles should tend to zero but the mean free path should be finite, this is known as Boltzmann-Grad limit [20]. For this reason, only binary collisions are considered, while collisions involving more particles are ignored.

2.3   Applicability of BTE for Phonons

Phonon is a discrete unit of vibrational energy in solids, that accounts for heat and sound transport in crystals. Phonon in any materials vibrate with random frequency and velocity at room temperature, but at elevated temperature they can oscillate at specific frequency together. For computing finite thermal conductivity 3 particle scattering process needs to be accounted for [3] in the collision term. As the BTE derived above is applicable only to binary collisions, modifications need to be done to make it applicable for phonons. Peierls *et. al.* [21] introduced a statistical heat transport equation that accounts for non-binary collisions for particles, which is called phonon BTE or Peierls-BTE (PBTE). Two additional assumtions are necessary for PBTE: 1) the states of phonons are quantized, and 2) all information about the system must be contained in terms of the occupation number of the particles. The phonon BTE can be represented as:

$$\frac{\partial f}{\partial t} + \frac{\partial \mathbf{r}(\mathbf{q_l})}{\partial t} \cdot \nabla_r f = \left. \frac{\partial f}{\partial t} \right|_{C_{coll}} \tag{2.5}$$

where $\mathbf{q_l}$ denotes the wavevector of the lth phonon mode. If the indexes of the modes are omitted for clarity, the collision integral can be expressed as:

$$\left. \frac{\partial f}{\partial t} \right|_{\text{coll}} = \sum_{m=1}^{M} \sum_{\mathbf{q}^1,\ldots,\mathbf{q}^m \subset \text{BZ1}} \Gamma_d \left(\mathbf{q}, \mathbf{q}^1, \ldots, \mathbf{q}^m, f\right) - \Gamma_r \left(\mathbf{q}, \mathbf{q}^1, \ldots, \mathbf{q}^m, f\right) \tag{2.6}$$

where M is the total number of phonon states, $BZ1$ is the first Brillouin Zone [22], $\Gamma_d(\mathbf{q}, \ldots)$ and $\Gamma_r(\mathbf{q}, \ldots)$ are the decrease and increase of the phonon population in state $\mathbf{q}$, respectively.

Similar to the original BTE, for obtaining a close-form solution of this equation severe approximations are needed either on the collision integral (relaxation time approximation [23]) or on the distribution function (near equilibrium approximation [24]), or on both. Therefore, both deterministic and stochastic algorithms have been implemented [25–29] to obtain numerical expressions of the distribution function. In deterministic approaches, only lower order moments of distribution functions are used along with different approximations.

2.4 Drift-Diffusion (DD) Model

The drift-diffusion model was introduced by Van Roosbroeck in 1950 [30] and has been one of the popular methods for device simulation in the last century. As the solution of BTE is complicated, particularly when coupled with field solvers for the simulation of a device, traditional device simulations involved around solving the drift-diffusion (DD) equations. In drift-diffusion model, the driving forces ($\mathbf{E}$ field and gradient of carrier density) are considered local, which means the current at any point is only a function of field and concentration gradient of that point [1].

2.4.1 Derivation of DD equations from BTE

For any continuous random variable r, if the probability distribution function is g(r), then n^{th} moment can be found as [31] :

$$\langle R^n \rangle = \int_{-\infty}^{\infty} r^n g(r) dr$$

Instead of $g(r)$, carrier distribution function can be used along with a nth order function of wavevector, k, $\phi(\mathbf{k}^n)$. The zeroth moment of the carrier distribution function can be obtained as:

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

The carrier density,n, can be found by dividing the distribution function by the unit volume of reciprocal space, Ω, such as:

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

Similarly, the zeroth moment of BTE can be obtained by multiplying Equation 2.4 by $\phi(\mathbf{k})/\Omega$ and integrating over momentum space such as [1]:

$$\frac{1}{\Omega}\frac{\partial}{\partial t}\sum_{\mathbf{k}}\phi(\mathbf{k})f = -\frac{1}{\Omega}\nabla_{\mathbf{r}}\cdot\sum_{\mathbf{k}}\phi(\mathbf{k})\mathbf{v}f + \frac{\mathbf{F}}{\Omega\hbar}\sum_{\mathbf{k}}f\nabla_{\mathbf{k}}\phi(\mathbf{k}) + \frac{1}{\Omega}\sum_{\mathbf{k}}\phi(\mathbf{k})\frac{\partial f}{\partial t}\bigg|_{\mathrm{coll}} \quad (2.7)$$

Setting $\phi(\mathbf{k}^0){=}1$, the zeroth moment can be obtained as:

$$\frac{\partial n}{\partial t} = \nabla_{\mathbf{r}} \cdot F_{\phi,0} + \mathcal{S}_n, \quad (2.8)$$

where $F_{\phi,0}$ is the carrier flux and $\mathcal{S}_n$ is the change in distribution function due to scattering events. The value of the flux can be obtained from: $F_{\phi,0} = -\frac{\mathbf{J}_n}{q}$, where $\mathbf{J}_n$ is the current density and q is the charge of the carrier.

Similarly, for the 1st order of moments, one has to set $\phi(\mathbf{k}^1){=}u(\mathbf{k})$, where $u(\mathbf{k})$ is a function of carrier velocity. After some manipulation the end result of the first moment can be found as:

$$\frac{\partial \mathbf{J}_n}{\partial t} = \frac{2q}{m^*}\nabla_{\mathbf{r}} \cdot \mathbf{W}_n + \frac{nq^2\mathbf{E}}{m^*} - \mathcal{S}_m \quad (2.9)$$

where m^* is the effective mass of the carrier, $\mathbf{E}$ is the electric field, $\mathbf{W}_n$ is the energy of the carrier and $\mathcal{S}_m$ is the change in momentum due to scattering events. For parabolic band structure, $\mathbf{E} = \hbar^2\mathbf{k}_i^2/2m^*$ and $\mathbf{W}_n = 1/\Omega\sum_{\mathbf{k}}E(\mathbf{k})f$.

For obtaining DD equations, Equation 2.8 and Equation 2.9 should be combined with the Poisson's equation. In DD model, the total energy of the carrier comes from the thermal energy as:

$$W = \frac{3}{2}nk_BT_C$$

8

where k_B is the Bolzmann's constant and T_C is the temperatur of the carrier. W has a equal contribution from x,y and z, and so $\mathbf{W}_n = W/3$. Replacing this into Equation 2.9, results in:

$$\frac{\partial \mathbf{J}_n}{\partial t} = \frac{2}{3}\frac{1}{m^*}\nabla_r W + \frac{nq^2}{m^*}\mathbf{E} - \mathcal{S}_m \qquad (2.10)$$

The term, $\mathcal{S}_m$ can be obtained by a relaxation time approximation: $\mathcal{S}_m = \frac{\mathbf{J}_n}{\tau_m}$, where τ_m is the relaxation time. Replacing this into Equation 2.10:

$$\mathbf{J}_n = \frac{\tau_m nq^2}{m^*}\mathbf{E} + \frac{2}{3}\frac{\tau_m}{m^*}\nabla_r W = \frac{\tau_m nq^2}{m^*}\mathbf{E} + \frac{\tau_m}{m^*}k_B T_C \nabla n + \frac{\tau_m}{m^*}k_B n \nabla T_C \qquad (2.11)$$

where μ_n is the mobility of electrons, D_n is the diffusion coefficient and D_T is the thermal diffusivity. The values of these quantities are: $\mu_n = (q\tau_m)/(m^*)$, $D_n = \mu_n k_B T_c/q$ and $D_T = \mu_n k_B n/q$. Replacing these terms into Equation 2.11, the final equation for the DD model for electrons can be obtained as:

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

Finally, similar equations can be obtained for holes and the final set of equations of DD model can be shown as:

Current equations:
$$\mathbf{J}_n = q\mu_n n\mathbf{E} + qD_n\nabla_r n + qD_T\nabla T_C,$$
$$\mathbf{J}_n = q\mu_p p\mathbf{E} - qD_p\nabla_r p - qD_T\nabla T_C.$$

Continuity equations:
$$\frac{\partial n}{\partial t} = \frac{1}{q}\nabla \cdot \mathbf{J}_n + \mathcal{S}_n,$$
$$\frac{\partial p}{\partial t} = -\frac{1}{q}\nabla \cdot \mathbf{J}_p + \mathcal{S}_p.$$

Poisson equation:
$$\nabla \cdot \varepsilon \nabla V = -\left(p - n + N_p^+ - N_A^-\right)$$

The continuity equation actually represents the conservation of the charge carriers. Any numerical technique should involve the following criterions for solving DD equations:

1) The amount of the total charge inside the device should be conserved.

2) density of carriers should be always positive.

3) The solution should be monotonous and shouldn't introduce any unwanted oscillations.

Scharfetter and Gummel introduced the numerical solution of DD equations [32], where they introduced a robust 1D discretization technique of DD equation. Later, Bank *et. al.* introduced a finite volume Scharfetter-Gummel method for higher-dimensional problems [33, 34]. This method is still one of the common techniques for semiconductor device simulation. On the other hand, different techniques have been adopted to solve DD equations with finite element method as well, such as : exponential fitting [35, 36], mixed finite element methods [37] and stabilization techniques [38, 39].

Along with the scaling of semiconductor devices into submicron regime, the assumptions on which the DD model stands don't hold true anymore. The high electric field over a small length of the device initiates nonlocal and hot-carrier effects, at the same time these effects become determinant factors for device performance. Different extensions have been adopted to make the DD model more applicable for smaller devices, such as: addition of a balance equation for average energy of carriers [40] and introduction of a term in current equation which is proportional to temperature gradient [23]. Although various extensions of the DD models have been performed, such as: addition of an balance equation for addressing average energy of carriers [40] and addition of a new term in current equation which is proportional to temperature gradient [23]. But still the limitations make DD model inadequate for modern device simulations.

2.5   Hydrodynamic Model

The hydrodynamic (HD) model is derived from the first 3 (0th, 1st and 2nd) moments of the distribution function in BTE. The necessity of a model other than DD for device simulation was first felt while analyzing the current drive capability of scaled n-MOSFET around 1970 [1]. Under the DD model, it was believed that the velocity of the carriers injected from the source to the channel would saturate and stay at that value for any **E**-field higher than the onset value for saturation. Some theoretical works like [41], reported velocity overshoot in Si raising questions about the validity of the DD model. Later experimental works confirmed observing velocity overshoot in Si as well [42, 43].

Apart from the fact that Hydrodynamic (HD) model takes macroscopic physical effects into account than DD model, the process of modeling physical parameters is another crucial advantage for HD models. In HD model, information about different parameters like mobility and energy relaxation time, are extracted from the temperature of the carrier. On the other hand, in DD model, the carrier temperature is set to be equal to the lattice and so these parameters need to be included in a different manner. Physical processes like impact ionization depend on distribution function of the particles rather the **E**-field, that's why to explain these, a local energy model models is better suited than a local **E**-field model. For these reasons mentioned above, efforts for formalizing HD model were initiated.

As HD model is not the primary method used for this study, the author intends to discuss only a brief overview of the HD model, rather than going through the step-by-step derivation from the BTE. To take advantage of the symmetric nature of the distribution function, Stratton *et.al.* [40] divided the distribution function into even and odd parts and thus creating two coupled equations for BTE. Later, a

quantity $\phi(\mathbf{k})$ is defined, which can have values $1, \mathbf{k}, \ldots$, etc., and the total averaged value of $\phi(\mathbf{k})$ can be obtained from:

$$n_\phi(\mathbf{r}, t) = \frac{1}{V} \sum_{\mathbf{k}} \phi(\mathbf{k}) f(\mathbf{r}, \mathbf{k}, t)$$

where n_ϕ can be represented as current density or carrier density. To incorporate the change in n_ϕ due to generation-recombination process, another term, $S(\mathbf{r}, \mathbf{k}, t)$, can be represented as:

$$S_\phi(\mathbf{r}, t) = \frac{1}{V} \sum_{\mathbf{k}} \phi(\mathbf{k}) s(\mathbf{r}, \mathbf{k}, t)$$

After some manipulation, the balance equation for n_ϕ can be written as follows:

$$\frac{\partial n_\phi}{\partial t} = -\nabla \cdot F_\phi + G_\phi - R_\phi + S_\phi \tag{2.12}$$

where F_ϕ is the flux associated with n_ϕ, G_ϕ and R_ϕ are the generation and re-combination terms for collision events, respectively. The HD model is comprised of 3 essential balance equations: 1) Carrier-density balance equation, 2) momentum balance equation and 3) energy balance equation.

Carrier-density balance equation: If $\phi(k)=1$, *i.e*, the zeoth order moment in k, then n_ϕ in Equation 2.5 represents the electron density, n. Furthermore, the flux F_ϕ can be replaced by $-\frac{1}{e} \mathbf{J}_n(\mathbf{r}, t)$. If only redistribution of carriers among states is considered due to scattering and electric-field rather than generation or recombination, the G_ϕ and R_ϕ can be neglected and the carrier-density balance equation can be represented as:

$$\frac{\partial n}{\partial t} = \frac{1}{e} \nabla \cdot \mathbf{J}_n + \mathcal{S}_n \tag{2.13}$$

which is also known as the continuity equation for the electrons, simply denoting the conservation of electrons in the entire system.

Momentum balance equation: Following similar approach and introducing a moment of order 1 in k $\phi(k) = k_z$, the balance equation can be formulated as follows:

$$\frac{\partial \mathbf{J}_n}{\partial t} = \frac{2q}{m^*} \nabla_\mathbf{r} \cdot \mathbf{W}_n + \frac{nq^2 \mathbf{E}}{m^*} - \mathcal{S}_m \tag{2.14}$$

where $\mathbf{W}_n$ is a component of the kinetic energy tensor. $\mathcal{S}_m$ is the momentum relaxation term, which can be obtained from the relaxation time approximation (RTA) as : $\mathcal{S}_m = \frac{\mathbf{J}_n}{\tau_m}$, where τ_m is the characteristic relaxation time. Under the following approximations, HD model can be simplified into DD model:

1) Drift energy of carrier is much smaller than thermal energy.

1) There is no temperature gradient in the system.

1) To obtain extended DD model (where the mobility and the diffusion coefficients are field dependent), these quantities should depend on local field only.

Energy balance equation: Again taking a moment of second order in k $\phi(k) = E_k$, the energy balance equation is derived as:

$$\frac{\partial W}{\partial t} = -\nabla_r \cdot \mathbf{F}_W + \mathbf{E} \cdot \mathbf{J}_n + \left.\frac{\partial W}{\partial t}\right|_{\text{coll}} \tag{2.15}$$

where W is the energy density and $\mathbf{F}_W$ is the energy flux. The collision term can be approximated using RTA such as:

$$\left.\frac{\partial W}{\partial t}\right|_{\text{coll}} = \frac{W - W_0}{\tau_W}$$

where W_0 is the equilibrium energy density and τ_W is the relaxation time needed for W to approach towards equilibrium. A similar set of equations can be obtained for holes. HD models are usually faster than Monte Carlo (MC) methods particularly in 2D simulations. However, in MC the parameters like effective mass and carrier mobility are the output of the simulation, whereas in HD model they are modeled as input [2] in balance equations.

HD model has been widely used in different illustrations [44–46]. An interesting and effective extension of electron energy balance equation is energy balance equation for phonons, which can represent the heat transfer in solid quite efficiently when coupled with electron Monte Carlo, as it will be discussed extensively in Chapter 4.

2.6 Necessity of a stochastic approach such as the Monte Carlo method

As the device dimensions continue to shrink, a continuous nature of the current can no longer capture the underlying physics anymore. Rather, a granular nature representing the charge particles that leads to the charge density in the active region of the device should be accounted for [1] as it models carrier dynamics more accurately. Obviously, simulations tools like DD or HD can't explain the operation of a single electron as in the single electron transistors [47]. For the case of solid and gas, Molecular Dynamics has been an effective tool, where the individual motion of the particles are tracked using Newton's laws of motion [48]. But for semiconductors, scattering events dominate the transport by altering the energy and momentum of the carriers, which needs to be modeled accurately for an acceptable result.

Furthermore, the DD and HD model are based on the lower order moments of the BTE, while the MC approach can provide a solution of the BTE (*i.e.* considering all the moments) that is exact in statistical terms. In addition, MC methods need fewer assumptions and can provide the transport parameters as output, while they are needed as closure relations in DD or HD models. The modern MC method was first used by Metropolis and Ulam [49] while they were working on neutron diffusion in the core of a nuclear weapon. Since then, it has been widely used for different applications, including semiconductor device simulations, where equilibrium conditions can't be met due to high **E**-fields [50].

2.7 Ensemble Monte Carlo (EMC) method for semiconductors

Usually, in particle-based Monte Carlo simulations, individual particles are tracked in real and momentum space and they go through scattering processes to achieve steady-state condition. In ensemble Monte Carlo (EMC), a synchronous ensemble of particles is used. These particles are called super-particle, since they represent a definite number of real particles.

The basic principle of the EMC method evolves around the generation of a random walks through random number generating algorithms, in order to simulate the collision events of the carriers. At the end of free motion of the particle, which is commonly known as free flight, the particles encounter random collisions with lattice vibrations (*i.e.* phonons), impurities, isotopes etc. The stochastic MC method chooses the type and value of the scattering, finds the new energy and momentum of the particles, updates the position of the particle in phase space and repeats the same procedure for the subsequent free flight. Motion of the particles are sampled after a specific time step, in order to obtain different physical parameters of the material [1].

Figure 2.2 represents the process flow of a general ensemble Monte Carlo device simulator, consisting of key stages such as: initialization of the system, free flight and tracking of the particles, modeling of the scattering events, solving the field equations and collecting and processing the output data. During the initialization, the carriers are initialized with a Maxwellian velocity distribution according to the initial temperature defined in the simulation [2]. Moreover, they are initially distributed according to the defined distribution to ensure charge neutrality, such that even with the presence of surface charge or doping levels, the overall charge is zero. With the progression of the simulation, the charge density is calculated from the position of the carriers.
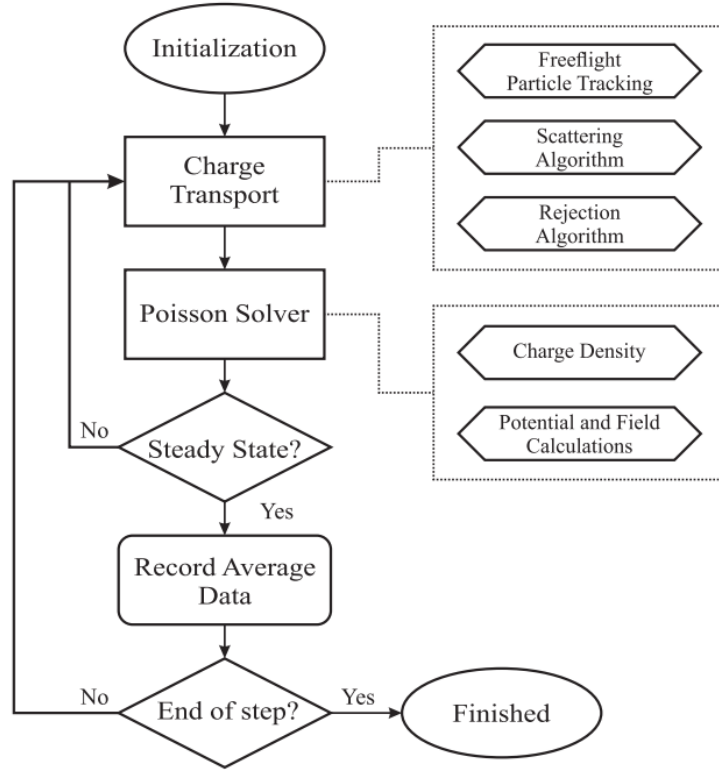
Figure 2.2: Flowchart representing different steps of a generic ensemble Monte Carlo device simulator [2].

Taking the charge density as input, the Poisson solver returns the electrostatic potential used to compute the electric field in each grid point of the simulation domain. Grid size determinant for the accuracy of the simulation. Two time-steps that are crucial for determining the accuracy and length of the simulation: Poisson and free flight time-step. The Poisson time-step, is the time interval between two successive Poisson solver calls. The value of this time-step is chosen smaller than $2/\omega_p$, where ω_p represents the plasma frequency such as [51]:

$$\omega_p = \sqrt{\frac{ne^2}{\epsilon_r\epsilon_0 m_{eff}}} \tag{2.16}$$

where n, e, ϵ_r and m_{eff} are the highest concentration of electron in simulation domain, electron charge, the dielectric constant and the effective mass of electron, respectively.

The positions of the particles are updated according to the **E**-field following a ray tracing algorithm.

At the end of each free flight time-step, a random number is generated for each carrier in the simulation domain in order to check if the carrier will go through a scattering event. There are two conditions that need to be met while choosing the duration of the free flight time-step: 1) should be smaller than the inverse of the highest scattering rate and 2) the free-flight time-step should be smaller than the Poisson time-step.

The stochastic Monte Carlo technique is employed to model the scattering as Markovian process, where the wavevector of the particle changes instantaneously. The scattering rates are pre-computed based on characteristics of the materials.

The usual scattering mechanisms considered in this Monte Carlo study are: 1) deformation potential scattering, 2) piezoelectric polar optical scattering and 3) impact ionization. The following discussions provide a brief discussion about these 3 scattering mechanisms.

Deformation potential scattering: Deformation potential scattering refers to the scattering due to non-polar interactions between electron and phonon. If an electron with a wave vector $\vec{k}$ interacts with a phonon with wave vector $\vec{q}$, the final wave vector will be $\vec{k'} = \vec{k} \pm \vec{q}$, where $+$ and $-$ indicate the absorption and emission of phonon, respectively. Hence, the deformation potential scattering rate from a point $\vec{k}$ in band ν to a region $\Omega_{\vec{k'}}$, centered around another point $\vec{k'}$ in band ν' can be derived as [10]:

$$P_{\nu\nu',\eta}^{def}\left(\vec{k}, \Omega_{\vec{k'}}\right) = \frac{\pi}{\rho\omega_{\eta\vec{q}}} \left|\Delta^{(\eta)}\left(\nu', \vec{k}, \vec{q}, \nu\right)\right|^2 \left|I\left(\nu, \nu', \vec{k}, \vec{k'}\right)\right|^2 D_{\nu'}\left(\epsilon', \Omega_{\vec{k'}}\right)\left(N_{\eta\vec{q}} + \frac{1}{2} \mp \frac{1}{2}\right)$$

(2.17)

where, $\Delta^{(\eta)}\left(\nu', \vec{k}, \vec{q}, \nu\right)$ is for the deformation potential [10], $I\left(\nu, \nu', \vec{k}, \vec{k'}\right)$ represents the overlap integral [10] and $D_{\nu'}\left(\epsilon', \Omega_{\vec{k'}}\right)$ is the density of states [10]. In this study,

17

Rigid-pseudo ion model [52] is usually used to calculate the deformation potential, where only full band structure and dispersion relation are necessary.

Piezoelectric Scattering: In polar materials, if the lattice is distorted due to phenomena like thermal fluctuations, the local charge neutrality can be altered. This can produce electric polarization along with the motion of electron in a macroscopic **E**-field [53]. If the disturbance of the motion of electron arises from an acoustic mode, it is called piezoelectric scattering. On the other hand, interactions with optical mode is called polar optical scattering (see below). The piezoelectric scattering rate can be represented as [53]:

$$P^{\text{piezo}}_{\nu\nu'}\left(\vec{k}, \Omega_{\vec{k}'}\right) = \frac{2\pi}{\hbar} K_{av}^2 \frac{e^2 k_B T}{q^2 \epsilon^*} \left| I\left(\nu, \nu', \vec{k}, \overrightarrow{k'}\right) \right|^2 D_{\nu'}\left(\epsilon', \Omega_{\vec{k}'}\right) \qquad (2.18)$$

where $I\left(\nu, \nu', \vec{k}, \overrightarrow{k'}\right)$ and $D_{\nu'}\left(\epsilon', \Omega_{\vec{k}'}\right)$ have similar meaning as in Equation 2.17, and K_{av}^2 is a coupling constant which can be represented as:

$$K_{av}^2 = \frac{\langle e_l^2 \rangle}{\epsilon^* c_l} + \frac{\langle e_t^2 \rangle}{\epsilon^* c_t}$$

where e and c are piezoelectric and elastic constants, respectively and the superscripts are for longitudinal and transverse waves. Finally $\langle . \rangle$ represents the spherical average [54].

Polar optical scattering: The interaction between electron and optical phonons is modeled following the approach as described by Lee *et. al.* [55], where uniaxial model and dielectric continuum model of Loudon [56] is used to derive the Fröhlich Hamiltonian and the scattering rates. The quantized form of the interaction Hamiltonian can be represented as follows:

$$H = \sum_{\overrightarrow{q}} -e\Phi(\overrightarrow{q})e^{i\overrightarrow{q}\cdot\mathbf{r}}\left(a_{\overrightarrow{q}} + a^{\dagger}_{-\overrightarrow{q}}\right),$$

where $\Phi(\overrightarrow{q})$ is the Fourier transform of the electrostatic potential, $\Phi(\mathbf{r})$, $a_{\overrightarrow{q}}$ and $a^{\dagger}_{-\overrightarrow{q}}$ are the annihilation and creation operator. The probability of electrons transitioning

from state $\overrightarrow{k}$ to $\overrightarrow{k'}$ per unit time can be calculated following Fermi's golden rule as:

$$W\left(\overrightarrow{k}, \overrightarrow{k'}\right) = \frac{2\pi}{\hbar} \left|M_{\overrightarrow{k}-\overrightarrow{k'}}\right|^2 \delta\left(E_{\overrightarrow{k'}} - E_{\overrightarrow{k}} \pm \hbar\omega_{\overrightarrow{k}-\overrightarrow{k'}}\right)$$

where $E_{\overrightarrow{k}}$ represents the energy of electron, $M_{\overrightarrow{k}-\overrightarrow{k'}}$ is the transition matrix built from the Hamiltonian, $\hbar\omega_{\overrightarrow{k}-\overrightarrow{k'}}$ is the energy needed for the transition and $\pm$ represents the emission and absorption of the phonon, respectively. Finally, the scattering rate $W(\overrightarrow{k})$ can be found as:

$$W(\overrightarrow{k}) = \sum_{\overrightarrow{k'}} W\left(\overrightarrow{k}, \overrightarrow{k'}\right).$$

Impact Ionization: Impact ionization rate is modeled based on two power-laws [57] such as:

$$\Gamma_{IMP}(E) = \sum_{i}^{2} \theta\left(E - E_{0i}\right) a_i \left[E - E_{0i}\right]^{r_i} \tag{2.19}$$

where $E_{0i} \geq E_{gap}$ is the threshold energy, θ is the step function, while the prefactor a_i and exponent r_i both are fitting parameters. Impact ionization happens at energies over E_{0i} and the energy of the ionized carrier is reduced by the amount of the band-gap.

2.8   Cellular Monte Carlo (CMC)

An improved and faster extension of the EMC approach was introduced by Saraniti *et. al.* [58], which mainly differs in the selection process of the final state of the carrier after scattering. In the traditional EMC, the following algorithms are executed during runtime: 1) after scattering, all final states inside first Brillouin Zone (BZ1) conserving energy and momentum are identified, 2) the scattering rate is computed for each final state and 3) the final state is chosen by a stochastic selection process. Scatterings in semiconductors are really common and all the runtime computations makes the traditional EMC computationally very expensive.

On the other hand, the CMC algorithm pre-computes all scattering rates from each initial state to every possible final states inside BZ1 for each scattering mechanism and
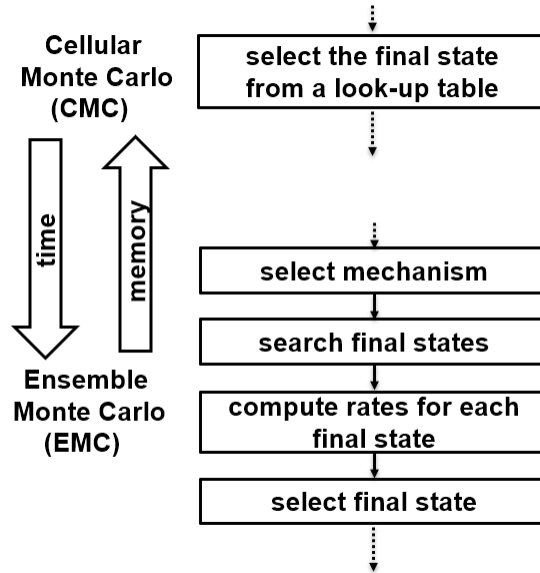
Figure 2.3: Algorithmic differences between EMC and CMC.

material. These scattering rates are stored in a large look-up table that is stored into RAM at the initialization stage of the simulation. Although the size of the table may be large depending upon the number of scattering mechanisms considered, the current availability of large memory makes the approach more than realistic. As a result, the CMC code is significantly more efficient for complex device simulation compared to EMC. Figure 2.3 illustrates the key algorithmic differences between EMC and CMC.

2.8.1   Structure of transition table (look-up table) in CMC

Figure 2.4 shows a schematic diagram of the structure of a typical transition table as it is used in CMC. Here, all the initial states are stacked up in a vertical array and each of them is connected to a horizontal array (ordered list), which represents the final states along with their corresponding scattering rate. After a scattering event, the final state or momentum is chosen directly from the horizontal array through a stochastic selection process. In order to keep the transition table within reasonable

20
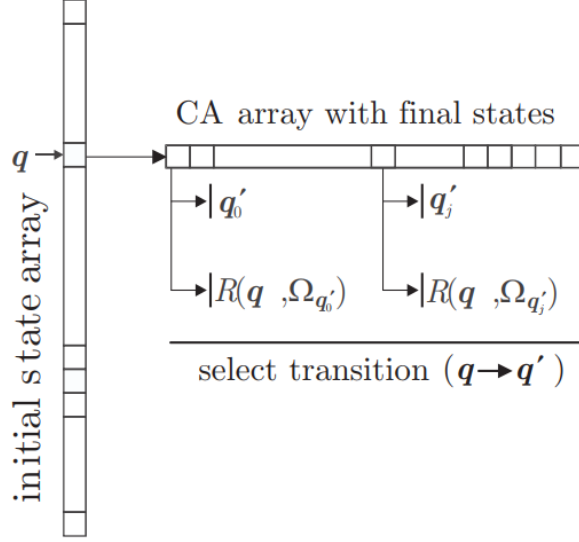
Figure 2.4: Structure of a transition table in CMC [3]

size, the whole BZ1 is divided into an inhomogeneous grid.

The scattering rate can also be represented in terms of probability of transition from an initial state $\mathbf{q}$ to a cell or region $\Omega_{\mathbf{q}'}$ around the point $\mathbf{q}'$ as:

$$R\left(\boldsymbol{q}, \Omega_{\boldsymbol{q}'}\right) = \int_{\Omega_{\boldsymbol{q}'}} d\boldsymbol{q}' R\left(\boldsymbol{q}, \boldsymbol{q}'\right);$$

if the cell is small enough, the expression becomes:

$$R\left(\boldsymbol{q}, \Omega_{\boldsymbol{q}'}\right) \simeq R\left(\boldsymbol{q}, \boldsymbol{q}'\right) D\left(\mathcal{E}\left(\boldsymbol{q}'\right), \Omega_{\boldsymbol{q}'}\right)$$

where $D\left(\mathcal{E}\left(\boldsymbol{q}'\right), \Omega_{\boldsymbol{q}'}\right)$ is the density of states at energy $\mathcal{E}(\boldsymbol{q}')$ in $\Omega_{\mathbf{q}'}$.

2.8.2 Rejection Algorithm

Rejection is one of the core algorithmic techniques implemented in CMC, which addresses among other, the need for multiple scattering tables at different temperature in an electrothermal simulation. During the initialization of an electrothermal simulation, the whole device is set at 300K and prior to that, the scattering table is also computed for different mechanisms for $T = 300K$. As the simulation progresses,

21
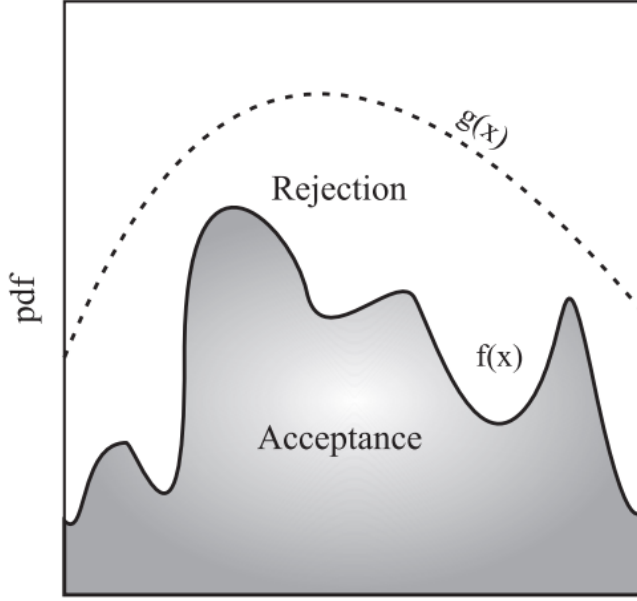
Figure 2.5: Schematic of the rejection algorithm, where g(x) is the proposed pdf and f(x) is the desired pdf, and $g(x) > f(x)$ at any value of x. [3]

electrons gain high kinetic energy because of the externally applied bias. Through electron-phonon scattering events, electrons start loosing energy to the crystal and in some part of the device the temperature reaches values well above the initial 300K. So if any part of the device reaches a higher temperature, we would need a series of scattering tables for collision events occuring in that region.

The primary reason behind switching to CMC, was the pre-tabulation of the scattering rates, which the speeds the simulations. Computing several tables for different temperatures prior to simulation and loading them is not realistic, as the range of the temperature can be wide depending on the simulation. For example, even if a table is computed at 50K intervals and performing some interpolation, a temperature range of 300K will require 6 scattering tables for a single material. The rejection algorithm as described in [59], can address the issue with significant accuracy and with a small price in terms of performance. Figure 2.5 illustrates the

principle of rejection algorithm, where the probability distribution function (pdf) is unknown initially. In this algorithm, two pdfs: desired distribution (f(x)) and proposed distribution (g(x)) are defined, such as, for any value of x, the g(x) is equal or greater than f(x). The rejection probability can be computed as $P_r = \frac{f(a)}{g(a)}$, which will range from 0 to 1. At the end, a random number between 0 to 1 is generated. If $R > P_r$, the event is rejected and accepted otherwise.

2.8.3 Application of Rejection in CMC

The rejection technique can be easily applied to the CMC scattering process, where g(x) will be representing the scattering rate at a pre-defined maximum temperature, while f(x) will be scattering rate at the local temperature in different parts of the device. As scattering rates are directly proportional to temperature, the scattering rate corresponding to maximum temperature will be higher than local scattering rates for any values of energy, which satisfies the $g(x) \geq f(x)$ condition of the rejection algorithm. Figure 2.6 shows an implementation of the rejection algorithm in CMC. Before the start of the simulation, the CMC code pre-tabulates the rate at 800K, which is the maximum temperature set for the simulation and represents g(x).

In a space of possible final state $\mathbf{q}'$, the proposal distribution $g(\mathbf{q}')$, target distribution $f(\mathbf{q}')$ and a constant c for initial state $\mathbf{q}^{ini}$ are defined. The purpose of defining c is to ensure that $cg(\mathbf{q}') > f(\mathbf{q}')$. The proposal pdf is calculated through the normalization to the maximum scattering rates $R_{max}(\mathbf{q}^{ini}, \mathbf{q}')$ such as:

$$g\left(\boldsymbol{q}'\right) = \frac{R_{\max}\left(\boldsymbol{q}^{ini}, \boldsymbol{q}'\right)}{\sum_{\boldsymbol{q}} R_{\max}\left(\boldsymbol{q}^{ini}, \boldsymbol{q}\right)}$$

Similarly, the target distribution can be obtained as:

$$f\left(\boldsymbol{q}'\right) = \frac{R_{loc}\left(\boldsymbol{q}^{ini}, \boldsymbol{q}'\right)}{\sum_{\boldsymbol{q}} R_{loc}\left(\boldsymbol{q}^{ini}, \boldsymbol{q}\right)}$$
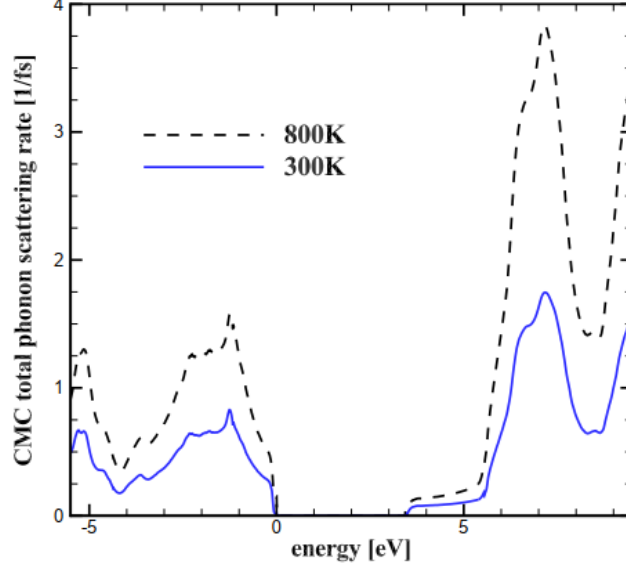
Figure 2.6: Phonon scattering rate for GaN obtained from CMC [2].

where $R_{loc}\left(\boldsymbol{q}^{ini}, \boldsymbol{q}\right)$ represents the rate of transition from $\boldsymbol{q}^{ini}$ to $\mathbf{q}'$ from local values of parameters. The constant c can be written as follows:

$$c = \frac{\sum_{\boldsymbol{q}} R_{\max}\left(\boldsymbol{q}^{ini}, \boldsymbol{q}\right)}{\sum_{\boldsymbol{q}} R_{loc}\left(\boldsymbol{q}^{ini}, \boldsymbol{q}\right)}.$$

From these computations, the rejection probability is expressed as follows:

$$P_{rej}\left(\boldsymbol{q}'\right) = \frac{f\left(\boldsymbol{q}'\right)}{c\, g\left(\boldsymbol{q}'\right)} = \frac{R_{loc}\left(\boldsymbol{q}'\right)}{R_{\max}\left(\boldsymbol{q}'\right)}.$$

In every scattering rates, there will be many common terms between $R_{loc}(\mathbf{q}')$ and $R_{max}(\mathbf{q}')$, which will cancel out. For example, for the case of deformation potential scattering rate in Equation 2.17, the rejection probability will be :

$$P_{rej}^{def} = \frac{\Gamma_{loc}^{def}\left(\boldsymbol{k}, \Omega_{\boldsymbol{k}'}\right)}{\Gamma_{\max}^{def}\left(\boldsymbol{k}, \Omega_{\boldsymbol{k}'}\right)} = \frac{\left(n_{loc} + \frac{1}{2} \mp \frac{1}{2}\right)}{\left(n_{\max} + \frac{1}{2} \mp \frac{1}{2}\right)}$$

where n_{loc} and n_{max} represent the number of phonons at local and maximum temperature with wavevector $\mathbf{q}$. Similarly, for piezoelectric scattering given in Equation 2.18, the

24
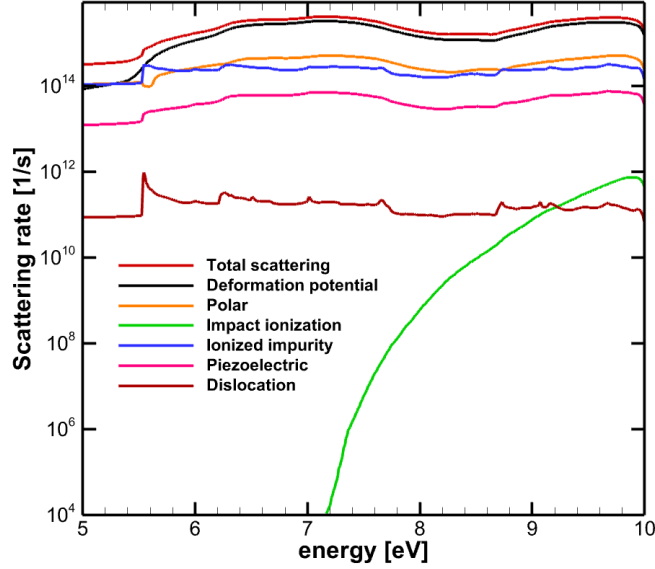
Figure 2.7: Different scattering rates obtained for wurtzite GaN by using the rejection algorithm.

probability is:

$$P_{rej}^{\text{piezo}} = \frac{\Gamma_{loc}^{\text{piezo}}\left(\boldsymbol{k}, \Omega_{\boldsymbol{k}'}\right)}{\Gamma_{\max}^{\text{piezo}}\left(\boldsymbol{k}, \Omega_{\boldsymbol{k}'}\right)} = \frac{T_{loc}}{T_{\max}}$$

where T_{loc} and T_{max} are the local and maximum temperature, respectively. Figure 2.7 shows the scattering rates for wurtzite GaN obtained by using the rejection algorithm implemented in CMC. Similar to scattering rates, the rejection can be applied to occupation number of phonons. The expected occupation number of phonon can be computed from the Bose-Einstein distribution such as:

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

from where the expectancy can be obtained for the highest temperature, P_{max} and the local temperature, P_{loc} can be obtained. Again, with the help of a random number,R, generated between 0 and 1, the scattering is accepted if $RP_{max} \leq P_{loc}$ and rejected otherwise.

25

2.9   Conclusion

In this chapter, the theoretical background of carrier transport in semiconductor devices has been discussed. The BTE supports the proper formalism to describe the carrier transport through a seven-dimensional distribution function that evolves in real and momentum space. Models based on lower order moments of the BTE can be faster, at least in 2D, but accuracy and robustness can be compromised for macroscopic device simulations.

On the contrary, Monte Carlo methods can provide a statistical and exact solution of the BTE considering all the moments. In these methods, scattering events are modeled as stochastic processes. Traditional EMC calculates the scattering rates during runtime, making it computationally very expensive. As an extension of EMC, the CMC mehtod pretabulates the scattering rates along with the final states in a large lookup table making it computationally faster than EMC. Two special components of the CMC: the transition table and rejection algorithm were discussed briefly. The theoretical aspects behind different scattering mechanisms considered in this study were also discussed.

Chapter 3

PORTABLE, EXTENSIBLE TOOLKIT FOR SCIENTIFIC COMPUTATION

(PETSC) AS A POISSON SOLVER

3.1   Introduction

Generally, any particle based simulator is a combination of self-consistent solution

of Boltzmann's transport equation (BTE) and Poisson's equation. In EMC or CMC,

BTE is solved following a stochastic process by tracking the particles in both position

and momentum space. After a specific time interval, the charge densities obtained

from BTE are used as an input to the Poisson solver. Later, the electrostatic potential

in each grid cell is computed by the Poisson solver and the carriers are accelerated

accordingly. There are primarily two discretization techniques available for solving the

Poisson equation numerically: finite-element and finite-difference [60]. Choice between

these two primarily depends on the procedure involved in solving the BTE. Drift

diffusion methods consider only lowest moments of BTE and finite-element methods

are used in solving Poisson's equation along with it [61]. In hydrodynamic models or

Monte Carlo methods, where higher order moments of BTE is used, finite difference

method is the ultimate choice. If an iterative solver is used as a Poisson solver, the

solver must use the previous solution as an initial guess. In order to track the carriers

in an accurate way, the electrostatic potential needs to be updated typically in the

order of every femtoseconds [62]. If the simulated geometry is complex enough, the

computer time taken by the Poisson solver in Monte Carlo method can be similar or

even higher than the time taken by solving of BTE [63]. Similar scenario arises for

the case of cellular automaton method. For these reasons, inclusion of a robust and

efficient Poisson solver is an integral part of any device simulation for both accuracy and speed of the simulation. In this chapter, PETSc, a robust computational toolkit from Argonne National Lab, has been incorporated into the CMC as a Poisson solver.

The chapter is organized as follows: section 3.2 discusses available Poisson solvers in CMC, while in section 3.3, the procedure to incorporate finite-difference to solve Poisson's equation is explained. On the next two sections the formation of Laplacian matrix and the forcing functions are discussed. Section 3.6 introduces PETSc along with its different components like the preconditioner and the Krylov subspace solver. Section 3.7 describes how PETSc is included in PETSc using the available routines in the CMC. On the last section, the results obtained using PETSc are compared with results from existing multi-grid solver in CMC, along with a performance comparison.

3.2   Available Poisson solvers in CMC

Currently in CMC, two types of mesh relaxation methods are available: multi-grid and successive over relaxation (SOR). The extension of 2D geometries to 3D is straightforward, that's why the discussion is limited here to solving Poisson's equation in two dimensions which can be expressed as:

$$\nabla^2 \phi = f(x, y), \tag{3.1}$$

where ∇^2 is the Laplacian operator, ϕ is the electrostatic potential and $f(x, y)$ is the source term or the forcing function. The above elliptic differential equation can be converted to an algebraic matrix equation using finite difference algorithm as:

$$A\mathbf{u} = \mathbf{f} \tag{3.2}$$

where the matrix A represents the Laplace operator, $\mathbf{u}$ is a vector representing the solution and $\mathbf{f}$ is the forcing function of the system. If $\mathbf{v}^i$ is the approximation of $\mathbf{u}$

after the ith iteration, then the residual $\mathbf{r}^i$ can be defined as:

$$\mathbf{r}^i = \mathbf{f} - A\mathbf{v}^i$$

and the residual equation is expressed as:

$$A\mathbf{e}^i = \mathbf{r}^i \tag{3.3}$$

where $\mathbf{e}^i$ is the algebraic error in the form of $\mathbf{e}^i = \mathbf{u} - \mathbf{v}^i$.

The successive overrelaxation (SOR) method assumes the iterative nature of the solution in such a way that $\mathbf{v}_k^{(i+1)} = \mathbf{v}_k^i + \omega \mathbf{r_k}^i$, where ω is a relaxation factor that ranges between 0 and 2 for SOR. The value of this factor completely depends on the geometry. For realistic devices, as the number of grid point grows along with inhomogeneous grids, the performance of SOR becomes unacceptable.

Multi-grid method uses coarser length scales compared to the original grid size to reduce the relative error of the solution, to be more specific solving Equation 3.3 on a grid: Ω_{n-1}, in lieu of original grid, Ω_n. The updated solution can be found from the correction of the previous solution such as:

$$\mathbf{v}^{i+1} = \mathbf{v}^i - \mathbf{e}^i$$

If the error $\mathbf{e}^i$ is expanded into Fourier series, it can be shown that its long wavelength components is only reduced slightly in fine grid as their spatial extent falls beyond the range of the relaxation operator. The coarser grid translates them into relatively shorter wavelength to be treatable by the relaxation operator. This reasoning provides the prime advantage to the multi-grid solver from the context of convergence, compared to single-grid techniques like SOR. Two-grid solver, which is the simplest multi-grid possible, is described in Figure 3.1.
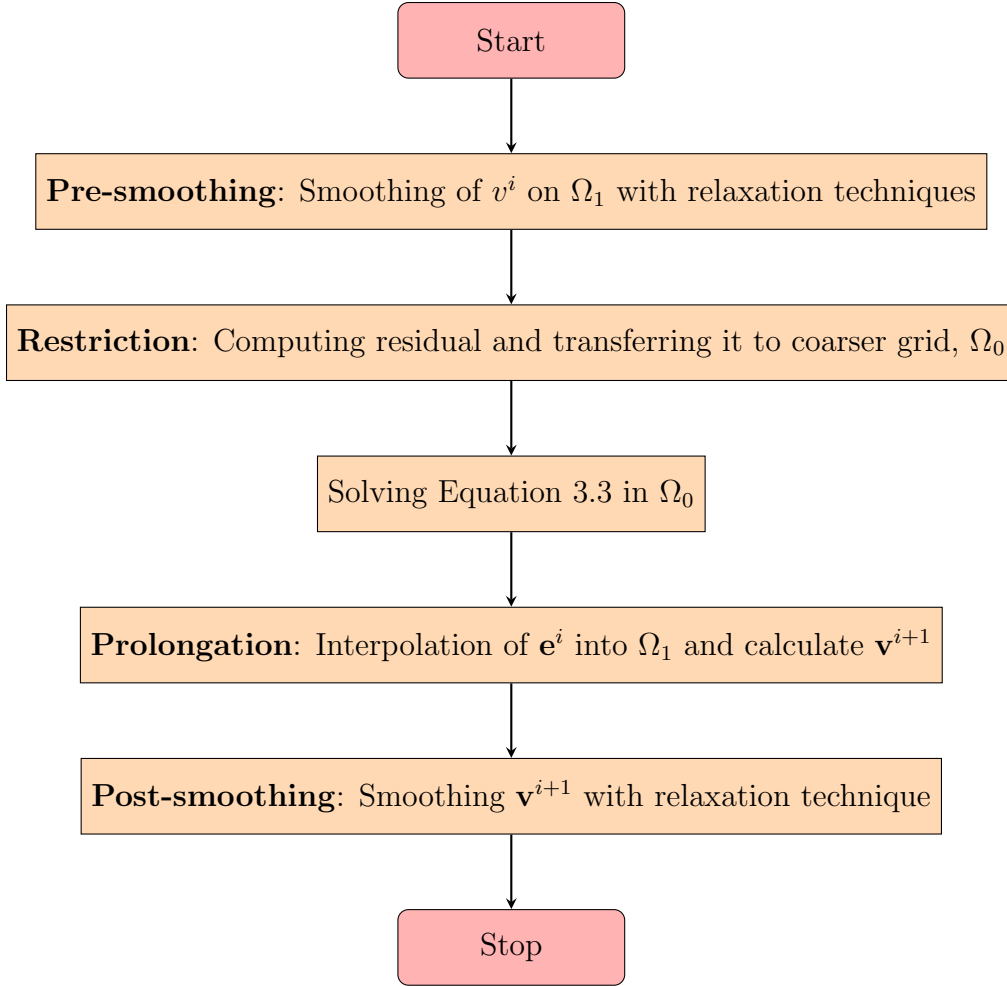
Figure 3.1: Process flow of a simple multi-grid solver.

The two-grid process can be extended to increasingly coarser grid case until the coarsest grid Ω_0 is reached. At that level, Equation 3.3 can be easily solved as it has only a few number of grid points and the whole process is repeated again till the desired convergence is achieved.

3.3 Finite Difference Scheme

Figure 3.2 shows the Finite difference gridding scheme adopted in CMC. For a specific cell c, the 4 neighbors (northern, eastern, western and southern) are represented
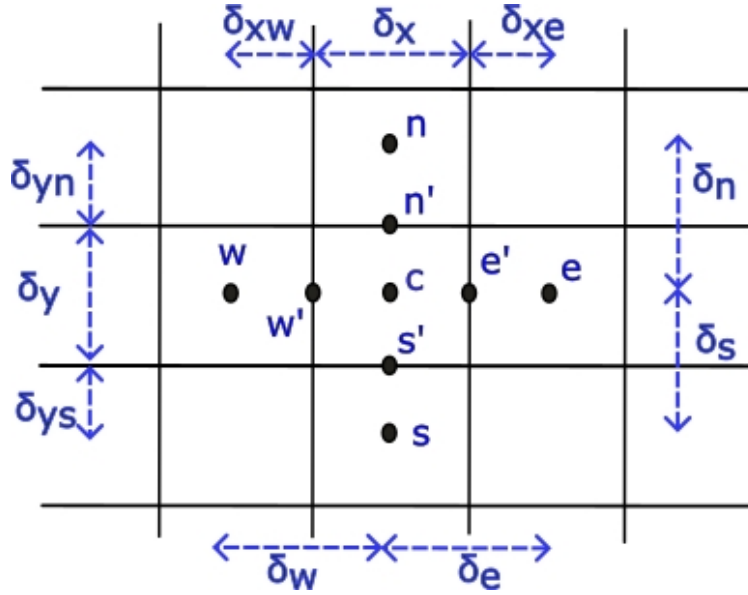
Figure 3.2: Finite difference grid used by Monte Carlo.

as n, e, w, s, respectively. The other points marked with a $'$ sign (n$'$, e$'$, w$'$, s$'$) are used for interpolation purpose. The dimensions of cell c are shown as δ_x and δ_y. The quantities, δ_n, δ_s, δ_w and δ_e represent the distances from the center of the cell c to its northern, southern, western and eastern neighbors, respectively. For performing interpolations from cell-to-cell, the distances labeled as δ_{xw}, δ_{xe}, δ_{yn} and δ_{ys} are being used, which represent the distances between the boundary of the cell c to the center of the neighboring cells. During the initialization, the nearest neighbors of each cell are computed and stored in linked lists. The boundary conditions of the geometry should be modeled in such a way that computations can be done for all grid points in the same manner. Figure 3.3 presents the process of setting coefficients for cells with fixed potentials in the boundary of the geometry (gray). For the left cell, the potential is fixed following the Dirichlet boundary condition and so the pointer to the northern neighbor is pointed to itself. The cell on the right has a null perpendicular field as it has a free surface on top, as a result it follows Neumann's boundary condition.

In this particular case, the southern cell is mapped as the northern cell. In this way a homogeneous boundary condition is implemented throughout the whole device geometry. The distances introduced in Figure 3.2 for the boundary cells can be
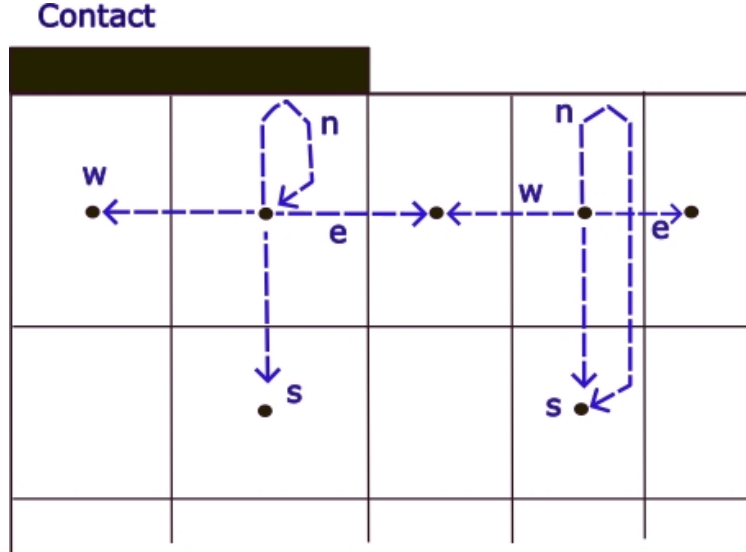
Figure 3.3: Implementation of Dirichlet and Neumann's boundary conditions

computed as:

$$
\delta_n = \begin{cases} |y_n - y_c| & y_n \neq y_c \\ 1 & \text{otherwise} \end{cases} \quad , \quad \delta_e = \begin{cases} |x_e - x_c| & x_e \neq x_c \\ 1 & \text{otherwise} \end{cases} \quad ,
$$

$$
\delta_w = \begin{cases} |x_c - x_w| & x_w \neq x_c \\ 1 & \text{otherwise} \end{cases} \quad , \quad \delta_s = \begin{cases} |y_c - y_s| & y_s \neq y_c \\ 1 & \text{otherwise} \end{cases} \quad .
$$

3.4 Laplacian Set-up

The Laplace operator in the 2D Poisson relation stated in eq. (3.1) can be approximated by the difference relation:

$$
\nabla^2 \phi_c \sim a_n \phi_n + a_e \phi_e + a_w \phi_w + a_s \phi_s + a_c \phi_c + O\left(\delta_x \delta_y\right)^m \tag{3.4}
$$

32

where the last term on the right is an error term which depends on the size of the cells. In an uniform grid the value of m is 3, while for non-uniform grids $m < 3$ [64]. The Laplacian matrix coefficients can be computed from the second order Taylor expansion of ϕ such as:

$$a_n = \begin{cases} 0 \\ \left(\frac{2}{\delta_n(\delta_n+\delta_s)}\right) \\ \left(\frac{2}{\delta_n(\delta_n+\delta_s)}\right)\left(1+\frac{\delta_{yn}}{\delta_y/2}\right)\left(\frac{\epsilon_n/\delta_{yn}}{(\epsilon_n/\delta_{yn})+[\epsilon_c/(\delta_y/2)]}\right) \end{cases}$$

$$a_e = \begin{cases} 0 \\ \left(\frac{2}{\delta_e(\delta_e+\delta_w)}\right) \\ \left(\frac{2}{\delta_e(\delta_e+\delta_w)}\right)\left(1+\frac{\delta_{xe}}{\delta_x/2}\right)\left(\frac{\epsilon_e/\delta_{xe}}{(\epsilon_e/\delta_{xe})+[\epsilon_c/(\delta_x/2)]}\right) \end{cases}$$

$$a_w = \begin{cases} 0 \\ \left(\frac{2}{\delta_w(\delta_e+\delta_w)}\right) \\ \left(\frac{2}{\delta_w(\delta_e+\delta_w)}\right)\left(1+\frac{\delta_{xw}}{\delta_x/2}\right)\left(\frac{\epsilon_w/\delta_{xw}}{(\epsilon_w/\delta_{xw})+[\epsilon_c/(\delta_x/2)]}\right) \end{cases}$$

$$a_s = \begin{cases} 0 \\ \left(\frac{2}{\delta_s(\delta_n+\delta_s)}\right) \\ \left(\frac{2}{\delta_s(\delta_n+\delta_s)}\right)\left(1+\frac{\delta_{ys}}{\delta_y/2}\right)\left(\frac{\epsilon_s/\delta_{ys}}{(\epsilon_s/\delta_{ys})+[\epsilon_c/(\delta_y/2)]}\right) \end{cases}$$

$$a_c = \begin{cases} 1 \\ -\left(\frac{2}{\delta_n+\delta_s}+\frac{2}{\delta_w+\delta_e}\right) \\ -\left(\frac{2}{\delta_n+\delta_s}+\frac{2}{\delta_w+\delta_e}\right)+p_n+p_e+p_w+p_s \end{cases}$$

where the first option in each case represents the Dirichlet boundary condition for fixed potential. The second option is for homo-junction devices and the third option is for devices with different materials, where the dielectric constant is varied in a piece wise constant manner with space [65]. The values of p_n, p_e, p_w and p_s for the central

33

cell can be found as:

$$p_n = \frac{2}{\delta_n \left(\delta_n + \delta_s\right)} \left[\left(1 + \frac{\delta_{yn}}{\delta_y/2}\right) \left(\frac{\epsilon_c/\left(\delta_y/2\right)}{\left(\epsilon_n/\delta_{yn}\right) + \left[\epsilon_c/\left(\delta_y/2\right)\right]}\right) - \left(\frac{\delta_{yn}}{\delta_y/2}\right)\right]$$

$$p_e = \frac{2}{\delta_e \left(\delta_e + \delta_w\right)} \left[\left(1 + \frac{\delta_{xe}}{\delta_x/2}\right) \left(\frac{\epsilon_c/\left(\delta_x/2\right)}{\left(\epsilon_e/\delta_{xe}\right) + \left[\epsilon_c/\left(\delta_x/2\right)\right]}\right) - \left(\frac{\delta_{xe}}{\delta_x/2}\right)\right]$$

$$p_w = \frac{2}{\delta_w \left(\delta_e + \delta_w\right)} \left[\left(1 + \frac{\delta_{xw}}{\delta_x/2}\right) \left(\frac{\epsilon_c/\left(\delta_x/2\right)}{\left(\epsilon_w/\delta_{xw}\right) + \left[\epsilon_c/\left(\delta_x/2\right)\right]}\right) - \left(\frac{\delta_{xw}}{\delta_x/2}\right)\right]$$

$$p_s = \frac{2}{\delta_s \left(\delta_n + \delta_s\right)} \left[\left(1 + \frac{\delta_{ys}}{\delta_y/2}\right) \left(\frac{\epsilon_c/\left(\delta_y/2\right)}{\left(\epsilon_s/\delta_{ys}\right) + \left[\epsilon_c/\left(\delta_y/2\right)\right]}\right) - \left(\frac{\delta_{ys}}{\delta_y/2}\right)\right]$$

It is noteworthy that these Laplacian coefficients are only computed once at the beginning of the simulation as they are only function of the geometry of the device and the dielectric constant. Both the multigrid solver and PETSc allows to store the matrix in memory and access it during runtime for reducing computational expenses.

3.5   Forcing Function set-up

The forcing function of the Poisson equation should account for both boundary condition and the surface charge (s_c) defined on boundary of each cell (points n$'$, e$'$, w$'$, s$'$ in Figure 3.2). The updated forcing function can be defined as :

$$f_c' = k_c' + k_c'' \rho_c \tag{3.5}$$

where ρ_c is the density of carriers and the constants k' and k'' can be computed from

$$k_c'' = \begin{cases} 0 & \text{(a)} \\ -\frac{q}{\epsilon_0 \epsilon_c} & \text{(b)} \end{cases}$$

$$k_c' = \begin{cases} \phi_c & \text{(a)} \\ g_n q_n + g_e q_e + g_w q_w + g_s q_s & \text{(b)} \end{cases}$$

34

where (a) represents the values to be used on the grid points with Dirichlet boundary condition. The other coefficients g_n, g_e, g_w and g_s are function of the grid geometry as:

$$g_n = \begin{cases} 0 & \text{(a)} \\ -\frac{2}{\delta_n(\delta_n + \delta_s)} & \text{(b)} \end{cases}, \quad g_e = \begin{cases} 0 & \text{(a)} \\ -\frac{2}{\delta_e(\delta_e + \delta_w)} & \text{(b)} \end{cases}$$

$$g_w = \begin{cases} 0 & \text{(a)} \\ -\frac{2}{\delta_w(\delta_e + \delta_w)} & \text{(b)} \end{cases}, \quad g_s = \begin{cases} 0 & \text{(a)} \\ -\frac{2}{\delta_s(\delta_n + \delta_s)} & \text{(b)} \end{cases}$$

$$g_c = \begin{cases} 0 & \text{(a)} \\ -\left(\frac{2}{\delta_e \delta_w} + \frac{2}{\delta_s \delta_n}\right) & \text{(b)} \end{cases} = \begin{cases} 0 & \text{(a)} \\ -(q_n + q_e + q_w + q_s) & \text{(b)} \end{cases}$$

where (a) and (b) conveys similar meaning as the earlier equation of k'_c and k''_c. The quantities, q_n, q_e, q_w and q_s are the interpolated effective charge at the cell boundary which can be computed as:

$$q_n = -\frac{q}{\epsilon_0} s_n \left(\frac{1 + [\delta_{yn}/(\delta_y/2)]}{(\epsilon_n/\delta_{yn})/[\epsilon_c/(\delta_y/2)]} \right),$$

$$q_e = -\frac{q}{\epsilon_0} s_e \left(\frac{1 + [\delta_{xe}/(\delta_x/2)]}{(\epsilon_e/\delta_{xe})/[\epsilon_c/(\delta_x/2)]} \right),$$

$$q_w = -\frac{q}{\epsilon_0} s_w \left(\frac{1 + [\delta_{xw}/(\delta_x/2)]}{(\epsilon_w/\delta_{xw})/[\epsilon_c/(\delta_x/2)]} \right),$$

$$q_s = -\frac{q}{\epsilon_0} s_s \left(\frac{1 + [\delta_{ys}/(\delta_y/2)]}{(\epsilon_s/\delta_{ys})/[\epsilon_c/(\delta_y/2)]} \right).$$

Similar to the formation of the Laplacian operator, the forcing function is also computed only once at the very beginning as all of its elements are dependent on the initial state of the system.

3.6   The PETSc software tool

The Portable, Extensible Toolkit for Scientific Computation (PETSc) [66] is a popular and large suite of scalable (parallel) data structures and routines like nonliear equation solvers, Ordinary Differential Equation (ODE) integrators and optimization

algorithms for C, C++, Fortan and Python based codes [66]. Developed by the Argonne National Lab, PETSc provides an absolute interface for the parallel management of PDE discretization, while maintaining parallel routines for matrix and vector assembly.
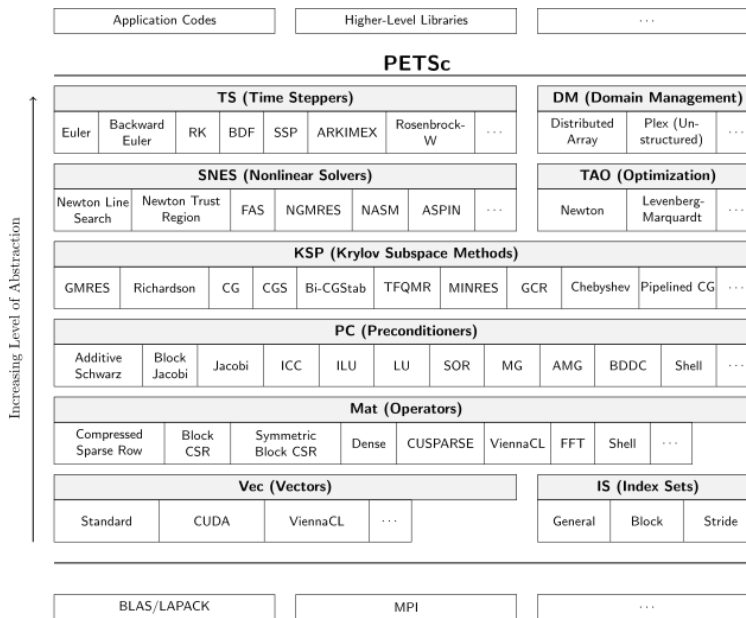
Figure 3.4: Different components of PETSc

For obtaining the solution of any linear equation like Equation 3.2, where $A \in \mathbb{R}^{n \times n}$ and $x, b \in \mathbb{R}^n$, PETSc uses the object KSP. This object can be used for solving all types of linear system solvers either it is parallel/sequential and direct/iterative.

For a three dimensional Partial Differential Equation (PDE), where the number of unknowns is N, a direct method would require:

$$work \geq CN^{\alpha \geq 2} \tag{3.6}$$

$$memory \geq CN^{\beta \geq 4/3} \tag{3.7}$$

which implies that the amount of work needed in direct method is at least the squared of number of unknowns. As the number of unknowns in any modern numerical codes

are really higher, most of the algorithms involve using iterative solvers. Similarly, PETSc uses the Krylov subspace method along with a preconditioner to obtain an iterative solution of the linear system.

3.6.1   Preconditioner in PETSc

In linear algebra, preconditioning usually refers to "the art of transforming a problem that appears intractable into another whose solution can be approximated rapidly" [67]. In most of the cases, the preconditioner is a matrix that converts Equation 3.2 in such a way that an iterative solver can converge faster. Generally, preconditioners are used for large and sparse matrix, resulting in most of the cases from the numerical discretization of PDEs [68]. For obtaining an excellent preconditioner three conditions should be fullfilled:

1)The preconditioned matrix should be close to the identity matrix making it somewhat easier to solve and ensuring an accelerated convergence.

2) The preconditioner should be computationaly inexpensive.

3) The application of the preconditioner to a vector should be inexpensive as well.

Equation 3.2 is multiplied by M^{-1}, where M is a nonsingular matrix,

$$M^{-1}Ax = M^{-1}b \tag{3.8}$$

if the resultant matrix $M^{-1}A$ is better conditioned than A for a Krylov subspace method (KSP) [69], then M^{-1} is the preconditioner and this system can be called as left preconditioned system. Similarly M^{-1} can be applied to the right like:

$$AM^{-1}y = b, \qquad x = M^{-1}y \tag{3.9}$$

which is called right preconditioned system. These two different approach can provide different advantages: the right preconditioning keeps the residual of the updated

Method	PCType	Options Database Name
Jacobi	PCJACOBI	jacobi
SOR	PCSOR	sor
Cholesky	PCCHOLESKY	cholesky
Incomplete LU	PCILU	ilu
Algebraic multi-grid	PCGAMG	gamg
Linear Solver	PCKSP	ksp
No preconditioning	PCNONE	none

Table 3.1: Partial list of Preconditioners available in PETSc.

preconditioned system same as the original one, while in case of left preconditioning the residual changes. For this reason, right preconditioning can be useful in residual minimization algorithms like GMRES [68]. For the right preconditioned system, computation of an additional variable, y, is a major disadvantage. A split preconditioning can be performed to use advantages of left and right preconditioning such as:

$$M^{-1} = M_2^{-1}M_1^{-1} \tag{3.10}$$

$$M_1^{-1}AM_2^{-1}y = M_1^{-1}b, \qquad x = M_2^{-1}y \tag{3.11}$$

PETSc has a really large collection of preconditioners that can be used for numerical methods. A partial list of preconditioners available in PETSc are shown in Table 3.1.

In this work, incomplete lower upper factorization (ILU) has been adopted as a preconditioning approach. The ILU method in representing a nonsingular matrix A in such way that:

$$A = LU \tag{3.12}$$

where L and U are two approximate triangular matrices and they are computed in a cost-effective way. A Gaussian elimination procedure is adopted to compute L and U, which are much less sparse than A. Further approximations of L and U, represented as $\tilde{L}$ and $\tilde{U}$ are defined by discarding some of the fill-in entries. The resulting preconditioner becomes:

$$M^{-1} = \tilde{L}^{-1}\tilde{U}^{-1} \tag{3.13}$$

At the very beginning of the ILU algorithm, a set $\mathcal{S}$ of position (i,j) is defined, where $1 \leq i$ and $j \leq n$, where n is the number of unknowns in the system. At the same time, either $\tilde{L}$ (if $i > j$) or $\tilde{U}$ (if $j > i$) has a nonzero entry. This set is called nonzero pattern for this algorithm, which is described as follows:

Data: Matrix A, the nonzero pattern $\mathcal{S}$

Result: Matrix A containing $\tilde{L}$ and $\tilde{U}$

for *each $(i,j) \notin \mathcal{S}$* **do**

$\quad \mid \quad a_{ij} = 0$

end

for *$i = 2, \ldots, n$* **do**

$\quad$ **for** *$k = 1, \ldots, i-1$ and $(i,k) \in \mathcal{S}$* **do**

$\quad\quad a_{ik} \leftarrow a_{ik}/a_{kk}$

$\quad\quad$ **for** *$j = k+1, \ldots, n$ and $(i,j) \in \mathcal{S}$* **do**

$\quad\quad\quad \mid \quad a_{ij} \leftarrow a_{ij} - a_{ik}a_{kj}$

$\quad\quad$ **end**

$\quad$ **end**

end

Algorithm 1: Algorithm used in ILU factorization

The algorithm above follows the Gauss elimination method. Further improvement techniques have been adopted to reduce its computational cost like KIJ/IJK/IKJ

elimination [70, 71].

3.6.2   Krylov Subspace method for iterative solver

The purpose of an iterative method [69] is to get closer to the original solution in each iteration. The fundamental equation representing classical iterative methods can be described as:

$$x^{k+1} = x^k + \alpha_k r^k \tag{3.14}$$

where α_k is a relaxation parameter and r^k is the residual after k^{th} iteration. The residual can be represented as:

$$r^k = b - Ax^k \tag{3.15}$$

Similar to x^{k+1}, the expression for x^k can be obtained after performing some manipulation:

$$x^k = x^{k-1} + \alpha_{k-1} r^{k-1}$$

$$b - Ax^k = b - Ax^{k-1} - A\alpha_{k-1} r^{k-1}$$

if the value of $r^k = b - Ax^k$ is replaced into the equation of x^{k+1} :

$$x^{k+1} = x^k + \alpha_k(b - Ax^k)$$

$$= x^k + \alpha_k(r^{k-1} - A\alpha_{k-1} r^{k-1})$$

$$= x^k + \alpha_k r^{k-1} - \alpha_k \alpha_{k-1} Ar^{k-1}$$

If the recursion is continued for k times it is evident that, the solution x^k will vary in such a way that $x^k \in$ span $\{r^0, Ar^0, A^2 r^0, \ldots, A^{k-1} r^0\}$. This subspace spanned by vectors is called Krylov subspace. The solver tries to find an optimal solution where the A-norm, $\|x - x^k\|_A$ is minimal.

40

3.7   Inclusion in Cellular Monte Carlo

The process flow describing the inclusion of PETSc into the cellular Monte Carlo (CMC) is shown on Figure 3.5. Initially, the values generated for the Laplace operator and the forcing function as described in section 3.4 and section 3.5, are passed to the PETSc solver function. Later, the initialization of PETSc, along with the formation of the Laplacian matrix and preallocation take place. The matrix components depend only on the device geometry, hence it is formed only once. On the next step, the vector representing the forcing function is formed along with its values obtained in section 3.5. The KSP solver is introduced and set up along with all of its attributes. After obtaining the solution, the values are inserted into the solution vector, x and returned to the function where the original PETSc function was called. During the last iteration of the Poisson solver, the solver and the vectors and the matrix are destroyed to deallocate memory.

3.7.1   PETSc Routines used

The inclusion of PETSc in CMC is performed using several routines included in PETSc. In this section all the necessary routines used for performing the addition of PETSc as a Poisson solver in the CMC are discussed. At the very beginning of the code, PETSC header file *petscksp.h* is added. This particular header file provides the interface functions necessary for the Krylov subspace accelerators.

PetscInitialize: PetscInitialize is used to initialize the PETSc database along with message passing interface (MPI). If MPI_Init() wasn't called prior to this, it calls MPI_Init() and so this routine is usually called at the very first line in most of the PETSc programs. In this study, it is used in the following way:
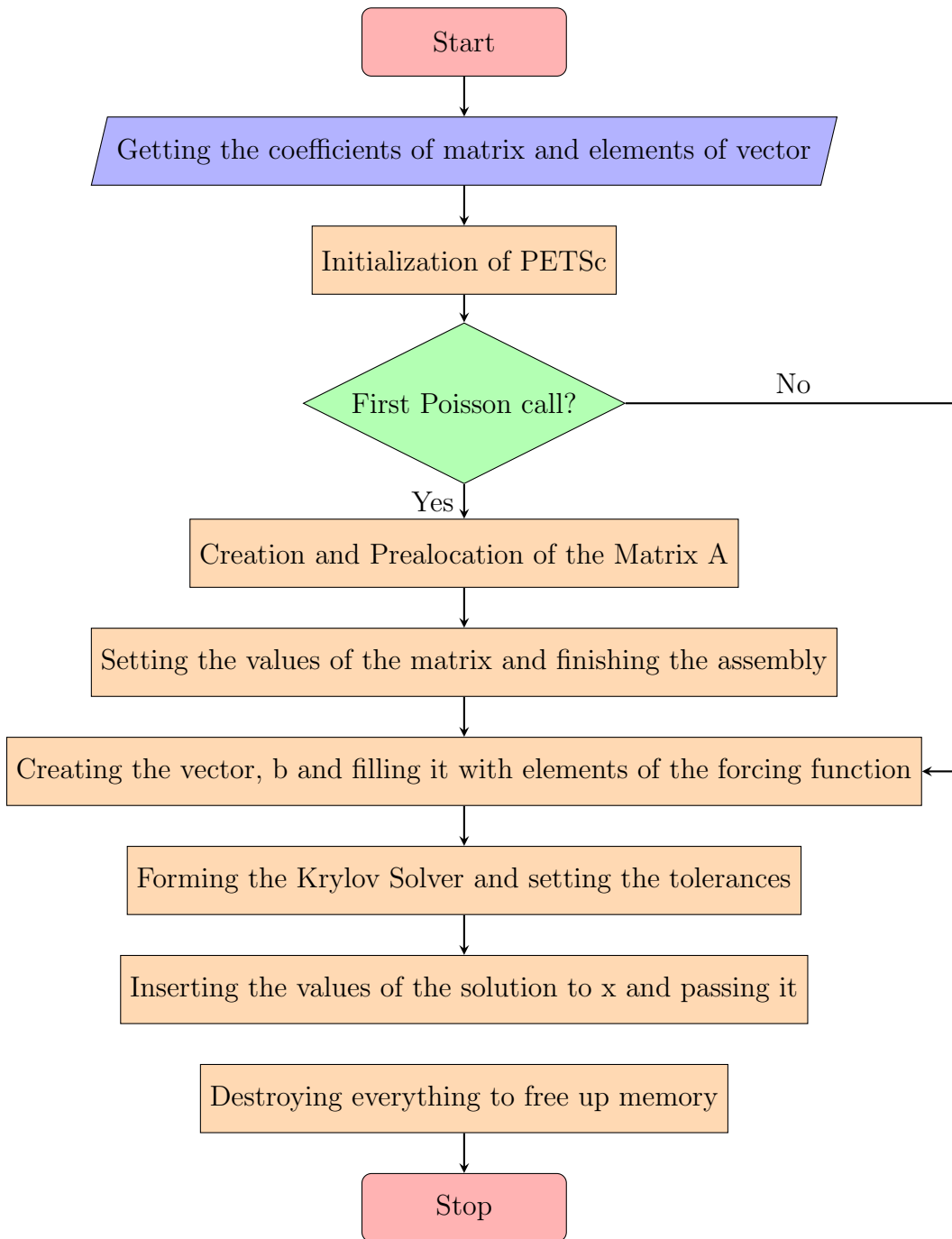
PetscInitialize(&argc, &argv, (char*)0, help);

Figure 3.5: Flow Chart representing PETSc as a Poisson solver in CMC.

Creating and filling the matrix: The matrix is created and preallocated by calling the following routines:

MatCreate(PETSC_COMM_WORLD, &A);

MatSetSizes(A, PETSC_DECIDE, PETSC_DECIDE, m*n*p, m*n*p);

MatSetFromOptions(A);

MatSeqAIJSetPreallocation(A, 7, NULL);

MatSetUp(A);

MatGetOwnershipRange(A, &Istart, &Iend);

MatCreate creates the matrix through the MPI_COMM_WORLD communicator, which contains all the processes available in PETSc to be used. MatSetSizes is used to set the local and global sizes of the matrix and performs a check in order to determine compability with the vectors. MatSetFromOptions is used to select the type of matrix from the available options database of PETSc, for example a parallel MPI matrix can be created through the options if there are more than one processor available for the simulation. If no options are selected, the default matrix type is MATAIJ where MatCreateSeqAIJ() routine is used. MatSeqAIJSetPreallocation performs the preallocation of the storage used by the matrix by setting the number of nonzero entities per row, as there can be seven neighbors at maximum for the 3D case, it is set to 7 in this study. It has been seen that the performance can be enhanced up to 50 times by setting the preallocation while assembling the matrix [66]. MatSetUp is used to setup internal data structures that can be used for later purposes. If there is a suitable choice of preallocation, this specific function is not needed. If the matrix owns values by row, MatGetOwnershipRange provides the range of the matrix rows. Here, Istart and Iend represent the global index of first row and global index of the last row, respectively.

MatSetValues(A, 1, &Ii, 1, &J, &(gdp -> socf), ADD_VALUES);

43

MatSetValues is used to add a block of values into the matrix, here gdp -> socf represents the south coefficients in Figure 3.2. By default in PETSc, the values are stored in a row-oriented manner. As the values can be ached, this command must be followed by MatAssemblyBegin().

MatAssemblyBegin(A,MAT_FINAL_ASSEMBLY);

MatAssemblyEnd(A,MAT_FINAL_ASSEMBLY);

There are two types of options of these two assembly: MAT_FLUSH_ASSEMBLY and MAT_FINAL_ASSEMBLY. MAT_FLUSH_ASSEMBLY is used when there is a necessity to switch between ADD_VALUES and INSERT_VALUES in the earlier MatSetValues() command, whereas MAT_FINAL_ASSEMBLY is used to for the last final assembly before actually using the matrix.

Vector Creation and assembly: The creation and insertion of the values for the vectors are done through following routines:

VecCreate(PETSC_COMM_WORLD, &b);

VecSetSizes(b, PETSC_DECIDE, m*n*p);

VecSetFromOptions(b);

VecDuplicate(b, &x);

VecCreate will create an empty vector object through the communicator for the vector object, where the local and global sizes of the vector are set by VecSetSizes. Here, m,n and p represent the grid sizes for x,y and z coordinates, respectively. Different options available in PETSc for the vector can be used by VecSetFromOptions. As the dimensions and other properties of the solution vector should be the same as the forcing function **b**, a new vector **x** is created as the same type of **b** with VecDuplicate command.

VecSetValue(b, Ii, gdp->fccf, INSERT_VALUES);

Similar to MatSetValues, VecSetValue inserts values into specific locations of the

vector. Here the forcing function components computed in section 3.5, are inserted into the vector **b**.

VecAssemblyBegin(b);

VecAssemblyEnd(b);

VecAssemblyBegin starts the assembly of the vector that ensures the storage of elements in correct MPI process and VecAssemblyEnd finalizes the assembly process.

VecAssemblyBegin(b);

VecAssemblyEnd(b);

On the next stage, the KSP solver is introduced along with the different options and properties of the solver, the routines used are shown below:

KSPCreate(PETSC_COMM_WORLD, &ksp);

KSPSetOperators(ksp, A, A);

KSPSetTolerances(ksp, PETSC_DEFAULT, 1.e-50, PETSC_DEFAULT, PETSC_DEFAULT);

KSPSetFromOptions(ksp);

KSPSolve(ksp, b, x);

KSPCreate creates the context of KSP, which is the primary object of PETSc that is used to manage all Krylov methods and the linear solvers. Even in the case of the use of a direct solver, the KSP object is actually used with KSPPREONLY, which means only the preconditioner is used as the linear solver. KSPSetOperators does the setup of the matrix associated with the linear system and the second one is to build the preconditioner associated with the solver. In this study, the preconditioner is built from the matrix **A** itself. The relative, absolute, divergence tolerances and maximum iteration associated with the KSP convergence tester are set by the KSPSetTolerances. The relative tolerance ($rtol$) refers to the relative decrease in the residual norm, while the absolute tolerance ($atol$) is for the absolute size of the residual norm. The di-

45

vergence tolerance (*dtol*) implies to the amount of increase in residual norm before the conclusion about the divergence of the method. In the solver the convergence is occurred when

$$\|r_k\|_2 < \max\,(\text{ rtol }*\|b\|_2,\,\text{atol })$$

where $r_k = b - Ax_k$. The solver concludes divergence if

$$\|r_k\|_2 > \text{ dtol }*\|b\|_2$$

The PETSC_DEFAULT values of these 4 parameters are: $rtol = 10^{-5}$, $atol = 10^{-50}$, $rtol = 10^{5}$ and $maxits = 10^{4}$.

KSPSetFromOptions is used here to set the options available in PETSc database. KSPSolve actually computes the solution of the linear solver and returns the solution vector x.

   KSPDestroy(&ksp);

   VecDestroy(&x);

   VecDestroy(&b);

At the end of the each iteration of the Poisson solver, the solver and the vectors are destroyed by KSPDestroy and VecDestroy routines to free up the memory. A similar operation is performed on the Laplacian matrix on the last iteration of the Poisson call, as the matrix is formed only once for the entire simulation.

3.8   Validation through 2D and 3D Simulations

In order to validate the inclusion of PETSc as a Poisson solver in the CMC code, a 2D GaN HEMT and a 3D MESFET using PETSc as a Poisson solver and the results are compared with the established multi-grid solver available in the CMC.
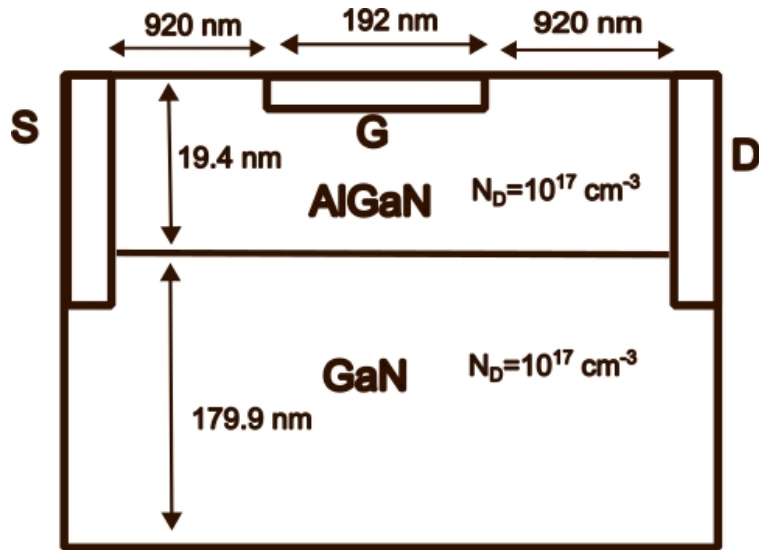
Figure 3.6: 2D GaN HEMT used for the validation.

2D GaN HEMT Results:

Figure 3.6 shows the device schematic of the simple 2D HEMT considered for the validation, while Figure 3.7 and Figure 3.8 show the potential plots obtained from the multi-grid and PETSc solver, respectively. It is evident from the plots that the potential distribution computed by both of the solvers are almost identical. To investigate the similarity between the results obtained by both of the solvers, profile along the channel for average x-electric field, average potential and e-total scattering has been compared in Figure 3.10, Figure 3.9 and Figure 3.11, respectively. All these results validates the accuracy of PETSc as a Poisson solver compared with multi-grid. Finally, the drain current for different bias points are compared these two solvers and it is evident from the plot that the currents are almost identical.
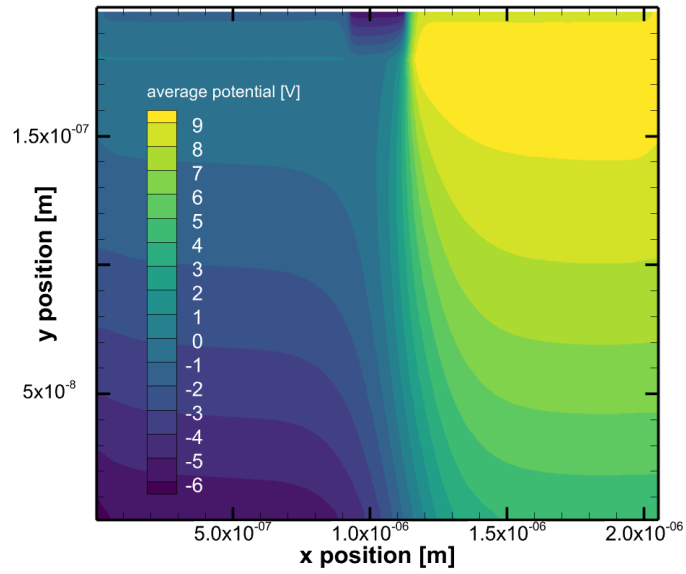
Figure 3.7: 2D potential plot obtained using the multi-grid solver at $V_{DS} = 10V$.
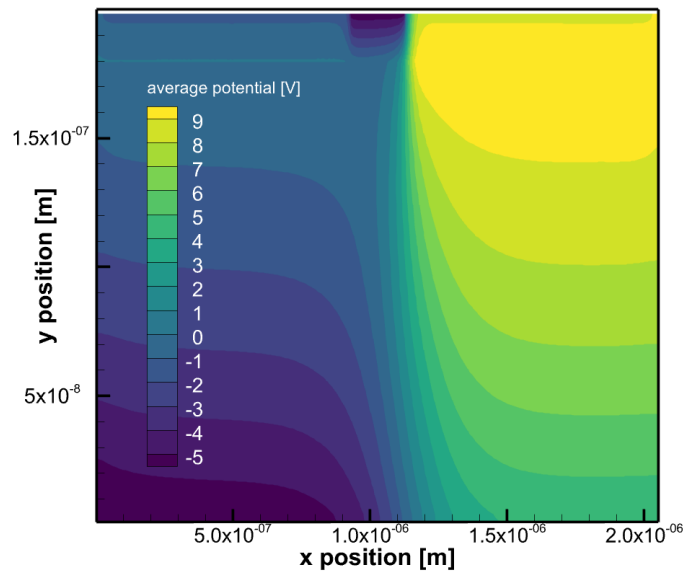
Figure 3.8: 2D potential plot obtained using the PETSc solver at $V_{DS} = 10V$.
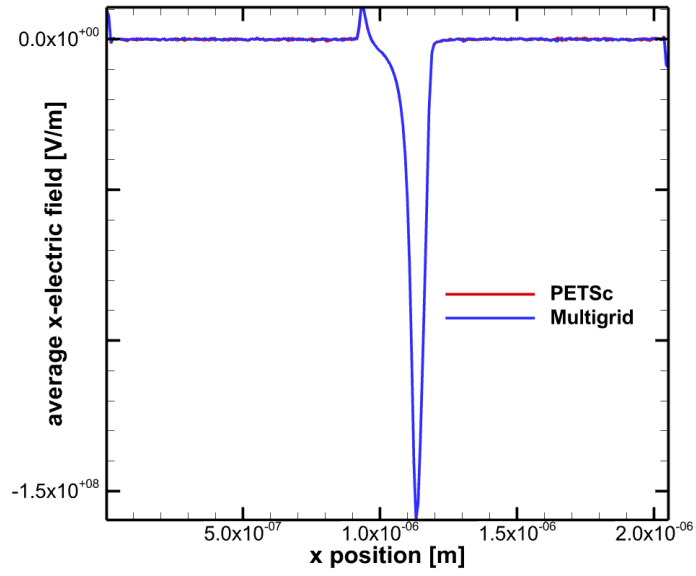
Figure 3.9: x-electric field along the channel obtained by PETSc and multi-grid solver.
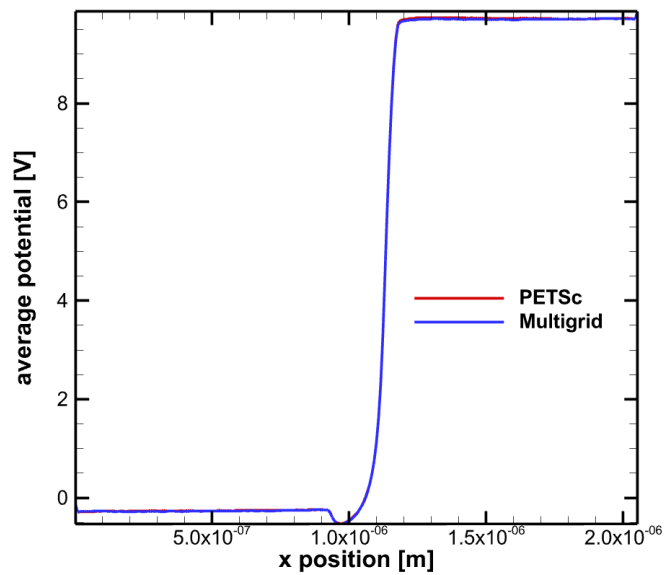
Figure 3.10: Potential along the channel obtained by PETSc and multi-grid solver.
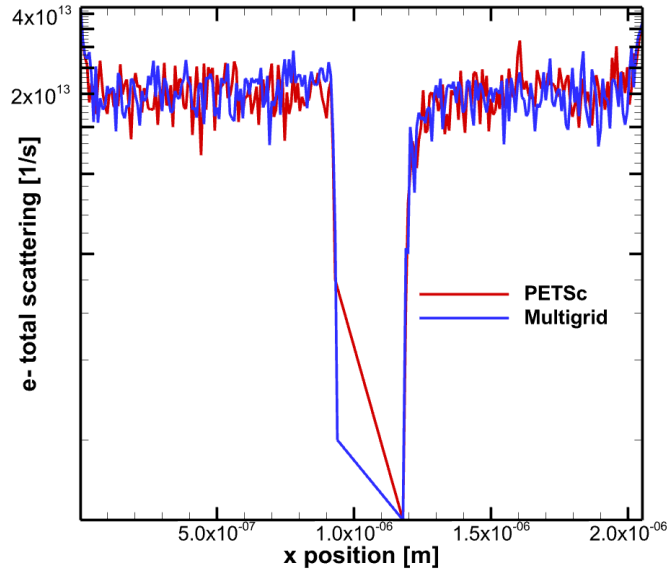
Figure 3.11: e-total scattering along the channel obtained by PETSc and multi-grid solver.
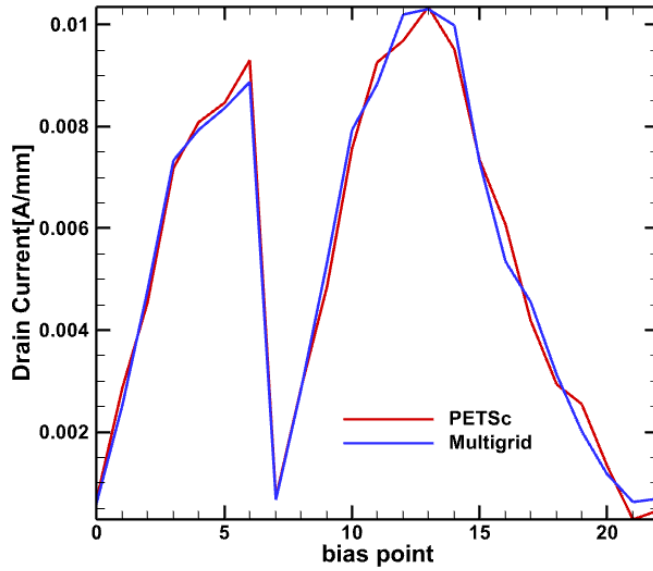
Figure 3.12: Drain current for the 2D HEMT for different bias points.
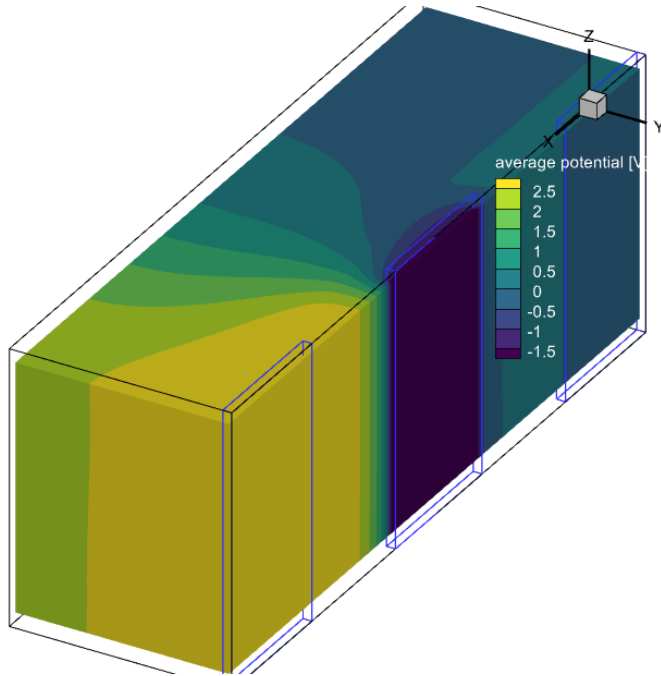
Figure 3.13: Potential plot obtained for the 3D MESFET using the multi-grid.
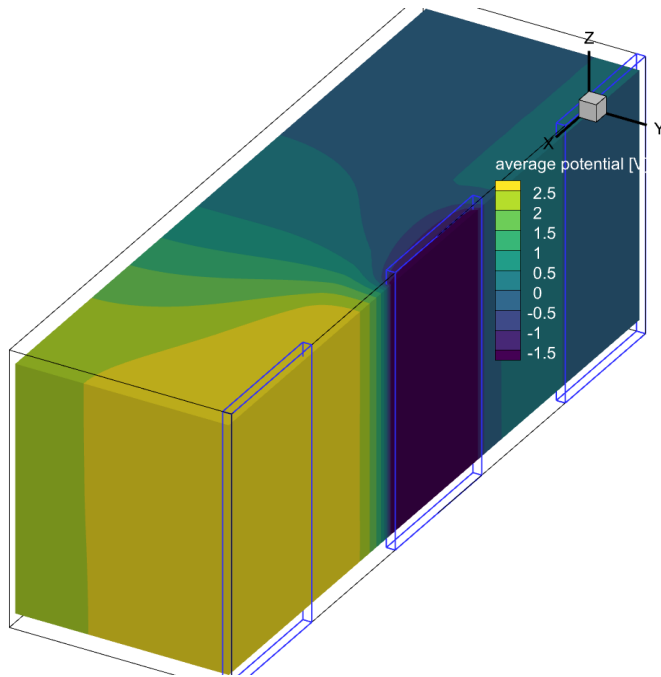
Figure 3.14: Potential plot obtained for the 3D MESFET using the PETSc.
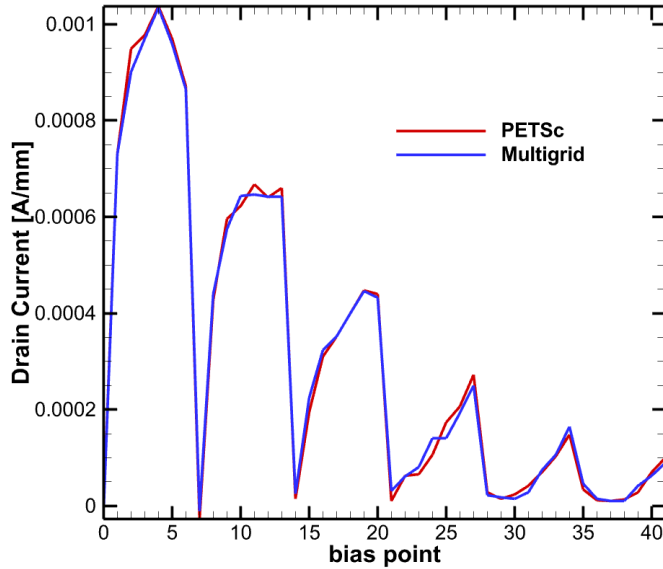
Figure 3.15: Drain current for the 3D MESFET obtained for different bias points using PETSc and multi-grid.

3D MESFET Results:

Figure 3.13 and Figure 3.14 show the 3D representation of the position averaged potential of a 3D MESFET obtained through multi-grid and PETSc solver. Similar to the case of 2D results, PETSc provides an excellent agreement with the multi-grid solver for the 3D average potential. The current shown on Figure 3.15 also shows great similarity between these two approaches. All these results convey the eligibility of PETSc to be used as a Poisson solver in CMC.

**Performance Comparison for 2D simulation**:

CMC has capability of determining time spent in different parts of the code while running any simulation, where the free flight time and Poisson solver consume higher time compared to other processes. Figure 3.16 and Figure 3.17 show time spent by CMC, while using multi-grid and PETSc solvers, respectively.
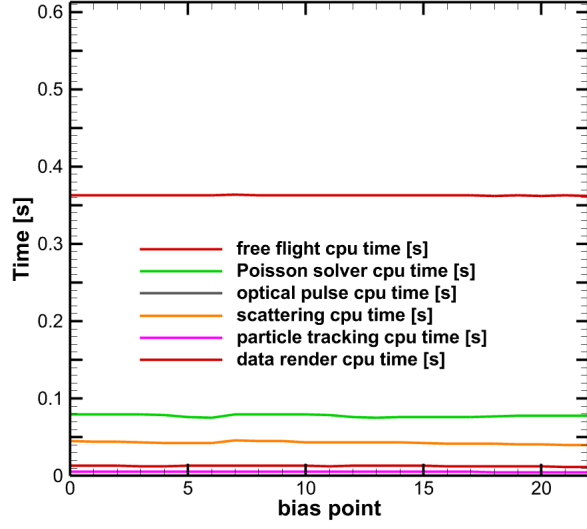
Figure 3.16: Time spent by CMC per iteration while using multi-grid in 2D GaN HEMT simulation.
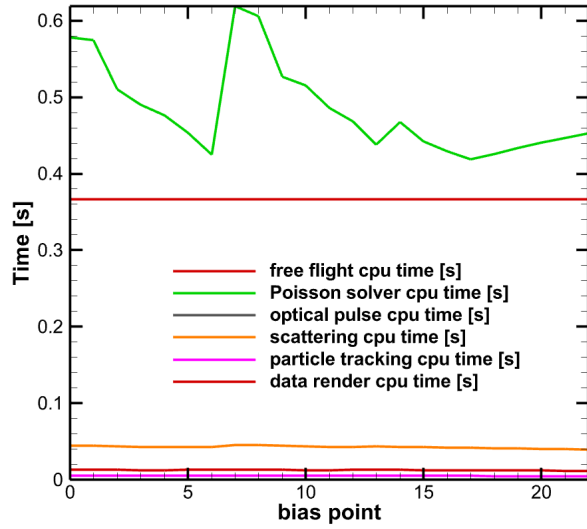
Figure 3.17: Time spent by CMC per iteration while using PETSc in 2D GaN HEMT simulation.

It is evident from the figures that, the current version of PETSc takes longer than multi-grid. On the contrary, the current sequential PETSc solver can be modified to parallelized PETSc , which should be much easier to implement compared to paralleling multi-grid and give a performance gain.

3.9 Conclusion

In this chapter, a popular computational tool, PETSc, has been incorporated as a Poisson solver into the CMC. PETSc solves the linear PDE with an Incomplete LU method as a preconditioner and Krylov subspace method as an iterative solver. The inclusion of PETSc in the CMC code is validated through 2D and 3D simulations with accuracy similar to the well established multi-grid method. Although the speed of PETSc is slower than multi-grid, parallelization of the PETSc solver can make it significantly faster when used with a high number of processors.

Chapter 4

ASSESSMENT OF GATE GEOMETRY IN HEMTS THROUGH CELLULAR

MONTE CARLO SIMULATIONS

4.1 Introduction

As Silicon-based power electronic devices show many limitations such as high leakage current in high temperature, wide bandgap (WBG) semiconductors, such as SiC, GaN and Ga_2O_3 along with their superior performances, are becoming the ultimate choice. The bandgap of these materials ranges from two to four times of silicon [72], which leads to properties like higher blocking voltage, better reliability and efficiency along with superior thermal performance [73]. High bandgap allows WBG-based devices to limit the free charge carrier and leakage current at high temperatures [72]. Additionally, GaN has a higher saturation velocity than Si, which results in higher saturation on-current. This also leads to higher switching speed along with shorter recovery time. Figure 4.1 compares some key parameters related to high temperature, high voltage and high switching operations [4]. From the figure, it is obvious that GaN provides advantages over Si in many aspects.

The growth techniques of group-III nitrides primarily are MOCVD (metal-organic chemical vapor deposition) and MBE (molecular beam epitaxy) [4]. In MOCVD, Ga is supplied from a metal organic compound such as trimethylgallium and it is transported by a gas like hydrogen. For this reason, the vapor pressure of the gas determines the concentration of the compound. In this technique, nitrogen is produced from ammonia. In the MBE process, nitrogen gas is passed through a plasma discharge to produce nitrogen atoms and molecules. Due to its superior thermal conductivity,
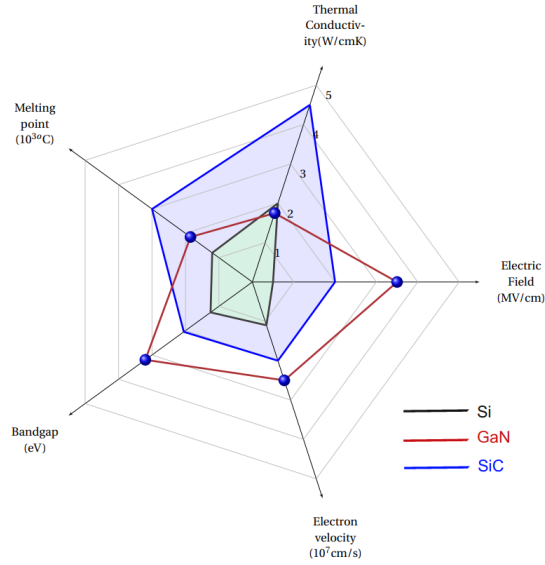
Figure 4.1: Comparison of different key parameters at room temperature between Si and wide band gap materials [4].

SiC is one of the popular substrate for GaN HEMTs [74], whereas Si and sapphire are also used for cost effectiveness [75, 76]. Usually, a nucleation layer of AlN, which is needed to provide electrical isolation, is added between device and SiC/Si substrate. Furthermore, a thin interlayer of AlN is added between AlGaN and GaN to reduce sheet resistance, and hence increasing mobility of the device [77]. The inclusion of the AlN interlayer actually reduce the alloy scattering and increases the amount of sheet charge at the interface due to higher conduction band discontinuity. Finally, the addition of a highly doped GaN cap layer leads to reduction in source to gate access resistance, which ultimately increases gain and efficiency [78].

4.1.1   Polarization in GaN

Figure 4.2 illustrates the hexagonal closed packed (hcp) crystal structure of Wurtzite GaN, where a and c represent the sides and height of the hexagon, with u being the distance between Ga and N neighbors. This structure is orginally formed by two
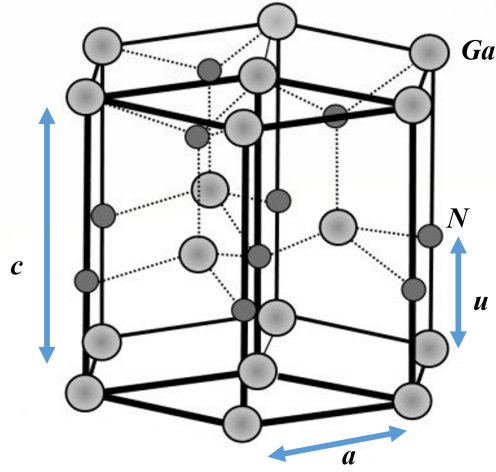
Figure 4.2: Hexagonal Closed Packed (HCP) crystal structure of Wurtzite GaN.

interpenetrating hcps, one for Ga and another for N. As N has a higher electronegativity than Ga, there is a net spontaneous polarization, P_{SP} inside the material along the c-axis [79]. The value for P_{SP} in GaN is reported to be -0.0029 C/m^2 [80].

Furthermore, the interface of AlGaN/GaN can also create a piezoelectric polariza-tion, P_{PE} due to the strain produced by lattice mismatch between the two materials. The value of P_{PE} can be obtained through the Ambacher formalism [81], such as:

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

where a_1 represents the lattice constant for the strained layer, e_{31} and e_{33} are the piezoelectric coefficients, c_{13} and c_{33} are the elastic constants. The values of these parameters are a function of the mole fraction x, of AlGaN, and can be determined by interpolation. Finally, the total amount of polarization charge can be determined from the sum of spontaneous and piezoelectric polarization.

4.2 Formation of two-dimensional electron gas

In conventional HEMTs, the two-dimensional electron gas (2DEG) is formed by the modulation-doped heterostructures, but for the case of AlGaN/GaN, intentional doping is not required due to the bandstructure of these two materials [82]. A detailed illustration of the formation of a 2DEG at AlGaN/GaN interface, along with an expression for the sheet charge density is stated in [83]. As a result of difference in electronegativities, a positive polarization charge is induced at the AlGaN/GaN interface, while negative polarization charge is found at the top of AlGaN layer [83]. This charge leads to the formation of an **E**-field inside AlGaN layer, where the interface and AlGaN surface can be treated as two planes with infinite amount of charge as the epilayer is really thin.

As a result of the **E**-field, the energy band and fermi level will tilt to the direction of the interface. Due to the presence of the field, electrons will cause accumulation near cathode, leaving space charges behind. This bends the energy band and the Fermi level will be flat again. This phenomena is illustrated in Figure 4.3.
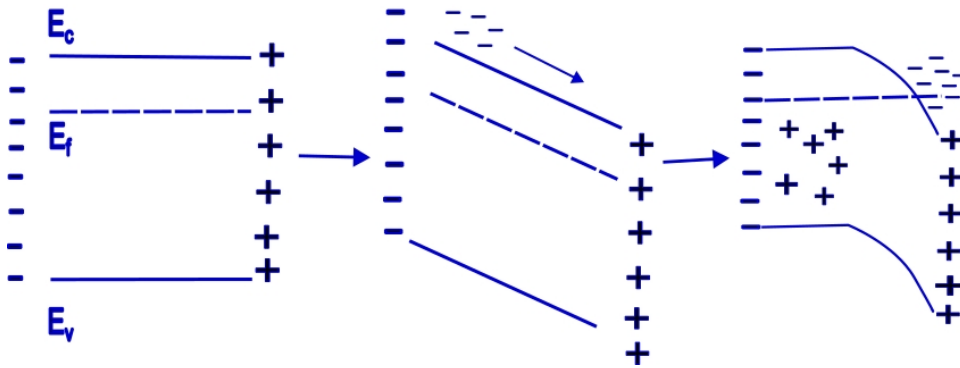
Figure 4.3: **E**-field formation and band bending in a freestanding AlGaN layer.

The formation of 2DEG is shown in Figure 4.4. Due to the band gap, the Fermi

level of GaN is lower than the AlGaN. For this reason, as soon as the contact is formed, the accumulated electrons will travel into GaN, which eventually formes the 2DEG. This migration process is continued untill the Fermi levels of the two materials are aligned.
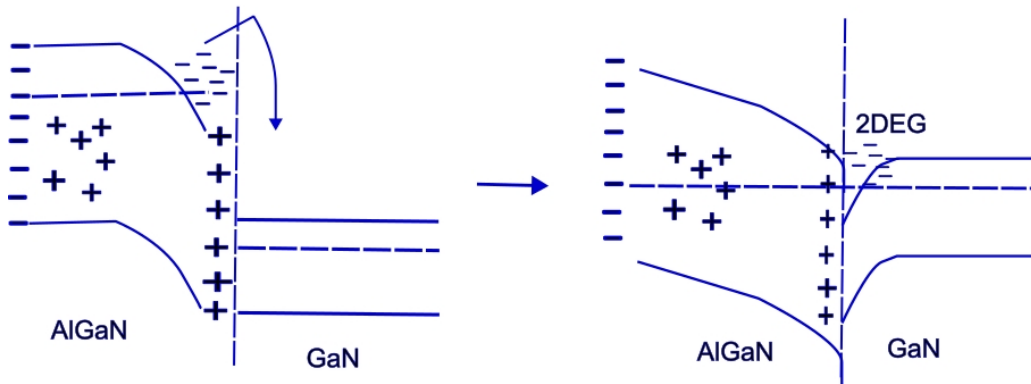
Figure 4.4: Band diagram of a heterojunction formed between AlGaN/GaN, where electrons will flow from AlGaN to GaN forming 2 dimensional electron gas (DEG) in the GaN side.

4.3   Reliability concerns in HEMT and scope of this work

Different physical mechanism can cause device degradation in GaN HEMTs leading to shorter lifetime than expected [84]. Some processes involved in device degradation are: 1) the degradation due to hot electrons, which occurs in ON-state and reducs the drain current because of accumulation of charges near gate-drain access region [85, 86], 2) the deterioration in Schottky contact of gate, which increases gate leakage current because of localized shunt paths near the gate edge [87, 88], 3) the delamination and cracking of the passivation layer due to high temperature or high power operations, which can cause further leakage or charge trapping [89], 4) the generation of defects within AlGaN/GaN heterostructure leading to time-dependent processes [90] and

5)breakdown mechanisms like source to drain breakdown, vertical breakdown because of compensation issues in the buffer and impact ionizations because of electron-hole pairs near gate [84].

Self-heating also causes major detoriation in the performance and reliability in high electron mobility transistors [91–95]. GaN HEMTs are susceptible to phenomena like current collapse [96], threshold voltage shift [97], surface oxidation [98] and creation of traps due dehydrogenation of Ga vacancies [99] which is responsible for further RF performance degradation [100].

Hot electrons are highly energetic electrons with a significantly higher effective temperature than the lattice are principal contributors to device degradation, and are caused by the presence of a strong electric field [9]. Trap states are created because of the injection of hot electrons in the AlGaN layer, the SiN passivation layer or the GaN buffer layer, which causes degradation in performance and reliability [101–103]. An innovative Π-gate HEMT concept was modeled using a Cellular Monte Carlo (CMC) in [104], where the standard T-gate stem was divided into two separate stems, leading to both symmetric and asymmetric gate contacts. Use of an asymmetric Π-gate configuration can reduce hot electron generation up to 40% under DC operation. It was demonstrated [104] that such a reduction in hot electrons was achieved while maintaining similar device performances as that of conventional T -gate. Self-heating effects were not included in that original study, but were incorporated in subsequent CMC device studies [105] by incorporating the temperature dependence of thermal conductivity and the effect of convective boundary conditions at material interfaces. These studies led to investigation of the effect of lateral scaling (scaling of Gate-Source and Drain-Source spacing) in T-gate AlGaN/GaN HEMT and its performance in terms of transit frequency, ON-resistance and DC output characteristics, which were compared with the experimental results reported in [106, 107]. Hot spot

characterization and acoustic and optical temperature maps were obtained in these studies through the modeling of electron-phonon interaction. These studies considered T-gate contacts only for thermal characterization. However, a comprehensive and conclusive study of Π-gate performance under self heating and comparison with T-gate under worst case scenarios is yet to be done.

Sehra *et. al.* have investigated Π-gate in a series of papers [108–112] using the Silvaco's TCAD ATLAS tool. It is worth mentioning that the T –gate geometry has been compared with Π-gate (both symmetric and asymmetric) and effects of the Π-gate pillar height is investigated along with a recessed version of it [108]. It was concluded that the symmetric Π-gate has 10% and 5% lower f_T and f_{max}, respectively with respect to a T–gate of same height (for asymmetric: 1.1% lower and 3% higher, respectively), but by increasing the stem height by nearly three times, an increase by 6.5 % and 5% respectively in f_T and f_{max} with similar DC performance. The recessed gate shows 25% and 13% improvement in f_T and f_{max}, respectively. Furthermore, electrothermal simulations reveal that Π-gate is about 5% cooler than the corresponding T –gate along the channel. In [109], Π-gate and T–gate were compared on linearity and intermodulation distortion (IMD) metrics showing superiority of Π-gate over T–gate by 10% and 30% respectively. Also, decrease in stem height reduces IMD at the cost of reduction in linearity. In [110], The symmetric and asymmetric recessed Π-gate were shown to have 64.5% and 77.5% higher intrinsic gain and improved noise figure (NF) metric over the T–gate, respectively. In [112] efficacy of Π-gate has been investigated taking advantage of thin AlGaN barrier as compared to a similar experimental T –gate showing enhanced Π-gate performance with respect to breakdown voltage and pulsed power applications.

The author would like to express the necessity of the study: at one hand, the commercial device simulators such as Silvaco's ATLAS tool models carrier transport

using drift-diffusion or hydrodynamic model which are obtained from the first and second moments of Boltzmann Transport Equation (BTE) and required an externally supplied closure relation to be solved. On the other hand, to simulate the device self-heating effects the Joule heating model is used with charge transport equation, which dictates heat generation rate as $W = \mathbf{J} \cdot \mathbf{E} \, W/m^3$ for a current density J and electric field E. It is however shown in [1], that the Joule heating model is valid for low electric field $\leq 10^6 \, V/m$ where the carrier – acoustic phonon temperature is nearly same and acoustic phonon are the prime carriers of heat. Under these circumstances, the phonon energy balance equation reduces to the Joule heating model, where the carrier mobility is approximated using at field within the relaxation time approximation. In this work, for device self-heating, we use the local temperature-dependent electron – phonon energy exchange rate and phonon – phonon energy exchange rate as forcing functions in the energy balance equation separately for acoustic and optical phonons yielding device acoustic and optical temperature maps separately, which gives enhanced insight into hot spot characterization and its spread. Finally, Cellular Monte Carlo algorithm solves the complete non-linear full band BTE yielding electron energy distribution function (EDF) as the solution which could be used for analyzing the hot electron effects [104], this can't be done using drift-diffusion or hydrodynamic model of carrier transport.

We feel that, an investigation is needed to test the efficacy of Π-gate vs T-gate under worst possible scenario for hot electron suppression taking self-heating into account to investigate its effect on device performance. The Π-gate does not alter the peak electric field [104] and with gate – drain length scaling the peak electric field increases. So we choose experimental T –gate reported by Altuntas *et. al.* in [6] which is scaled version of T–gate reported by Fitch *et. al.* [107] to be compared with corresponding Π-gate. The device layouts chosen for present study are shown in

Figure 4.5 and Figure 4.6. Further, since the electric field along the channel increases with decrease in stem height, we also consider a shorter version of the Π-gate device with 60 nm stem height. In summary, we compare the T –gate and Π-gate geometries with stem height 120 nm and 60 nm, respectively. The symmetric version of Π-gate has $L_{G1} = L_{G2} = 75$nm (Figure 4.6) whereas the two asymmetric version are: (1) L_{G1} = 50 nm and $L_{G2} = 25$nm (2) $L_{G1} = 25$ nm and $L_{G2} = 50$nm. The Π-gate has a spacer length d = 50 nm separating the two gate stems, therefore we simulate T –gates with an added 50 nm in the source–drain region in order to equate device width for comparison.
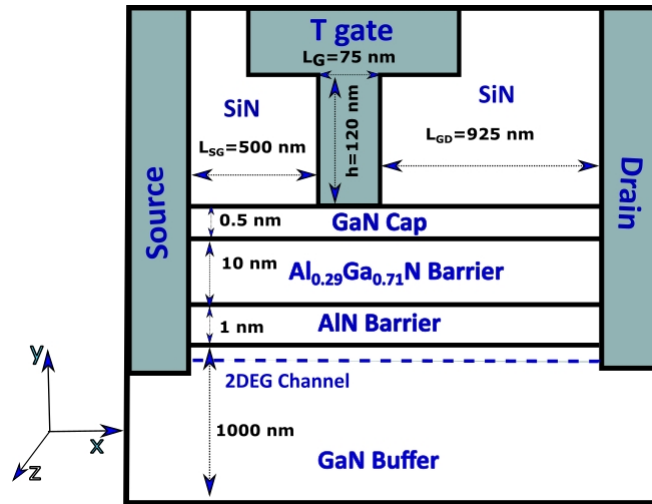
Figure 4.5: Device layout of a T-gate GaN/AlGaN HEMT with a stem length of 120 nm [5].
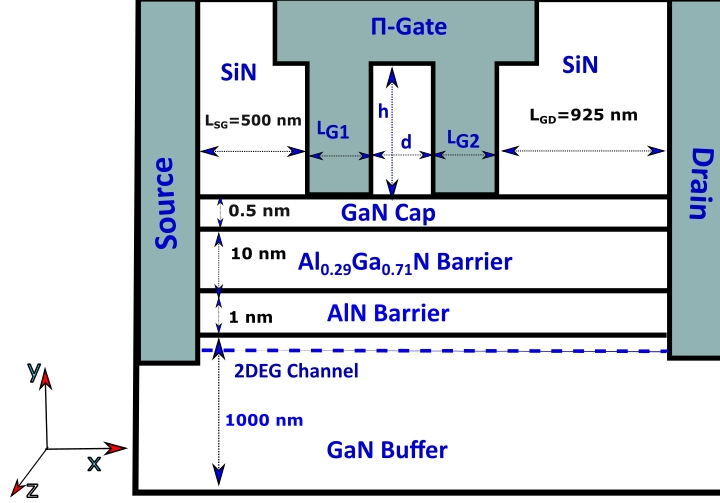
Figure 4.6: Device layout of a Π-gate HEMT, formed by splitting the stem of the T-gate into two stems [5].

4.4   Simulation Setup

The device geometry is represented with a nonuniform tensor product grid. The modeling of the charge carrier dynamics is performed within the CMC framework, in which the trajectories of electrons are tracked in the first Brillouin Zone (BZ1) [22] and scattering events due to piezoelectric (acoustic) phonons, ionized impurities, deformation potential, polar optical phonons and impact ionization are considered. To model the strain effects caused by GaN-on-Si technology, a density of threading dislocations, $T_{DD} = 5 \times 10^9 cm^{-2}$ [113] is taken into account. The charge distribution is modeled as homogeneous along the depth or z-direction of the device, so the corresponding component of the electric field is effectively zero. Therefore, the computational cost is minimized by modeling only the charge transport in 3-D momentum space, while the electrostatics is solved in 2-D on the xy-axes of Figure 4.5, with an arbitrary depth used for normalization. Some of the material parameters used for this study are tabulated in Table 4.1.

Schottky Barrier Height	0.86 eV [114]
Polarization sheet charge density	0.1055 C/m^2 [81]
Surface Charge density in AlN/AlGaN interface	-0.0826 C/m^2 [81]
Surface Charge density in AlGaN/GaN cap interface	-0.0229 C/m^2 [81]
Surface Charge density in GaN cap/SiN interface	0.04 C/m^2 [81]
Longitudinal elastic constant	265 GPa [115]
Transverse elastic constant	44.2 GPa [115]
Optical phonon energy	0.091 eV

Table 4.1: Table with different material parameters used for this study.

4.4.1   Energy Balance Equation

The electrons in the channel loose energy to highly energetic optical phonons and to some extent to acoustic phonons. The optical phonons decay into acoustic phonons due to anharmonicity of the crystal at timescales higher than electron-phonon scattering rate. Optical phonon decay into acoustic phonons having a higher group velocity and contribute more to the thermal conductivity before being absorbed at the heat sink [12].

The process of heat generation and propagation can be accurately represented with a particle-flux electro-thermal simulator by solving the Energy Balance Equation

(EBE) for phonons [116] in each cell of the real space grid of the device, while for electrons traditional MC algorithm is applied. The phonon EBE is derived directly from the phonon Boltzmann Transport Equation [21]. Furthermore, the EBE can be solved for each individual phonon mode or for a group of modes, *e.g.*, for the acoustic and optical phonon modes resulting in both an acoustic and a optical temperature map. The EBE for each phonon mode, or group of phonons, μ is expressed as

$$\frac{\delta W_\mu}{\delta t} = -\nabla \cdot \mathbf{F}_\mu + \left.\frac{\delta W_\mu}{\delta t}\right|_{e-p} + \left.\frac{\delta W_\mu}{\delta t}\right|_{p-p} \tag{4.1}$$

where $W_\mu(\mathbf{r}, t) = (1/\Omega) \sum_{\mathbf{k}} (E_\mu(\mathbf{k}) f_\mu(\mathbf{r}, \mathbf{k}, t)$ is the ensemble energy density within the volume Ω of the reciprocal space, $\mathbf{F}_\mu(\mathbf{r}, t) = (1/\Omega) \sum_{\mathbf{k}} \nu(\mathbf{k})(E_\mu(\mathbf{k}) f_\mu(\mathbf{r}, \mathbf{k}, t)$ is the energy flux, and two partial derivatives on the right-hand side represents contributions to W_μ due to electron-phonon $(e-p)$ and phonon-phonon $(p-p)$ interactions, respectivley.

If steady-state conditions are considered, the left-hand side of Equation 4.1 is null and the total heat generation rate, P_μ can be found from the sum of the electron-phonon and phonon-phonon interaction terms

$$\nabla \cdot (\kappa_\mu(\mathbf{r}, T) \nabla T) = -\left(\left.\frac{\delta W_\mu}{\delta t}\right|_{e-p} + \left.\frac{\delta W_\mu}{\delta t}\right|_{p-p} \right) = -P_\mu \tag{4.2}$$

where κ_μ is the thermal conductivity of phonon mode μ and the energy flux has been approximated by using Fourier's law: $\mathbf{F}_\mu(\mathbf{r}) = -\kappa_\mu(\mathbf{r}, T) \nabla T$.

An efficient elliptical partial differential equation (PDE) solver is included in most device simulators for solving Poisson's equation. Manipulation of Equation 4.2 into the form of an elliptical PDE therefore leads to an easier implementation. In the finite difference framework, the thermal conductivity $\kappa_{\mu,C}$, which is originally dependent on both position and temperature, is considered to be constant within the cell and to vary from one cell to another. To take the temperature dependence of the

thermal conductivity into account, the Kirchhoff transformation [117] can be used to manipulate Equation 4.2 into an elliptical Poisson-like PDE. Using the Kirchhoff transformation a new variable Θ, can be defined as the "apparent" temperature as [117]

$$\Theta_{\mu,C}(T) = T_{ref} + \frac{1}{\kappa_{\mu,C}(T_{ref})} \int_{T_{ref}}^{T} \kappa_{\mu,C}(\tau)\partial\tau, \tag{4.3}$$

where $\kappa_{\mu,C}(T_{ref})$ is the independently known thermal conductivity in a particular cell C at a reference temperature T_{ref}. In this work, a power law of the form $\kappa_{\mu}(T) = \kappa_{ref}(T/T_{ref})^{\alpha}$ is adopted to model the temperature-dependent thermal conductivity, where α is a fitting parameter [118].

The EBE can then be expressed as a elliptical PDE for the apparent temperature Θ

$$\nabla^2 \Theta_{\mu,C} = -\frac{P_\mu(\mathbf{r})}{\kappa_{\mu,C}(T_{ref})} \tag{4.4}$$

The linearity of Equation 4.4 is subject to the conditions that the temperature and its derivative are continuous in the normal direction of the material boundaries [119]. As a consequence, hetero-structures are subject to the condition that the value of α in the temperature dependence must be the same in every material.

The forcing function, $P_\mu(r)$ in Equation 4.4, can be replaced with the RHS of Equation 4.2 in Equation 4.4, where the electron-phonon $(e-p)$ interaction is computed from the actual energy exchanged during acoustic or optical phonon scattering, rather than being approximated with the Relaxation Time Approximation (RTA) [18] as done in earlier studies [12]. The phonon-phonon $(p-p)$ interaction is modeled by using an RTA as

$$\frac{\partial W_\mu}{\partial t}\Big|_{p-p} = C_i \left(\frac{T_i - T_j}{\tau_{i-j}}\right), \tag{4.5}$$

where C_i is the heat capacity of i_{th} phonon mode, $T_{i,j}$ is the temperature of two phonon

modes and τ_{i-j} is the decay time from i^{th} to j^{th} mode. In this work, Equation 4.5 represents the expression of the anharmonic decay of optical phonons into acoustic ones. In this approach, simultaneous solutions of two sets of Equation 4.4 are carried out, one for the acoustic phonon mode group and another for the optical phonon mode group.

Different mechanisms involving electron scattering are discussed in Chapter 2. The electron scattering rates in each cell are updated using rejection algorithm according to the new temperature and then a new heat generation is calculated. This process is continued until heat generation rate is stabilized and convergence is achieved for the temperature maps in an iterative loop.
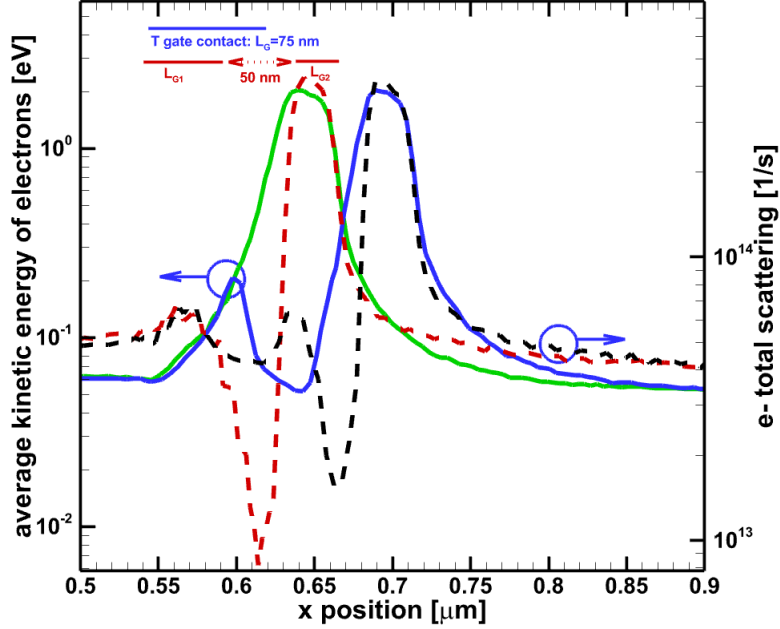
Figure 4.7: Kinetic energy (green (T-gate) and blue ($\Pi^{(50,25)}$-gate) solid lines) and electron scattering rate (dashed red (T-gate) and black ($\Pi^{(50,25)}$) lines) along the channel at bias point $V_{DS} = 10$ V, $V_{GS} = 2$ V. Both gates have a stem length of 60 nm.

Figure 4.7 shows the kinetic energy and scattering plots of electrons along the channel for one T-gate and one Π-gate device. The average electron energy under the gate for T-gate is 0.24eV, which reduces to 0.21eV for Π-gate, whereas it remains at 0.047eV outside the gate region for both devices. These high energetic electrons are emitting optical phonons and eventually lose energy at the end of the gate region. The temperature maps for the acoustic and optical modes, respectively, obtained at the bias point V_{DS}=10 V, V_{GS}=2 V are shown in Figure 4.8 and Figure 4.9 for a T-gate of gate

length $L_G = 75$ nm and stem height 60 nm and a Π-gate having gate lengths $L_{G1} = 50$ & $L_{G2} = 25$ nm and stem height of 60 nm, respectively. For these bias conditions both high current and electric field are present, while the DC power remains at a moderate level. The peak temperature is on the drain side of gate due to increased electron-optical phonon scattering rates as shown in Figure 4.7. The maximum acoustic and optical mode temperature within the entire device are extracted, and are shown in Table 4.2 for each device layout. For the acoustic phonon mode, the peak temperature is around 370 K in all devices but the optical temperatures are significantly different in the T-gate and Π-gate devices. The peak optical temperature for T-gate devices is around 550 K , while for the Π-gate it is around 10 K lower. It has been found that the difference of the optical temperatures between the T-gate and Π-gate comes from the extra 50 nm length between the two stems of the Π-gate, which allows for increased relaxation of optical phonons. Therefore, for comparative purposes we show simulations results of a modified version of the T-gate layout fabricated by Altuntas *et. al.*, where the source-to-gate access region length has been increased by 50 nm (T^*_{120}). It is seen that T^*_{120} has the lowest optical temperature of 534K, whereas the asymmetric $\Pi^{(50,25)}_{120}$ has value of 539K. Also, for a reduced stem height $\Pi^{(50,25)}_{60}$ shows an increase of optical temperature to 543K, while reducing the acoustic temperature to 366K as compared to 371K for a stem height of 120nm. This behavior is due to an increase in electric field along the channel favoring electron-optical phonon energy exchange over acoustic phonons. This trend of acoustic/optical temperature variation with stem height is same across all devices.
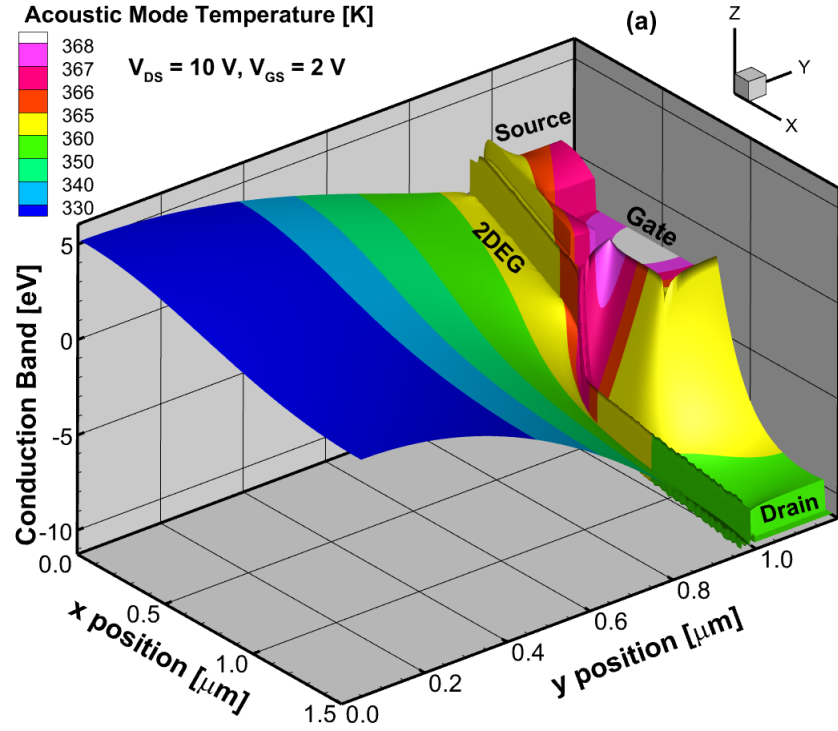
(a) Acoustic mode
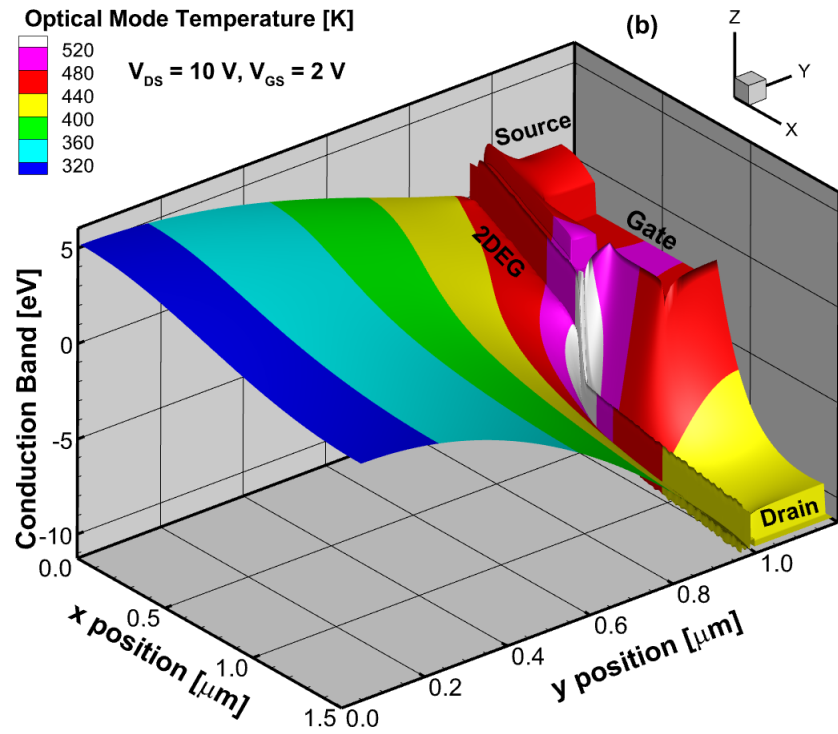
(b) Optical Mode

Figure 4.8: Temperature Map of acoustic (a) and optical (b) mode for a T-gate
HEMT with 60 nm stem length.
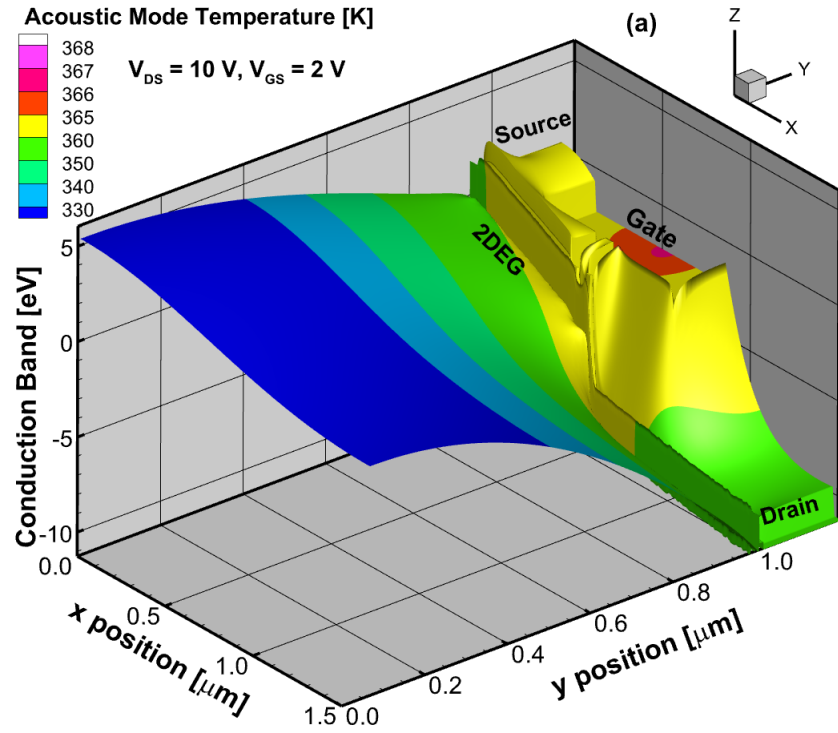
(a) Acoustic mode
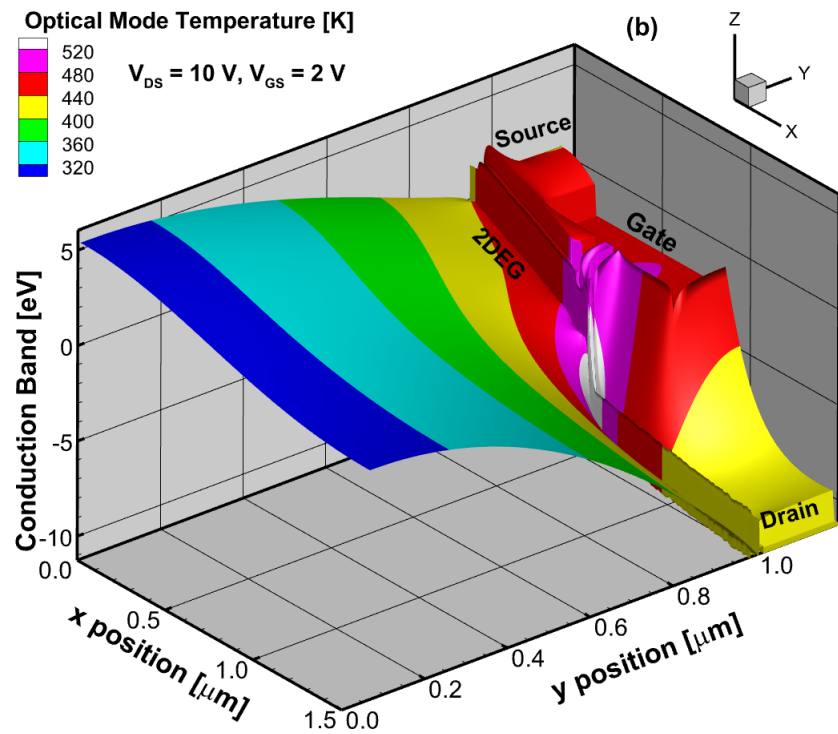
(b) Optical Mode

Figure 4.9: Temperature Map of acoustic (a) and optical (b) mode for a Π-gate
HEMT with gate lengths of 50 & 25 nm and a 60 nm stem length.

Gate Architectures	Peak acoustic temperature (K)	Peak optical temperature (K)
T-gate (Stem length 120 nm)	373	549
T-gate (Stem length 60 nm)	368	556
T-gate (Stem length 120 nm, with an extra 50 nm on L_{SG})	367	534
T-gate (Stem length 60 nm, with an extra 50 nm on L_{SG})	365	541
Π-gate (Gate lengths 25,50 nm & Stem length 120 nm)	369	539
Π-gate (Gate lengths 25,50 nm & Stem length 60 nm)	367	547
Π-gate (Gate lengths 37.5, 37.5 nm & Stem length 120 nm)	369	539
Π-gate (Gate lengths 37.5, 37.5 nm & Stem length 60 nm)	366	546
Π-gate (Gate lengths 50, 25 nm & Stem length 120 nm)	371	539
Π-gate (Gate lengths 50, 25 nm & Stem length 60 nm)	366	543

Table 4.2: Table with the peak temperatures for acoustic and optical phonon modes for different devices.
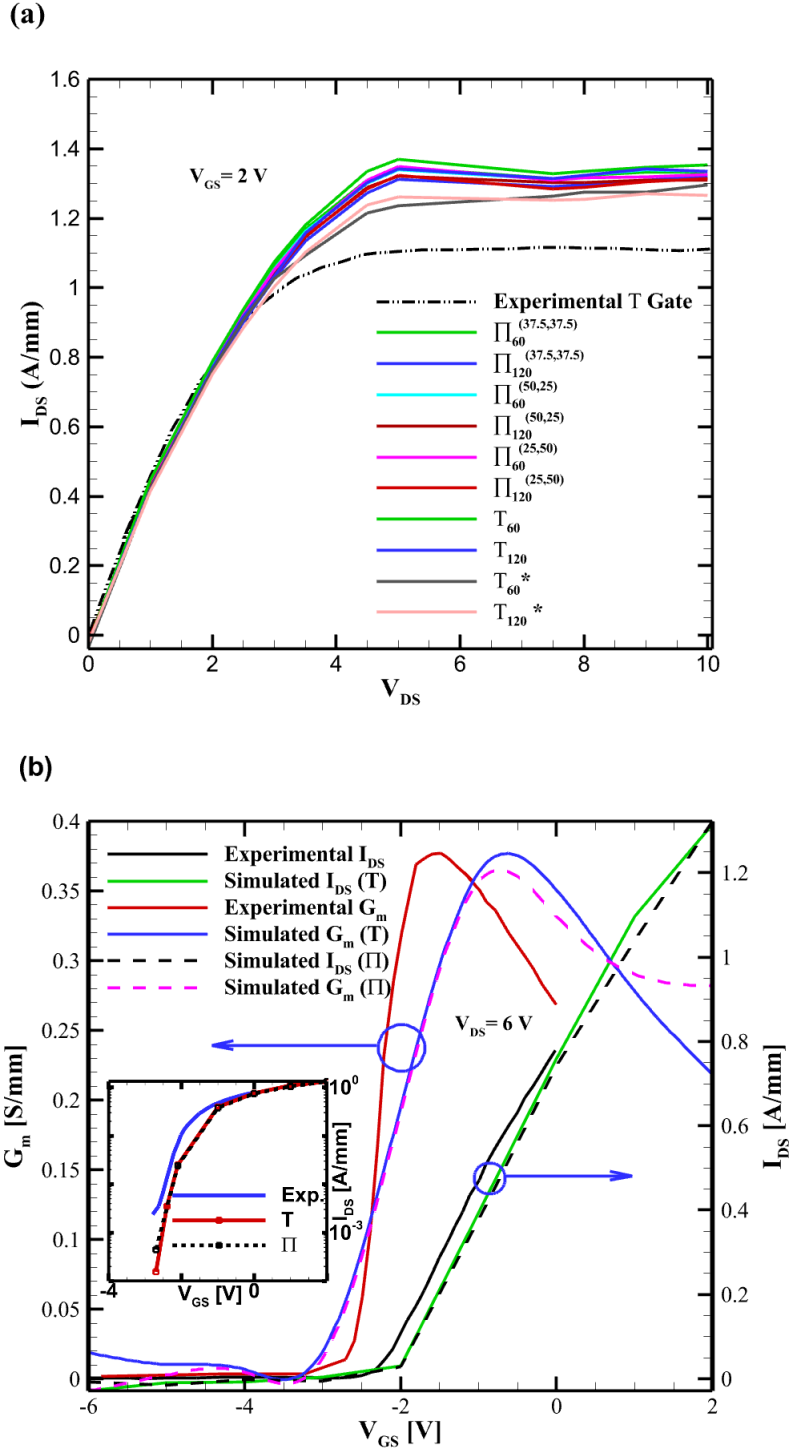
Figure 4.10: DC output characteristics (a) of different T and Π-gates at $V_{GS} = 2$ V, where the subscripts indicate the height of the stem and the superscripts indicate the lengths of the gate. Transfer characteristics (b) shown on linear and semi-logarithmic scale (inset) for a simulated T-gate and a Π-gate with a 120 nm stem length for $V_D = 6$ V. The experimental values refer to the T-gate fabricated by Altuntas *et. al.* [6]

74

Figure 4.10 shows the DC output characteristics of each of the ten devices we studied (four T-gate and six Π-gate structures), with the gate bias held constant at $V_G = 2$ V. It is evident that they show similar $I - V$ characteristics and the simulated current is slightly higher compared to the experimental data observed by Altuntas *et. al.* [6]. The reason behind the overestimation of current is due to a lower threshold voltage due to absence of the substrate in the present simulation, resulting in a higher charge density in the channel at the same gate voltage when compared to the experimental structure. For T_{120}^* a reduction of 7% is seen in the drain saturation current due to an increase in the source access resistance causing a reduction of the effective V_{GS}. Figure 4.10 also shows the transfer characteristics along with the transconductance G_m of both T-gate and Π-gate devices with a stem length of 120 nm, which shows that both I_{DS} & G_m closely follow the experimental measurement of Altuntas *et. al.* [6]. The experimental measurement shows the threshold voltage of the device as -2.3 V, while our simulated threshold voltage is around -2.1 V, with the slight difference again arising from the absence of the substrate.

Using the DC electrical and thermal results as a starting point, small-signal AC simulations were performed of all 10 layouts. The current gain as a function of the frequency is shown in Figure 4.11 as obtained by Fourier decomposition (FD) [120]. In CMC, any field effect transistor (FET) is represented as a two-port network. The Y-parameters are then extracted by applying two perturbations separately to gate and drain, respectively. The sinusoidal perturbations applied on top of DC operating point $V_{DS} = 10$ V and $V_{GS} = 2$ V.

From these RF simulations, it is seen that the current gain is consistently reduced in a more realistic electrothermal (self-heating is included) simulation as compared to an isothermal (self-heating is not included) 300K simulation due to an increase in total scattering. As a consequence, the cutoff frequency is around 20GHz lower
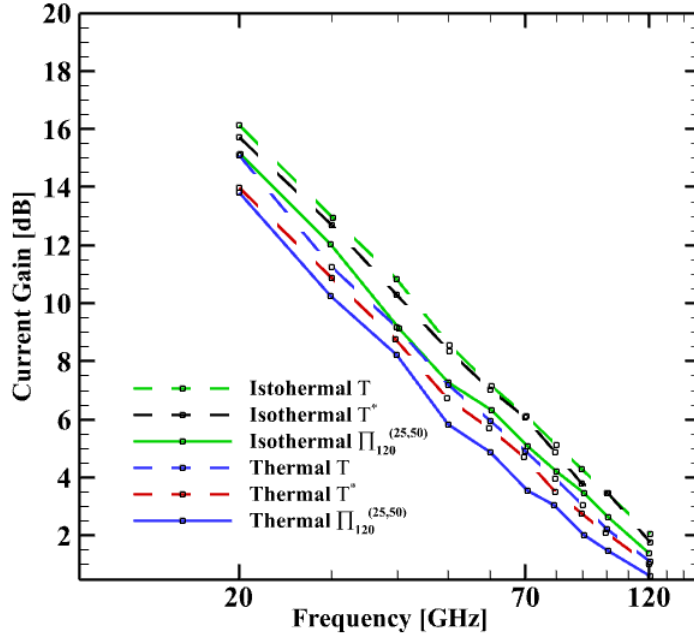
Figure 4.11: Current gain vs frequency for a T-gate of stem length 120nm and gate length 75 nm and for a Π-gate HEMT with 25 & 50 nm gate lengths and a stem length of 120 nm .The symbol * represents the T gate with extra 50nm added on L_{SG}.

than the isothermal case for all of the devices considered in this study. Figure 4.12 represents the cutoff frequencies extracted from the simulations of different devices. These results can be explained with the definition of the cutoff frequency,

$$f_T = \frac{g_m/(2\pi)}{[C_{gs} + C_{gd}] \cdot [1 + (R_s + R_d)/R_{sd}] + C_{gd} \cdot g_m \cdot (R_s + R_d)} \tag{4.6}$$

where g_m, C_{gs}, C_{gd}, R_s, R_d and R_{sd} are transconductance, gate-to-source and gate-to-drain capacitance, source and drain parasitic resistance and source to drain resistance, respectively [121]. If the parasitic resistances R_s and R_d are considered negligible, then (1) can be expressed as:

$$f_T = \frac{g_m}{(2\pi) \cdot [C_{gs} + C_{gd}]} \tag{4.7}$$

76
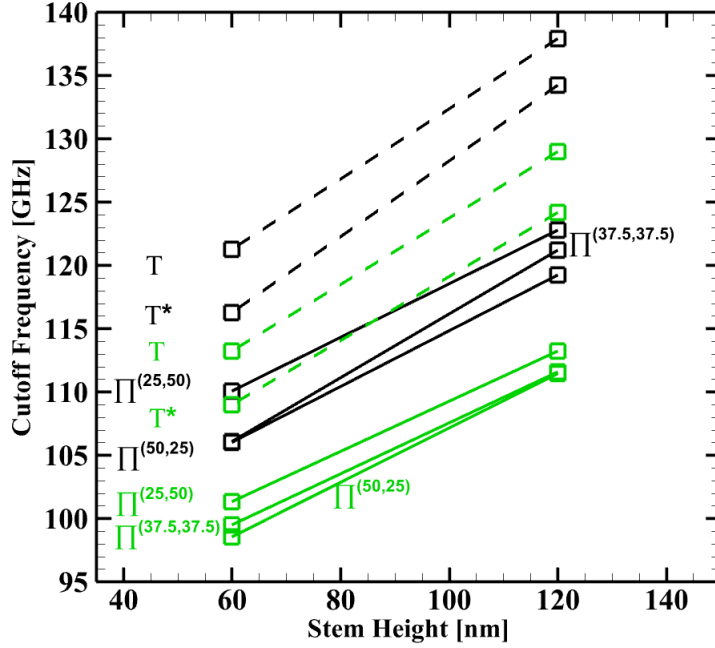
Figure 4.12: Cutoff frequencies for different T-gate and Π-gate layouts, the symbol *
represents the modified devices with an extra 50 nm on L_{SG} .

Therefore, with an increase in g_m and decrease in transit time under the gate, the
cutoff frequency f_T can be improved. The Π-gate devices have lower carrier velocity
(87.8 Km/s) than the T-gate devices (89.6 Km/s) and so have higher (5.3%) transit
time under the gate. From Figure 4.12 it is evident that T-gate device with same stem
length shows better cutoff frequency than the Π-gate device. Devices with 60nm stem
height have lower cutoff frequency than their 120nm counterparts due to increased
capacitance (C_{gs} and C_{gd}). The Π-gate with gate lengths of 50 & 25nm and a stem
length of 60nm shows the lowest cutoff frequency of 98 GHz for the electrothermal
case, while the T-gate with a stem length of 120nm shows the highest cutoff of 138
GHz for the isothermal case. The electrothermal cutoff frequency for the T-gate with a

77

stem length of 120nm is around 128.97 GHz, where the experimental cutoff frequency is 116 GHz. Therefore, with the introduction of electrothermal solver, the simulated RF performance becomes significantly closer to the experimentally measured value. The slight over estimation of the cutoff frequency is due to the increased G_m as shown in Figure 4.10 resulting from the absence of substrate. All devices show a significant decrease of about 7% in cuttoff frequency from isothermal to electrothermal simulation, with the highest one being 7.95%.
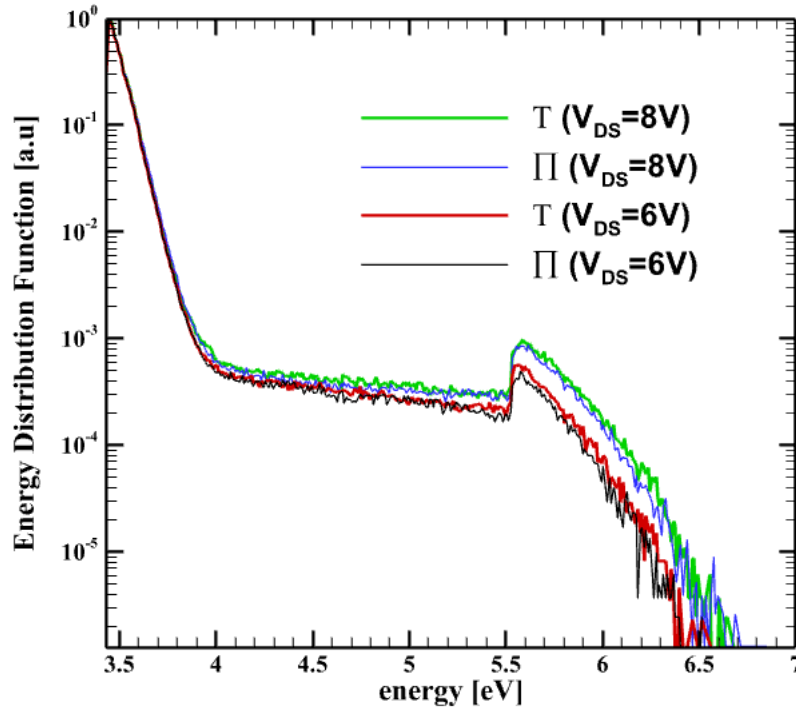
Figure 4.13: Electron energy distribution function for $V_{GS} = $ -0.645V and $V_{DS} = $ 6V & 8V.

As it has already been shown, the asymmetric Π gate geometry with $L_{G1} > L_{G2}$ is most effective over other Π-gate configurations in hot electron suppression, hence we choose the asymmetric Π gate layout to evaluate hot electron mitigation [104]. In

GaN, the hot electrons responsible for trap generation are with energy greater than 5.5 eV [104]. The DC power (P_{dc}) increases with V_{DS}, so we compare the Energy Distribution Function (EDF) for T-gate and Π–gate for energies exceeding 5.5 eV as shown on Figure 4.13. The Π –gate shows 19.75% and 18.07% decrease over T–gate for $V_{DS} = 6$ V and 8 V, respectively corresponding to a DC power of 2.5 W/mm and 3.6 W/mm, respectively. Although the DC power extracted by the device depends on the DC load line imposed by external circuit, we have chosen V_{GS} corresponding to device maximum transconductance and varied V_{DS} from 6V – 10V.

4.6   Conclusion

In this chapter, we compared of various T and Π-gate HEMT architectures through electrothermal simulation using CMC the device simulator. Reduced stem height geometries were considered to facilitate hot electron generation so as to assess Π-gate's effectiveness in their suppression. Similarly, an extra T –gate configuration with added 50nm (T_{120}^*) in source –gate was chosen to make the source-drain distance similar for both T and Π-gate. The DC characteristics of all gate configuration were found to be similar, except that for a drain current decrease of 7% in the T_{120}^*- due to the increase in source access resistance and reduced effective V_{GS}. In RF simulations, the transit cut-off frequency shows 7% drop for the electrothermal case as compared to the isothermal case. T–gate device with 120 nm stem height shows highest cut-off frequency of 137.8 GHz/128.97 GHz for isothermal/electrothermal scenarios. The asymmetric Π-gate shows hot electron suppression up to 19.75% corresponding to a DC power of 2.5 W/mm . The temperature maps show that the Π-gate devices have nearly the same acoustic temperature but lower optical temperature than the corresponding T –gate structures. The peak optical temperature for Π-gate increases with decreasing stem height due to increased electric field along the channel facilitating

electron – optical phonon scattering. Though the T_{120}^* gate shows the lowest optical temperature of 534K and cutoff frequency of 129 GHz, its apparent advantage as a power transistor is reduced due to lower transconductance (due to reduced effective V_{GS}) and its inability to suppress hot electrons.

Chapter 5

DETERMINATION OF THERMAL PROPERTIES OF DIAMOND USING CELLULAR MONTE CARLO DEVICE SIMULATOR

5.1 Introduction

Carbon is the 6th most commonly found material in the entire universe [122]. Due to its position in periodic table, it has the unique capability of forming tetrahedral, hexagonal and linear bonds. The valence electrons present in carbon can form hybrid orbitals, which leads to this unique bond formation [123]. sp^3, sp^2 and sp hybridization lead to tetrahedral, hexagonal and linear bonds, respectively. For sp^3 hybridization, four of each valence electrons forms σ-bond with four neighboring atoms. In sp^2, three electrons form a σ-bond but the fourth one forms a π-bond. In sp hybridization, two electrons form σ-bond and other two form π-bond. These differently bonded structures show different physical and chemical properties, and this phenomena is called as allotropism. Diamond and graphite are two most common allotropes of carbon. While in diamond the hybridization is sp^3 and the structure is tetrahedral, on the contrary graphite is sp^2 hybridized and so it has a trigonal planar structure [124]. Due to their crystal structure, diamond has high thermal conductivity and negligible electric conductivity, while graphite shows high electrical conductance but lower thermal conductance.

As an ultra wide-bandgap (UWBG) material, diamond has been a material of interest due to its unique properties. Some notable properties of diamond include high electron and hole mobility of about 2000 cm^2/Vs, breakdown field of 10^7 V/cm, Debye temperature of around 1900K and thermal conductivity around 25 W/cm·K [125]. A

comparison of different properties of diamond with other commonly used semiconductors is shown in Table 5.1. It is obvious that diamond posess a significant advantage for thermal applications over its counterparts.

Properties/Materials	Diamond	Si	4H-SiC	GaN
Bandgap (eV)	5.47	1.12	3.26	3.5
Electron Mobility (cm^2/Vs) at 300 K	$1,900 - 2,300$	1,500	900	1,250
Hole Mobility (cm^2/Vs) at 300 K	$1,500 - 2,300$	600	100	200
Dielectric Constant	5.7	11.9	9.7	9.5
Thermal Conductivity W/cm $\cdot$ K)	25	1.48	4.9	1.3
Electron Saturation Velocity (10^7 cm/s)	2.7	1	2.7	2.7
Breakdown Field (10^5 V/cm)	100	3	30	30
Debye Temperature (K)	1,860	645	1,200	608
Hardness (kg/mm^2)	10,000	1,000	4,000	
Johnson's FoM	81,000	1	278	215
Baliga's FoM	25,100	1	125	187
Bipolar Power Switching Product	$1,426,711$	1	748	560

Table 5.1: Comparison of material parameters of diamond with other popular semiconductors [126]

Natural diamond and sysnthetic diamond occur in face centered cubic structure, with a lattice constant of 3.56 Å. Diamond technology is becoming more mature due to continued research over the last 60 years. The growth mechanisms of synthethic diamond primarily involve two techniques: High-Pressue High-Temperature (HPHT) method [127] and Chemical Vapor Deposition method [128]. In HPHT, graphite is converted to diamond either directly or with the help of any catalyst, taking advantage

of high thermal stability of diamond. There are couple of CVD techniques available for diamond growth, such as: plasma-jet CVD [129], microwave plasma CVD [130] and hot filament CVD [131]. In these methods, a diamond seed is placed in a highly heated carbon rich gas. Due to ionization of the gas, pure carbon molecules from the gas are attached to the seed forming the synthetic diamond.

A strong isotope effect is observed on the materials that have a lower atomic mass, along with the compounds with higher ratio of mass difference between isotopes to the atomic mass [132]. Diamond formed with light carbon atoms, has mass difference of 8.3% (while it is 3.5% for Si) between its two isotopes found in nature (98.9% ^{12}C and 1.1% ^{13}C). Furthermore, strong interatomic bond, low anharmonicity of interatomic potential and unsually weak Umklapp three-phonon process at room temperature contribute more to pronounced isotope effect in diamond [133]. Due to higher atomic density and bond energy density, a larger amount of energy is needed for defect formation or elemental substitution in diamond, which makes it only susceptible to easily dissoluble elements like hydrogen, nitrogen and boron [134].

Figure of merit (FOM) is commonly used to quantify material's eligibility to be used in high frequency, high power and high temperature operations [135]. Chaudhary *et. al.* [136] reviewed the values of figure of merit like Baliga's FOM [137] and Huang's FOM [138] in details, from where it is obvious that diamond has high values of FOM compared to its counterparts. To take the advantage of its high thermal conductivity, diamond is being used for different thermal applications in semiconductor devices. Particularly for AlGaN/GaN High Electron Mobility Transistors (HEMT), where self-heating leads to a decrease in electron mobility along with reduced power added efficiency (PAE), a device level heat spreader induced near the hot spot can play a major role minimizing the temperature of the channel [139, 140]. In [141], it was shown that use of diamond instead of SiC as a substrate, can lead to 40-45% decrease in

operating junction temperature. At the same time, GaN HEMT on diamond provided 3.6 times higher RF power than GaN HEMT-on-SiC within same active area [142, 143]. Lu *et. al.* presented a detailed review on the integration of Diamond as heat spreader in HEMT for techniques like GaN epitaxial on-diamond, transfer of substrate with diamond, capped diamond on top and dual side heat spreader [144]. Diamond Schottky barrier diode (SBD) are being studied extensively, where a maximum breakdown field of 9.5 MV/cm [145], high operation current greater than 20A [146], high blocking voltage above 10KV [147] and a low on-resistance have been obtained [148]. Diamond SBD showed a leakage current of less than $10^{-7} A/cm^2$, which is 2-4 orders of magnitude lower than SiC SBD due to larger barrier height [149]. Ongoing research works related to transistors based on diamond are [150]: Hydrogen terminated accumulation FETs (H-FETs), Oxygen terminated inversion channel FETs (I-FETs), metal-semiconductor FETs (MESFETs) and junction FETs (JFETs). Prior efforts to show the bipolar mode of operation using diamond failed due to high resistivity of n-type base layer along with low diffusion length of minority carriers (holes) [151], but due to significant progress in n-type doping technology there has been successful fabrication of BJTs [152]. Deep depletion metal-oxide-semiconductor FET (D3MOSFET) has been fabricated recently and a critical electric field of $5.4 MV cm^{-1}$ has been achieved at a drain-source bias of -175V [150]. A two-dimensional hole gas (2DHG) can be formed near the hydrogen-terminated surface to obtain close to zero activation energy hole channel, a vertical p-channel MOSFET based on 2DHG is fabricated here [153] where a on/off ratio close to 10^8 has been demonstrated.

In a nutshell, the theoretical interpretation of calculation of isotope effect on thermal conductivity mostly revolves around Debye model, Callaway model [154] , numerical solution of Boltzmann transport equation (BTE) and kinetic collective model [155, 156]. Normal (N) scattering doesn't contribute to the thermal resistance directly

and it was ignored in the Debye's model. An excellent integration of N scattering into the Debye's model using Callaway formula [154] was done by Novikov *et. al.* [157], where thermal conductivity was computed over a wide range of temperature range along with the effect of isotope and grain size on it. Asen-Palmer *et. al.* also modified the Debye-Callaway model to incorporate the effect of both longitudinal and transverse phonons [158], which was later used by Morelli *et. al.* to show estimate the isotope effect on diamond, germanium and silicon where Gruneisen constants and sample diameter were used as fitting parameters [159]. Han *et. al.* used a two-step model for explaining contribution of N scattering on thermal conductivity of dielectric materials [160] and diamond [161]. Firstly, the low frequency phonons are transferred to a high frequency region because of N processes and on the next step they go through resistive processes. Effect of isotope scattering on thermal conductivity of diamond was also analyzed by Sparavigna *et. al.* through an iterative solution of phonon Boltzmann transport equation (BTE) without any approximation about relaxation time [162], where three-phonon normal and umklapp process are treated rigorously as an effect of real lattice dynamics. In [163], a detailed derivation for phonon-phonon scattering rate has been given using anharmonic continuum theory, where it is also shown that inclusion of drift term for normal process into the Callaway's model leads to a significant contribution to the thermal conductivity above 150K. A first-principle based approach to solve BTE to compute thermal conductivity of diamond has been used by Ward *et. al.* [133], where density-functional perturbation theory [164] has been used to calculate the harmonic and anharmonic interatomic force constants which are used as input to BTE. Recently, Inyushkin *et. al.* measured thermal conductivity of polycrystalline diamond along with natural and isotopically enriched form, also using a detailed Callaway's theory considering scatterings due to external boundaries, point defects, dislocations, dopants and grain boundary validated his experimental

85

results [165].

While all of these theoretical analysis provide background for assessing the quantitative effect of isotope solely on thermal conductivity in diamond, the authors feel the scarcity of study that sheds light on isotope effect on different macroscopic properties of phonon like mean free path, relaxation time etc. that eventually lead to thermal conductivity. For micro-scale devices, if the characteristic device length is much larger than MFP, then heat transport is entirely diffusive and the effective thermal conductivity is similar to the bulk thermal conductivity [166]. If the characteristic length is less MFP of phonon, the heat transport becomes nondiffusive due to boundary scattering. For instance in diamond nanowires, [167] shows that the the effective thermal conductivity reduces half of its bulk value when the diameter is 670nm. For this reason, this study involves effort to define maximum achievable thermal conductivity in terms of mean free path in semiconductor devices. Moreover, to the best of authors knowledge, there is no study discussing the specific contribution from different phonon modes and their change along with isotope composition. Mode-wise contribution is particularly important for band structure engineering. Qian [168] *et. al.* summarized the effort on band structure engineering of phonons that ultimately lead to a increase in thermal conductivity of BAs to 1200 $Wm^{-1}K^{-1}$ [169] from 200 $Wm^{-1}K^{-1}$ [170] at room temperature. It was possible because of some first principles calculations [171, 172] that investigate details of phonon-phonon interactions in BAs. In this study, for similar reason, specific contribution from different modes of phonon are studied so that further investigation can be conducted to increase thermal conductivity of diamond.

A particle-based Ensemble Monte Carlo (EMC) method [10] can provide a space-time solution of BTE, where trajectories of simulated particles are tracked in momentum and real space. A Monte Carlo algorithm is used to determine the scattering

86

process and the final states after scattering along with probability are computed during the run time. As an extension of EMC, Cellular Monte Carlo (CMC) method [58] stores the final states along with probabilities in a large lookup table which makes it computationally faster than EMC. A Cellular Monte Carlo (CMC) simulator, which solves BTE including three phonon interactions was presented in [173], which was validated using thermal conductivity of Silicon as a function of isotopic composition, temperature and film thickness. In this study, the same tool have been used for the case of diamond.

5.2   Simulation Setup

Under semi-classical approach, charge transport can be accurately depicted by a seven-dimensional distribution function $f(\mathbf{r}, \mathbf{q}, t)$, where $\mathbf{r}$ and $\mathbf{q}$ are the position and wavevector of the carrier at a specific time t. Original Boltzmann Transport Equation (BTE) assumed binary collisions only, while Peierls $et.$ $al.$ [21] implemented a method for heat transport by particles, where non binary collisions are included. It is called as Peierls-BTE (PBTE) or phonon BTE, which can be represented as:

$$\frac{\partial f}{\partial t} + \frac{\partial \mathbf{r}(\mathbf{q_l})}{\partial t} \cdot \nabla_r f = \left|\frac{\partial f}{\partial t}\right|_{C_{col}} \tag{5.1}$$

where $\mathbf{q_l}$ denotes the wavevector for lth mode of phonon. If we ignore the index of modes for clarity, the collision integral can be expressed as:

$$\left|\frac{\partial f}{\partial t}\right|_{\text{Col}} = \sum_{m=1}^{M} \sum_{\mathbf{q}^1,\ldots,\mathbf{q}^m \subset \text{BZ1}} \Gamma_d\left(\mathbf{q}, \mathbf{q}^1, \ldots, \mathbf{q}^m, f\right) - \\ \Gamma_r\left(\mathbf{q}, \mathbf{q}^1, \ldots, \mathbf{q}^m, f\right) \tag{5.2}$$

where M is the total phonon states, $\Gamma_d(\mathbf{q},\ldots)$ and $\Gamma_r(\mathbf{q},\ldots)$ are the decrease and increase in phonon population respectively in state $\mathbf{q}$. It is impossible to obtain a closed-form solution of the above equation without approximations on collision integral

or distribution function. In this study, the stochastic particle-based CMC presented in [173] is used for obtaining a solution of BTE.

In this study, 2 types of phonons scattering are considered: 1) Anharmonic Scattering and 2) Isotope Scattering. At room temperature, other anharmonic effects except three-phonon scattering are negligible [174], for this reason this study only considers three-phonon scattering and is ultimately referred to as anharmic scattering. Anharmonic scattering of phonons can be classified into two categories: 1) Recombination, where two phonons are combined into a third phonon conserving energy and momentum and 2) Decay, where a single phonon leads to the creation two new phonons. The decay rate at which phonon with wavevector $\mathbf{q}$ generates two phonons with wavevectors $\mathbf{q}'$ and $\mathbf{q}''$ can be expressed as

$$\Gamma_d(\mathbf{q}, \mathbf{q}'', \mathbf{q}'') = n(n'+1)(n''+1)A(\mathbf{q}, \mathbf{q}'', \mathbf{q}'') \tag{5.3}$$

where n, n' and n'' represents the number of phonons with wavevector $\mathbf{q}$, $\mathbf{q}''$ and $\mathbf{q}''$ respectively and A is a factor representing the strength of three-phonon process, which depends on the elastic properties but not on the phonon population. Similarly, if two phonons with wave vector $\mathbf{q}'$ and $\mathbf{q}''$ recombines into a phonon having wavevector of $\mathbf{q}$, the recombination rate will be

$$\Gamma_r(\mathbf{q}, \mathbf{q}'', \mathbf{q}'') = (n+1)n'n''A(\mathbf{q}, \mathbf{q}'', \mathbf{q}'') \tag{5.4}$$

The recombination and decay rates become equal during thermal equilibrium, where the phonon population in each state can be described with Bose-Einstein distribution as [174]

$$n_0(\mathbf{q}, T) = \left(exp\left(\frac{\hbar\omega}{k_B T} \right) - 1 \right)^{-1} \tag{5.5}$$

where T represents the absolute temperature, $\hbar$ stands for the reduced Plank constant, k_B is the Boltzmann constant anad ω represents the frequency of phonon. From

[175, 176], factor A can be computed as

$$A(\mathbf{q}, \mathbf{q}'', \mathbf{q}'') = \frac{2\hbar}{M^3} \frac{C^2(\mathbf{q}, \mathbf{q}'', \mathbf{q}'')}{\omega\omega'\omega''} \delta(\omega - \omega' - \omega'') \tag{5.6}$$

where M is for the average atomic mass and $C^2(\mathbf{q}, \mathbf{q}'', \mathbf{q}'')$ represents the Fourier component of anharmonic coupling. It is possible to obtain an expression for this term applicable for acoustic phonons using [176] as

$$C^2(\mathbf{q}, \mathbf{q}'', \mathbf{q}'') = (\frac{4\gamma^2}{3N_c})(\frac{M^2}{v^2})(\omega\omega'\omega'')^2\delta(\mathbf{q} - \mathbf{q}' - \mathbf{q}'') \tag{5.7}$$

where γ is the Grüneisen constant, N_c is the numner of unit cells in the lattice and v is the sound velocity.

Isotopes are atoms of the same element having same number of protons but different atomic mass due to different number of neutrons. Although isotopes doesn't alter crystal or electrical structure, the difference in mass can disturb the periodicity of the crystal to cause large phonon scattering. The increase in phonon scattering ultimately leads to a significant reduction in thermal conductivity tensor. Isotopic enrichment can lead to increase in thermal conductivity in semiconductors like 30% in germanium, 60% for Silicon and 50% for diamond [159], so this isotope scattering events should not be overlooked.

The matrix element necessary for isotope scattering can be found from [174, 177] as

$$\Xi_{iso}(\mathbf{q}, \mathbf{q}') = \frac{\hbar}{4\rho N_c}\sqrt{(\omega, \omega')}(\mathbf{e}_0 \cdot \mathbf{e}_1^*)M_{\mathbf{q},\mathbf{q}'}a(-\mathbf{q})a(\mathbf{q}'), \tag{5.8}$$

where ρ is the density of material, N_c is the number of elementary cells, $\mathbf{e}_0$ and $\mathbf{e}_1^*$ are the polarization vectors of $\mathbf{q}$ and $\mathbf{q}'$, $M_{\mathbf{q},\mathbf{q}'}$ can be defined as

$$M_{\mathbf{q},\mathbf{q}'} = \sum_j (\overline{M} - M_j)exp[i(\mathbf{q} - \mathbf{q}') \cdot \mathbf{r}_j] \tag{5.9}$$

where $\overline{M}$ is the average mass, $\mathbf{r}_j$ and M_j are the position and mass of the jth atom respectively. Fermi's Golden rule can be used to replace creation and annihilation

operators with matrix representations and thus obtaining the scattering rate as

$$\Gamma_{iso}(\mathbf{q}, \mathbf{q}') = \frac{\pi K_{iso}}{2N_a} \omega \omega'(\mathbf{e}_0 \cdot \mathbf{e}_1^*)^2 \delta(\omega - \omega')n(n' + 1) \tag{5.10}$$

where K_{iso} is a mass weighting factor defined as

$$K_{iso} = \sum_i f_i \left(\frac{M_i - \bar{M}}{\bar{M}}\right)^2, \tag{5.11}$$

where i represents the index of isotope species and f_i stands for the fraction of the solid made with that species.

In this study, using CMC, exactly same approach as discussed in [173] has been used to form the lookup table or scattering table, which contains all the final states after different scattering mechanisms along with probabilities. For dispersion relation, 14 parameter valence shell model is used and phonon modes are segmented by a non-uniform grid. For reducing the size of the lookup table, similar rejection algorithm as [173] is used, where the maximum scattering rate is compared to the actual value based on local conditions.

Different tensors of thermal conductivity are computed using [176] as:

$$\lambda_{k,j} = \sum_{q \subset BZ1} v_k v_j \tau(\mathbf{q}) \frac{(\hbar\omega)^2}{k_B T^2} n(\mathbf{q}, T)(n(\mathbf{q}, T) + 1) \tag{5.12}$$

where k and j are spatial directions, v, τ and ω are the velocity, average time between scattering events and frequency of phonon.

5.3   Simulation Results

In our simulations we have modeled bulk diamond without any defect/dislocations and only isotope of carbon ^{13}C which substitutes the native ^{12}C atoms at their respective lattice sites. We consider three different samples with isotope concentration of 0.001% (ultrapure), 0.1% and 1.07% (natural) respectively. The simulation temperature range is chosen between 40K to 500K. Below 40K, the dominant scattering

mechanism would be the device boundary (surface scattering) and beyond 500K the higher order coupling between phonon modes (four-phonon process) would dominate [178]. From power transistor/diode perspective the temperature region between 300K – 500K is of interest.

Firstly, we plot the steady state phonon energy distribution function (EDF) as shown in Figure 5.1 is obtained through solution of full band nonlinear Peierls-Boltzmann transport equation in Equation 5.1. The total phonon concentration at a given position in real-space of the bulk diamond is obtained by integrating the EDF with respect to phonon energy plotted on the x-axis of Figure 5.1. The phonon concentration shows an increase of 80 times for 0.1% ^{13}C and $\sim$ 300 times for 1.07% ^{13}C with respect to 0.001% ^{13}C which is $1.8 \times 10^{18} m^{-3}$ at 80K. The steady state phonon concentration so obtained above is independent of isotope concentration.

Secondly, we discuss the scattering rate regimes across temperatures and isotope concentrations. Figure 5.2 shows the total anharmonic scattering (inelastic) rates and the isotope scattering rates (elastic) plots for all the three isotope concentrations for 80K, 300K and 500K temperatures respectively. Clearly, the anharmonic rate is independent of isotope concentration at all temperatures. In the low temperature regime (80K) it is approximately $\sim 9.2 \times 10^9 sec^{-1}$ on the average. It is observed that, for ultrapure diamond, the anharmonic rate dominates the isotope rate ($\sim 7.4 \times 10^7 sec^{-1}$) but when isotope concentration increases to 0.1% then the isotope rate ($\sim 7.42 \times 10^9 sec^{-1}$) becomes comparable to the anharmonic rate. For the natural diamond (1.07%), the isotope rate ($\sim 7.9 \times 10^{10} sec^{-1}$) dominates the anharmonic rate. At room temperature, the anharmonic rate increases to $\sim 1.3 \times 10^{10} sec^{-1}$, whereas the isotope rate for the three samples (in the order of increasing concentration) are $\sim 7.5 \times 10^7 sec^{-1}$, $7.5 \times 10^9 sec^{-1}$ and $7.9 \times 10^{10} sec^{-1}$. Thus, only for the natural diamond we see that the isotope rate dominates the anharmonic rate ($\sim$ 6 times).
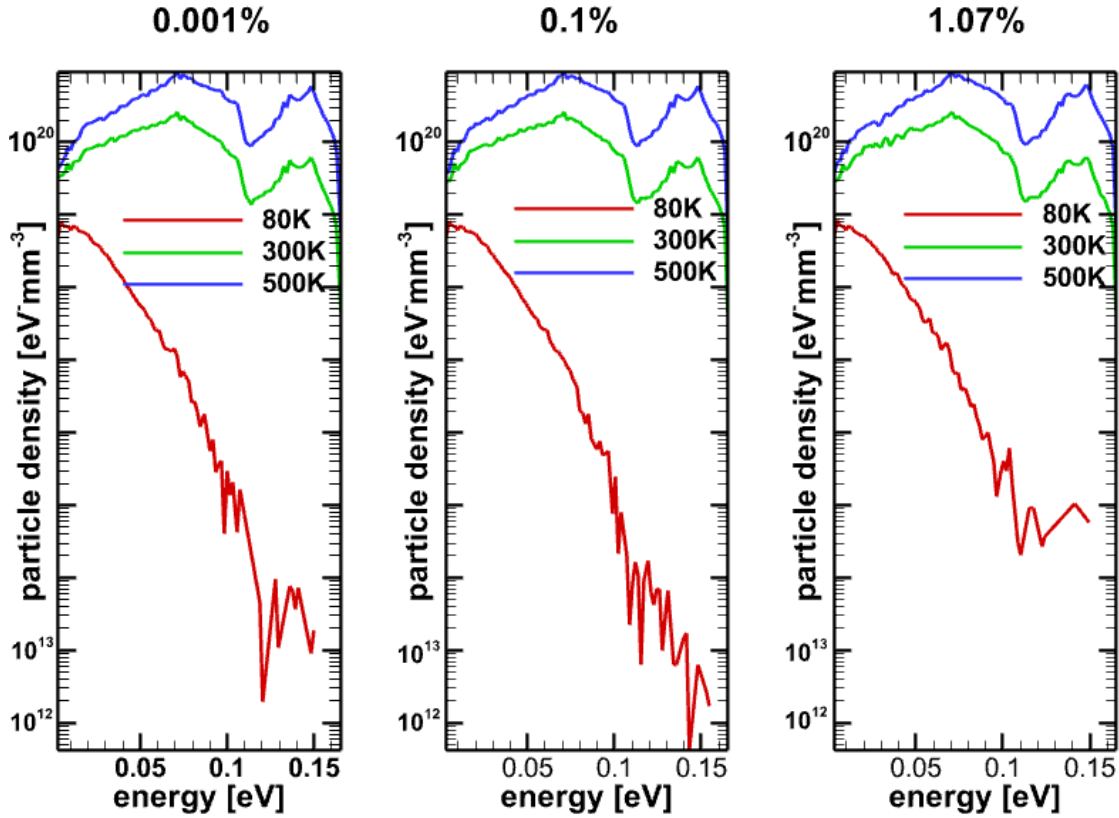
Figure 5.1: Energy distribution Function for acoustic phonon modes computed for 3 different temperatures for each isotope composition.

At 500K, the anharmonic rate approximately doubles its room temperature value to $\sim 2.3 \times 10^{10} sec^{-1}$, however the isotope rate shows a very small increment (in the order of increasing concentration) are $\sim 7.9 \times 10^7 sec^{-1}$, $7.9 \times 10^9 sec^{-1}$ and $8.5 \times 10^{10} sec^{-1}$. Hence, for 500K the anharmonic rate is the dominant scattering mechanism for ultrapure and 0.1% sample but for natural diamond the isotope rate dominates with nearly 4 times higher scattering rate.

Thirdly, Figure 5.3 shows the total thermal conductivity of bulk diamond for three ^{13}C concentration 0.001%, 0.1% and 1.07% respectively as a function of temperature obtained through CMC simulations. It is seen that the simulated values of thermal
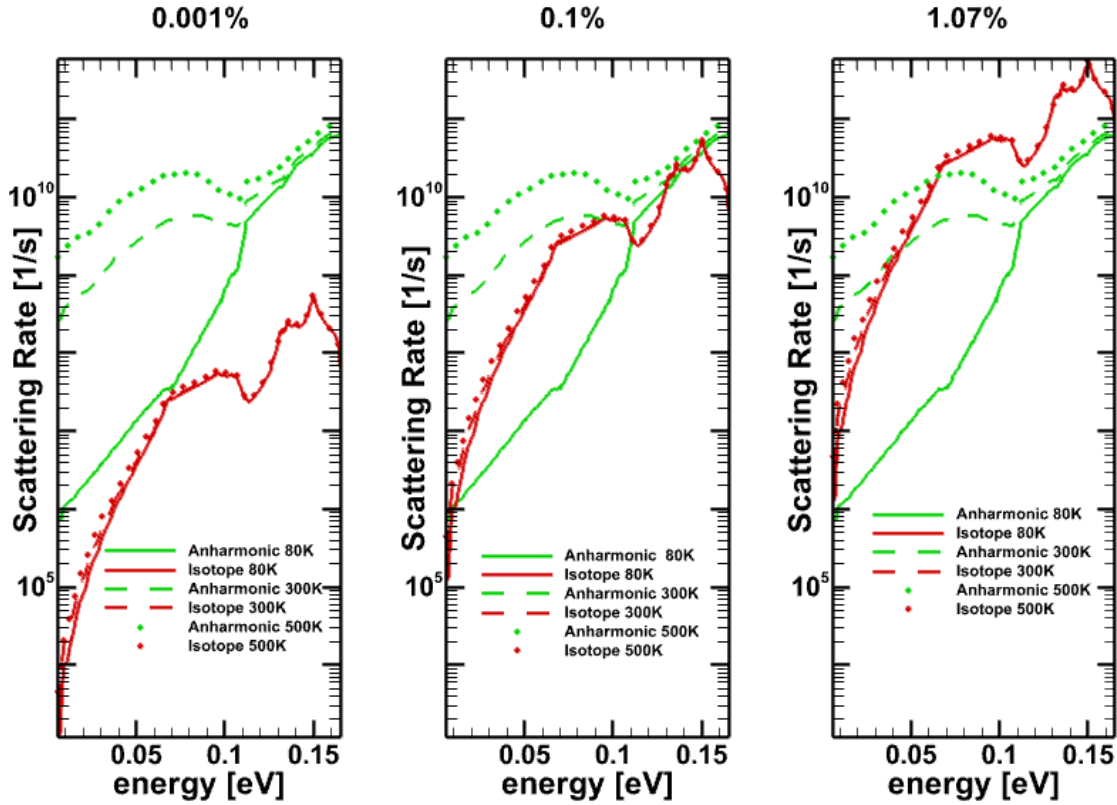
Figure 5.2: Anharmonic and Isotope scattering rate for 3 different temperature for each isotope concentration.

conductivity is in good agreement with the experimental values. The thermal conductivity for all three different isotope concentrations is maximized at a temperature of 80K. For 0.001% concentration, the maximum value is 231, 469 W/mK which drops by 74.2% to 59,681 W/mK for 0.1% concentration and drops by 94.2% to 13,410 W/mK for the natural diamond with 1.07% ^{13}C concentration. However, the deterioration in thermal conductivity at room temperature and 500K is much less. The thermal conductivity at room temperature for ultrapure diamond is 3890 W/mK which drops by 8.8% and 38.4% respectively for 0.1% and 1.07% isotope concentrations. At 500K the thermal conductivity of ultrapure sample is 2000 W/mK which drops by 6.8% and
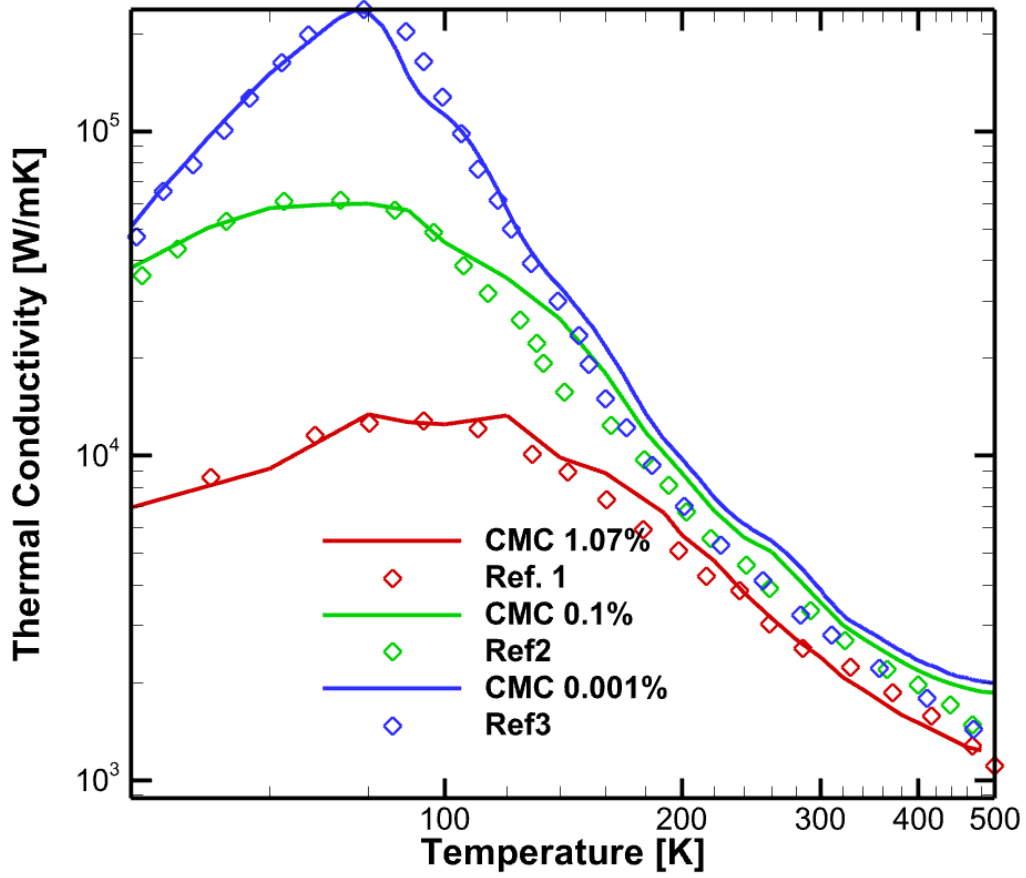
Figure 5.3: Thermal Conductivity of diamond with different isotope composition. Ref 1 refers to [7] and Ref 2 & Ref 3 indicates [8].

44.5% for 0.1% and 1.07% concentrations. To explain these observations, we compute overall mean free path (contributed from all the phonon branches and across all states in 1st Brillouin Zone) in Figure 5.4. At 80K, the overall mean free path for the 0.001% C13 is 45.1 µm which drops by 35.5% to 29.6 µm and 85.2% to 6.7 µm for 0.1% and 1.07% respectively.

To gain further insights we obtain the determinants of thermal properties for each phonon branch. As we know, diamond is a material with cubic crystal structure having,

Figure 5.4: Average mfp for different concentration

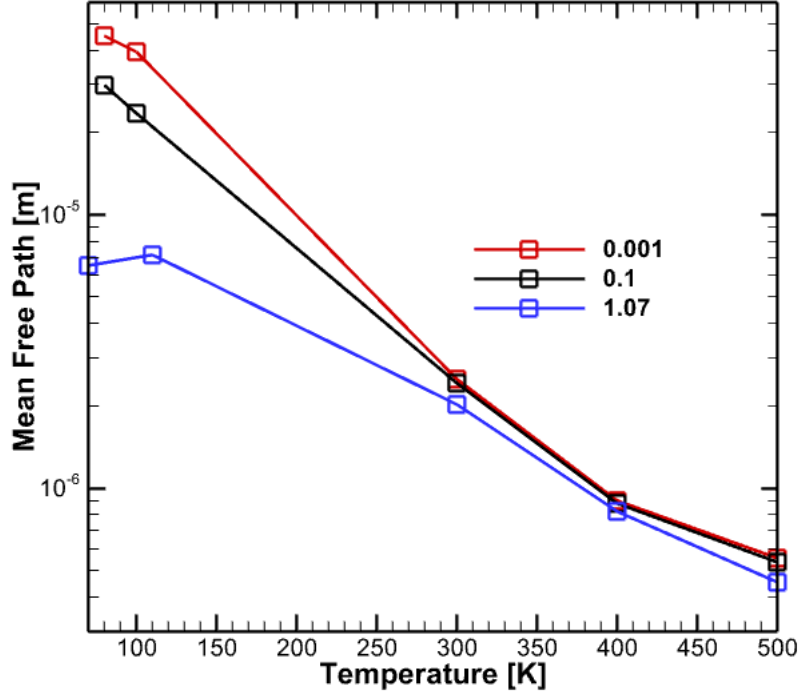basis atoms s = 2 in the unit cell, which gives 3s-3 = 3 optical branch and 3 acoustic phonon branches. The acoustic phonons along [100] directions have pure polarization and accordingly they are identified as Transverse acoustic (TA) and Longitudinal acoustic (LA) as shown in Figure 5.5. For optical phonons we adopt the convention that the branch with highest energy at the Γ will be called as Longitudinal optical (LO) phonon and the two degenerate low energy optical phonon will be termed as Transverse acoustic (as their plane of vibration being perpendicular to the direction of propagation of energy). Figure 5.6, Figure 5.7 and Figure 5.8 show the breakup of thermal conductivity attributed to each phonon branch at 300K and 500K respectively for all three diamond samples. As is expected, the contribution of optical phonons at room temperature is   1% whereas at 500K it is   5%. For the ultrapure diamond,
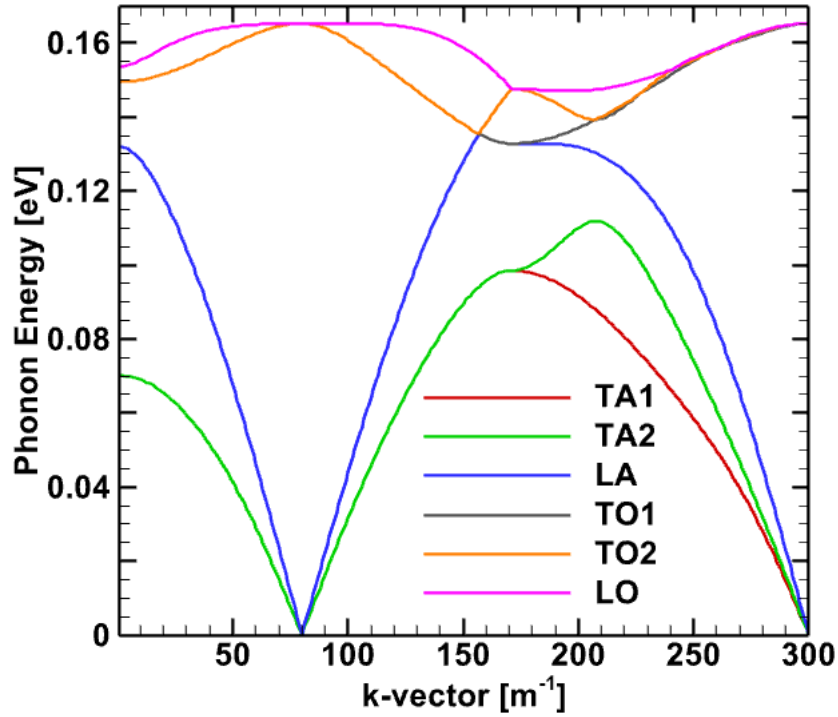
Figure 5.5: Phonon dispersion relation of diamond.

the TA2 branch contributes a maximum of 41% at room temperature which decreases to 37% at 500K whereas the LA branch's contribution is 26% and 35% at 300K and 500K respectively. The contribution of TA1 branch decreases from 32% to 23% as temperature increases from 300K to 500K. For the 0.1% and 1.07% isotope sample we see a similar trend of TA2 being the dominant contributor and TA1 contribution is the 2nd highest and LA is the 3rd highest contributor. The percentage contribution of the above acoustic branches are nearly same across all isotope concentrations and show a nearly same increase/decrease percentage as the temperature increases from 300K to 500K.
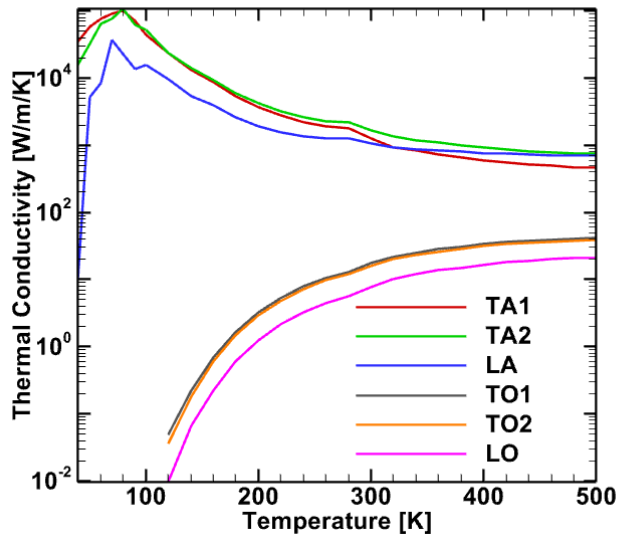
(a) Modewise thermal conductivity of 0.0001% isotope
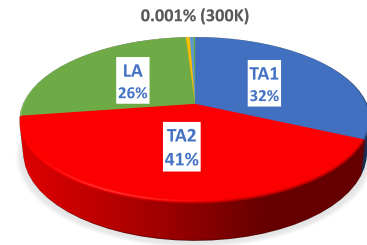
(b) Percentage contribution at
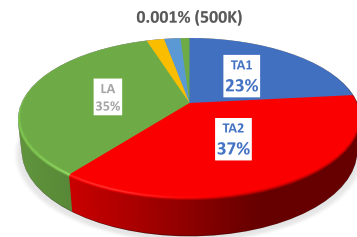
300K

(c) Percentage contribution at

500K

Figure 5.6: Modewise thermal conductivity and percentage contribution at 300K and

500K for 0.001% isotope concentration.
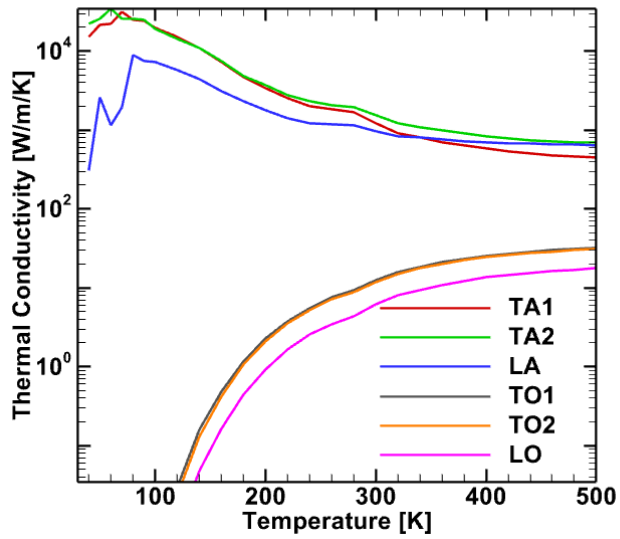
(a) Modewise thermal conductivity of 0.1% isotope
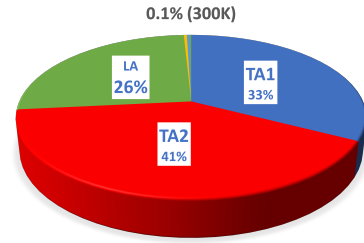
(b) Percentage contribution at 300K
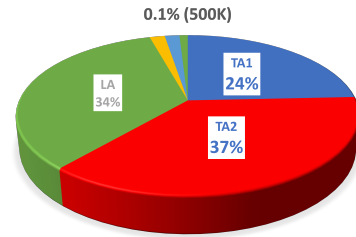
(c) Percentage contribution at 500K

Figure 5.7: Modewise thermal conductivity and percentage contribution at 300K and 500K for 0.1% isotope concentration.

(a) Modewise thermal conductivity of 1.07% isotope

(b) Percentage contribution at 300K

(c) Percentage contribution at 500K
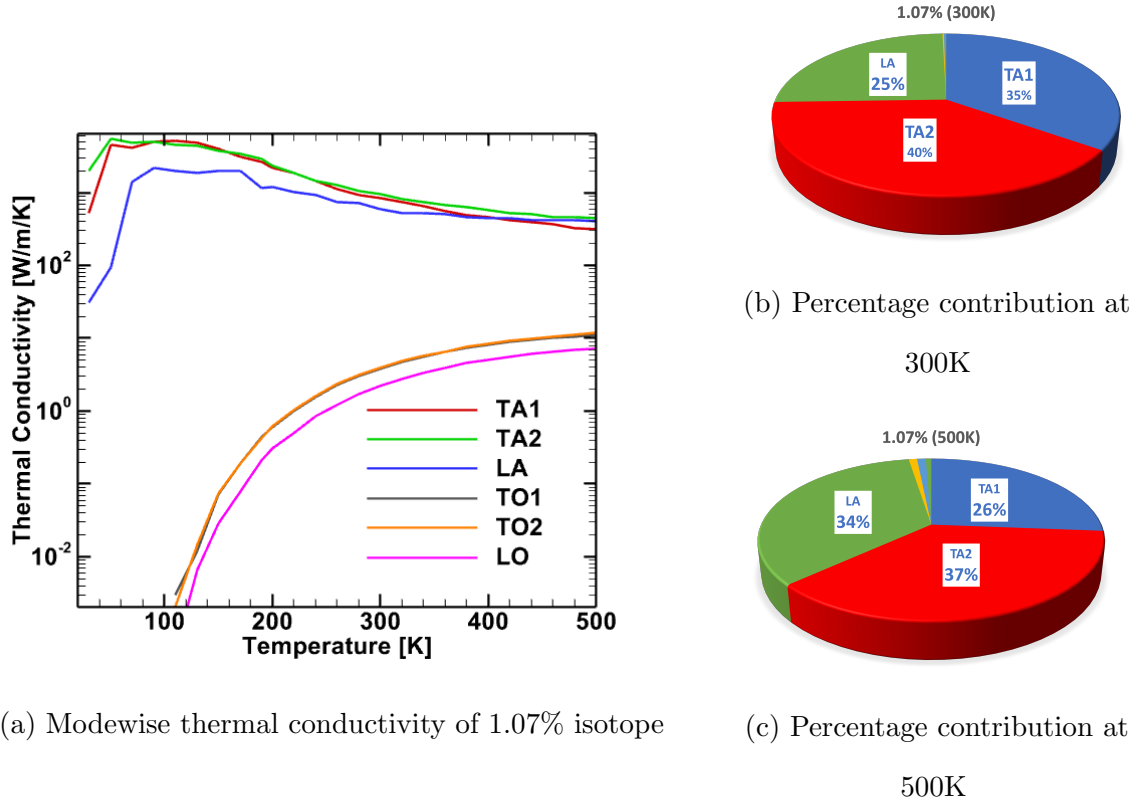
Figure 5.8: Modewise thermal conductivity and percentage contribution at 300K and 500K for 1.07% isotope concentration.

As mentioned above that the acoustic branches are the principal contributors to thermal conductivity, we plot the EDF attributed to these branches for all isotope concentrations by solving the Peierls-Boltzmann equation of Equation 5.1 as shown in Figure 5.9 for temperatures 300K and 500K respectively. Also, Figure 5.10, Figure 5.11 and Figure 5.12 show the mean free path attributed to these branches at 300K and 500K respectively. We first discuss the plots at 300K temperature, for the ultrapure diamond, the steady state phonon concentration for the two transverse phonons TA1 and TA2 are nearly same $\sim 6.5 \times 10^{18} m^{-3}$ and $5.5 \times 10^{18} m^{-3}$ respectively whereas for the LA phonon it is about ~ 4 times lower at $1.5 \times 10^{18} m^{-3}$. The average mean free path for TA1, TA2 and LA are 3.5µm, 3.6µm and 4µm respectively. The steady

99

state phonon concentration are independent of isotope concentrations. The average mean free path for TA1, TA2 and LA phonon for 0.1% isotope concentration are 3.5μm, 3.5μm and 3.7μm respectively, these values decrease to 2.8μm, 2.7μm and 3.1μm respectively for the natural diamond (1.07%). We observe that, for a given phonon branch the average mean free path remain nearly same for all three isotope concentrations hence their contribution to total thermal conductivity does not change with isotope concentration. Further, the 4 times lower phonon concentration for the LA branch with respect to the transverse branch explains its relatively lower contribution to the overall thermal conductivity compared to its transverse counterparts. Also, the TA1 and TA2 branches have nearly same phonon concentration and average mean free path, but the $7 - 8$ % higher contribution of TA2 branch with respect to TA1 branch can be attributed to higher average group velocity of TA2 with respect to TA1 branch.
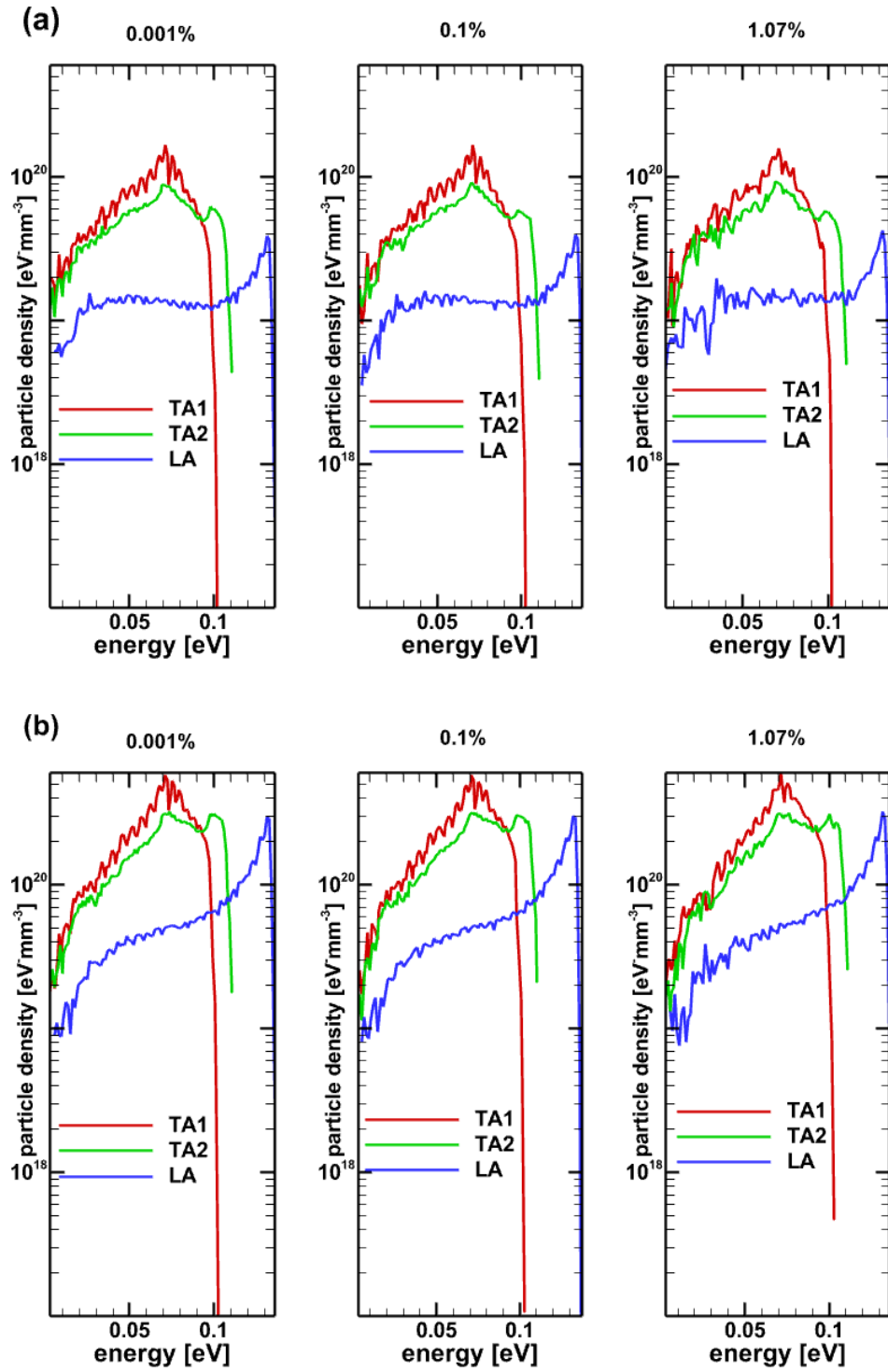
Figure 5.9: Particle density for different isotopic composition at (a) 300K and (b) 500K.
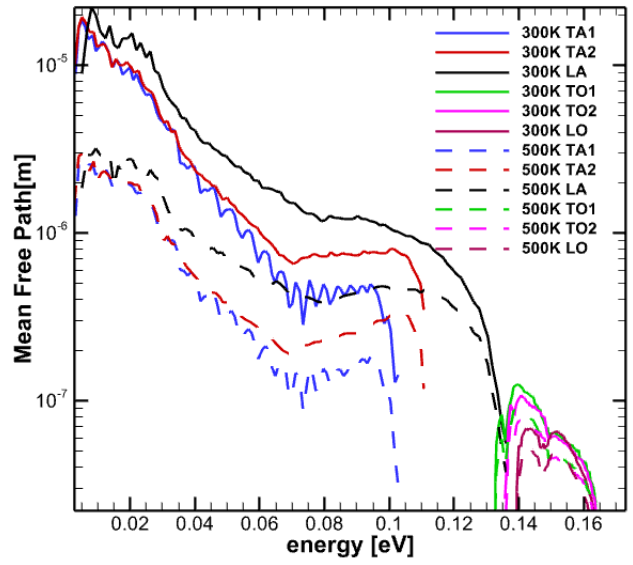
Figure 5.10: MFP for 0.001% isotope
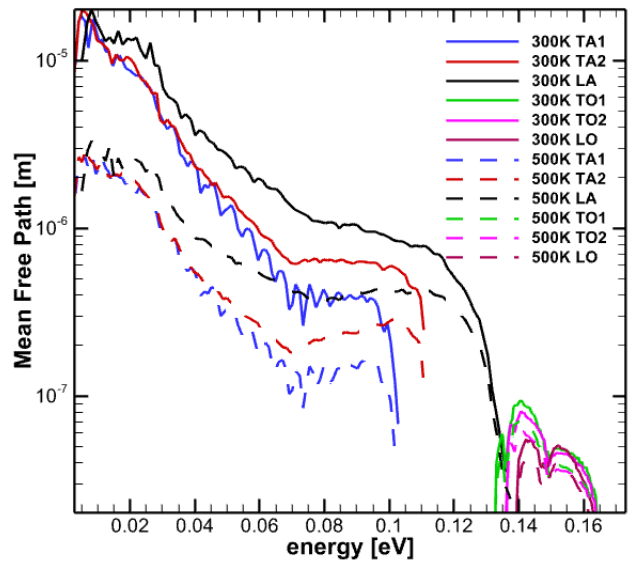
Figure 5.11: MFP for 0.1% isotope
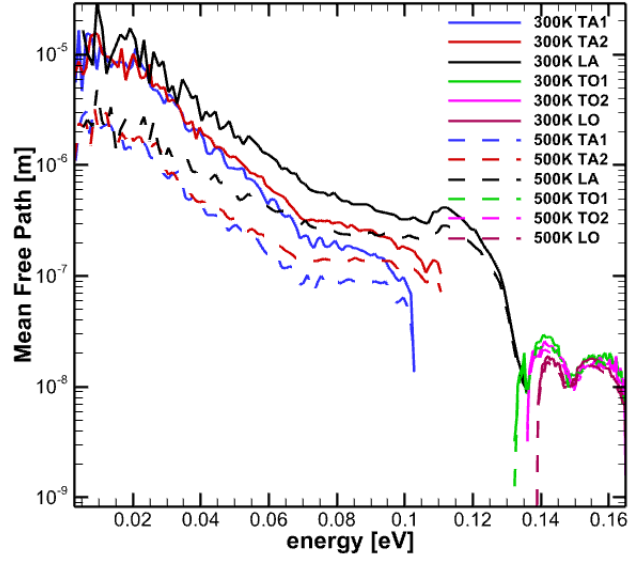
Figure 5.12: MFP for 1.07% isotope

Now, we discuss the thermal properties of the three acoustic branches at 500K. The phonon concentration for TA1, TA2 and LA phonons are $\sim 2.1 \times 10^{19} m^{-3}$, $1.7 \times 10^{19} m^{-3}$ and $6.2 \times 10^{18} m^{-3}$ respectively. Whereas, the average mean free path for the ultrapure diamond for the above three phonons are 0.72µm, 0.74µm and 0.87µm respectively. For the 0.1% sample the values of average mean free path are 0.70µm, 0.73µm and 0.85µm respectively. Finally, for the natural diamond the average mean free path for the three modes are 0.62µm, 0.63µm and 0.71µm respectively. We see that the decrease in mean free path with increase in isotope concentration (from ultrapure to natural) is around 100nm. The phonon concentration for the LA phonon is ~ 3 times lower than the transverse branches whereas it was 4 times lower than the transverse branches at 300K. But, compared to the phonon concentration at 300K we see an increase by 4 times at 500K, this explains its increase in contribution from $\sim 26\%$ at 300K to $\sim 35\%$ at 500K. We again see that, the phonon concentration and

103

mean free path for both the transverse branch are nearly same at 500K, the TA2 branch's 7 – 8% higher contribution to the total thermal conductivity is attributed to its higher average group velocity.

Although optical phonon's contribution to the total thermal conductivity for bulk diamond is insignificant but we give the values of their average mean free path for ultra pure diamond for the Transverse optical TO1, TO2 and Longitudinal optical LO branch for 300K (500K) as follows: 36.7 nm (49.7nm), 32.5nm (43.8nm) and 24.2nm (32.3nm). For 0.1% these values are 29.4nm(37.6nm), 26.2nm(33.3nm) and 20.1 nm(25.5nm). Similarly, for the natural diamond it is 10.9nm(12.5nm), 9.8nm(11.1nm) and 8nm(9nm) respectively. The extremely low mean free path for the optical phonons are responsible for their low thermal conductivity.

5.4   Conclusion

Bulk diamond simulations were performed for three isotope concentrations namely 0.001%, 0.1% and 1.07% for determination of thermal properties. We investigate the thermal properties attributed to each phonon branch in order to understand the deterioration in thermal conductivity due to increase in isotope concentrations. The isotope scattering dominates the anharmonic rate for room temperature and above and is the principal determinant of the thermal conductivity. The increase in isotope concentration from 0.001% to 0.1% does not causes as much deterioration as we observe for an increase from 0.1 to 1.07% concentration, and it is the dominant isotope effect which is responsible for it.

Chapter 6

SUMMARY & FUTURE WORK

6.1   Summary

Thermal management of continuously shrinking semiconductor devices has been a major challenge of device development. Modeling heat transport can be viewed as a vital tool for the device design. The exact solution of BTE can lead to proper representation of transport inside the device, but integro-differential nature of the equation leads to significant challenge to solve it. Some methods consider only the lower order moments of the BTE, but they are applicable for near-equilibrium scenarios. As the simulation conditions deviate from equilibrium in modern devices almost all the times, MC methods take care of this issue by providing a solution in exact statistical terms. As an extension of the traditional EMC method, the CMC pretabulates all the scattering rates along with the final states, which significantly reduces the simulation time. CMC also allows the use of rejection algorithm for the adaptation to local runtime condition.

The solution of Poisson's equation is an integral part of any device simulation tool. In this work, the multi-purpose library PETSc has been employed as a Poisson solver into CMC. PETSc provide a variety of options for preconditioner and solvers to solve any linear PDE. For this particular study, Incomplete LU method is chosen as preconditioner and KSP method is adopted as solver. Under this circumstance, results are obtained for 2D and 3D geometries and compared with existing multigrid method available in CMC. The results found from PETSc closely follows the multigrid method.

EBE is derived using the first two moments of the BTE for phonons and can effectively be used for modeling heat transport when it is properly coupled with the electron dynamics. For solving EBE, CMC adopts the Kirchoff transformation, which effectively converts EBE into an elliptical PDE and later solved with the Poisson solver. This study uses this existing thermal solver to compare the effects of different gate architectures and the performance of GaN HEMT. Acoustic and optical temperature profiles of each studied device geometries have been obtained. Later, DC and RF performances were studied along with proper validation with experimental results. Furthermore, hot electron minimization was obtained through the simulated energy distribution function.

Thermal properties of diamond have been evaluated through CMC. Diamond is being effectively used for the solution of thermal management problem due to its high thermal conductivity. The effects of isotope on the deterioration of thermal conductivity of diamond have been studied. Thermal properties of each phonon mode have been obtained to identify the dominant mode contributing to the thermal conductivity. At room temperature anharmonic scattering is dominant over isotope scattering, but at elevated temperature isotope scattering becomes significantly higher than anharmonic scattering.

6.2 Future Work

The planned future work related to this study can be summerized as follows:

1) Implementation of a parallel PETSc solver to make the existing solver faster.

2) Inclusion of substrate and associated thermal interface in the study of GaN HEMT to evaluate the effect on the temperature maps. Currently the drain current is overestimated along with the cutoff frequency due to the absence of the substrate. With an inclusion of substrate the mismatch should be taken care of.

106

3) Effect of n-type and p-type impurity in thermal conductivity of diamond.

4) Similar to diamond, thermal conductivity analysis of cubic boron nitride.

REFERENCES

[1] D. Vasileska, S. M. Goodnick, and G. Klimeck, *Computational Electronics: semiclassical and quantum device modeling and simulation.* CRC press, 2017.

[2] K. E. Merrill, "Nonlinear Electrothermal Monte Carlo Device Simulation," Ph.D. dissertation, Arizona State University, 2020.

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

[4] U. K. Mishra, L. Shen, T. E. Kazior, and Y.-F. Wu, "GaN-based RF power devices and amplifiers," *Proceedings of the IEEE*, vol. 96, no. 2, pp. 287–305, 2008.

[5] J. Acharjee, R. Singh, K. Merill, S. Goodnick, and M. Saraniti, "Assessment of T-gate and π-gate HEMT through Cellular Monte Carlo simulations," *IEEE Transactions on Electron Devices*, 2023.

[6] P. Altuntas, F. Lecourt, A. Cutivet, N. Defrance, E. Okada, M. Lesecq, S. Rennesson, A. Agboton, Y. Cordier, V. Hoel, and J. De Jaeger, "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, 2015.

[7] R. Berman, P. Hudson, and M. Martinez, "Nitrogen in diamond: evidence from thermal conductivity," *Journal of Physics C: Solid State Physics*, vol. 8, no. 21, p. L430, 1975.

[8] L. Wei, P. Kuo, R. Thomas, T. Anthony, and W. Banholzer, "Thermal conductivity of isotopically modified single crystal diamond," *Physical review letters*, vol. 70, no. 24, p. 3764, 1993.

[9] S. M. Sze, Y. Li, and K. K. Ng, *Physics of semiconductor devices.* New York: Wiley, 2021.

[10] C. Jacoboni and P. Lugli, *The Monte Carlo method for semiconductor device simulation.* Springer Science & Business Media, 2012.

[11] K. Fushinobu, A. Majumdar, and K. Hijikata, "Heat generation and transport in submicron semiconductor devices," 1995.

[12] 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, 2008.

[13] N. D. Mermin, "Lindhard dielectric function in the relaxation-time approximation," *Physical Review B*, vol. 1, no. 5, p. 2362, 1970.

[14] T. Sadi, R. Kelsall, and N. Pilgrim, "Simulation of electron transport in in-gaas/algaas hemts using an electrothermal monte carlo method," *IEEE transactions on electron devices*, vol. 53, no. 8, pp. 1768–1774, 2006.

[15] V. Efimov and L. Mezhov-Deglin, "Phonon scattering in diamond films," *Physica B: Condensed Matter*, vol. 263, pp. 745–748, 1999.

[16] D. Ferry and S. M. Goodnick, *Transport in nanostructures*. Cambridge university press, 1999, no. 6.

[17] M. Lundstrom, "Fundamentals of carrier transport, 2nd edn," *Measurement Science and Technology*, vol. 13, no. 2, pp. 230–230, 2002.

[18] S. Selberherr, *Analysis and simulation of semiconductor devices*. New York: Springer, 2012.

[19] D. K. Ferry, *Semiconductors*. Bristol: IoP Publishing, 2013.

[20] C. Cercignani, "On the Boltzmann equation for rigid spheres," *Transport Theory and Statistical Physics*, vol. 2, no. 3, pp. 211–225, 1972.

[21] R. E. Peierls, *Quantum theory of solids*. Oxford University Press, 1955.

[22] N. W. Ashcroft and N. D. Mermin, *Solid state physics*. Cengage Learning, 2022.

[23] 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, pp. 251–274, 2003.

[24] S. Sinha, E. Pop, R. Dutton, and K. Goodson, "Non-equilibrium phonon distributions in sub-100 nm silicon transistors," *Journal of Heat Transfer*, vol. 128, no. 7, pp. 638–647, 12 2006.

[25] A. Joshi and A. Majumdar, "Transient ballistic and diffusive phonon heat transport in thin films," *Journal of Applied Physics*, vol. 74, no. 1, pp. 31–39, 1993.

[26] J. C. Chai, H. S. Lee, and S. V. Patankar, "Finite volume method for radiation heat transfer," *Journal of Thermophysics and Heat Transfer*, vol. 8, no. 3, pp. 419–425, 1994.

[27] G. Chen, "Ballistic-diffusive equations for transient heat conduction from nano to macroscales," *Journal of Heat Transfer*, vol. 124, no. 2, pp. 320–328, 2002.

[28] K. Kukita and Y. Kamakura, "Monte Carlo simulation of phonon transport in silicon including a realistic dispersion relation," *Journal of Applied Physics*, vol. 114, no. 15, 2013.

[29] S. Mei, L. Maurer, Z. Aksamija, and I. Knezevic, "Full-dispersion Monte Carlo simulation of phonon transport in micron-sized graphene nanoribbons," *Journal of Applied Physics*, vol. 116, no. 16, 2014.

[30] W. Van Roosbroeck, "Theory of the flow of electrons and holes in germanium and other semiconductors," *The Bell System Technical Journal*, vol. 29, no. 4, pp. 560–607, 1950.

[31] F. Beichelt, *Applied probability and stochastic processes.* CRC Press, 2018.

[32] D. L. Scharfetter and H. K. Gummel, "Large-signal analysis of a silicon read diode oscillator," *IEEE Transactions on electron devices*, vol. 16, no. 1, pp. 64–77, 1969.

[33] Q. Zhang, Q. Wang, L. Zhang, and B. Lu, "A class of finite element methods with averaging techniques for solving the three-dimensional drift-diffusion model in semiconductor device simulations," *Journal of Computational Physics*, vol. 458, p. 111086, 2022.

[34] R. E. Bank, W. Coughran, Jr, and L. C. Cowsar, "The finite volume Scharfetter-Gummel method for steady convection diffusion equations," *Computing and Visualization in Science*, vol. 1, no. 3, pp. 123–136, 1998.

[35] P. Bochev, K. Peterson, and X. Gao, "A new control volume finite element method for the stable and accurate solution of the drift–diffusion equations on general unstructured grids," *Computer Methods in Applied Mechanics and Engineering*, vol. 254, pp. 126–145, 2013.

[36] F. Brezzi, L. D. Marini, and P. Pietra, "Numerical simulation of semiconductor devices," *Computer methods in applied mechanics and engineering*, vol. 75, no. 1-3, pp. 493–514, 1989.

[37] F. Brezzi, L. Marini, S. Micheletti, P. Pietra, R. Sacco, and S. Wang, "Discretization of semiconductor device problems (i)," *Handbook of numerical analysis*, vol. 13, pp. 317–441, 2005.

[38] C. Johnson, U. Navert, and J. Pitkaranta, "Finite element methods for linear hyperbolic problems," *Computer methods in applied mechanics and engineering*, vol. 45, pp. 285–312, 1984.

[39] G. Carey and M. Sharma, "Semiconductor device modeling using flux upwind finite elements," *COMPEL-The international journal for computation and mathematics in electrical and electronic engineering*, vol. 8, no. 4, pp. 219–234, 1989.

[40] R. Stratton, "Diffusion of hot and cold electrons in semiconductor barriers," *Physical Review*, vol. 126, no. 6, p. 2002, 1962.

[41] J. G. Ruch, "Electron dynamics in short channel field-effect transistors," *IEEE Transactions on Electron Devices*, vol. 19, no. 5, pp. 652–654, 1972.

[42] S. Chou, D. A. Antoniadis, and H. I. Smith, "Observation of electron velocity overshoot in sub-100-nm-channel MOSFET's in silicon," *IEEE Electron Device Letters*, vol. 6, no. 12, pp. 665–667, 1985.

[43] G. Shahidi, D. A. Antoniadis, and H. I. Smith, "Electron velocity overshoot at room and liquid nitrogen temperatures in silicon inversion layers," *IEEE Electron Device Letters*, vol. 9, no. 2, pp. 94–96, 1988.

[44] P. Degond and P. A. Markowich, "On a one-dimensional steady-state hydrodynamic model for semiconductors," *Applied Mathematics Letters*, vol. 3, no. 3, pp. 25–29, 1990.

[45] R. Thoma, A. Emunds, B. Meinerzhagen, H.-J. Peifer, and W. L. Engl, "Hydrodynamic equations for semiconductors with nonparabolic band structure," *IEEE Transactions on Electron Devices*, vol. 38, no. 6, pp. 1343–1353, 1991.

[46] A. Anile and O. Muscato, "Improved hydrodynamical model for carrier transport in semiconductors," *Physical Review B*, vol. 51, no. 23, p. 16728, 1995.

[47] M. A. Kastner, "The single-electron transistor," *Reviews of modern physics*, vol. 64, no. 3, p. 849, 1992.

[48] M. P. Allen *et al.*, "Introduction to molecular dynamics simulation," *Computational soft matter: from synthetic polymers to proteins*, vol. 23, no. 1, pp. 1–28, 2004.

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

[50] K. Hess, *Monte Carlo device simulation: full band and beyond.* Springer Science & Business Media, 2012, vol. 144.

[51] R. W. Hockney and J. W. Eastwood, *Computer simulation using particles.* crc Press, 2021.

[52] S. Zollner, S. Gopalan, and M. Cardona, "Intervalley deformation potentials and scattering rates in zinc blende semiconductors," *Applied Physics Letters*, vol. 54, no. 7, pp. 614–616, 1989.

[53] B. K. Ridley, *Quantum processes in semiconductors.* Oxford University Press, USA, 2013.

[54] J. D. Zook, "Piezoelectric scattering in semiconductors," *Physical Review*, vol. 136, no. 3A, p. A869, 1964.

[55] B. Lee, K. Kim, M. Dutta, and M. Stroscio, "Electron–optical-phonon scattering in wurtzite crystals," *Physical Review B*, vol. 56, no. 3, p. 997, 1997.

[56] R. Loudon, *The quantum theory of light.* OUP Oxford, 2000.

[57] M. V. Fischetti, N. Sano, S. Laux, and K. Natori, "Full-band-structure theory of high-field transport and impact ionization of electrons and holes in Ge, Si, and GaAs," *Journal of Technology Computer Aided Design TCAD*, pp. 1–50, 1996.

[58] M. Saraniti and S. M. 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.

[59] D. K. Ferry, *Semiconductors: Bonds and Bands.* IoP Publishing, 2013.

[60] G. Dahlquist and Å. Björck, *Numerical methods in scientific computing, volume i.* SIAM, 2008.

[61] C. M. Snowden, *Introduction to semiconductor device modelling.* World Scientific, 1998.

[62] R. Hockney, "Measurements of collision and heating times in a two-dimensional thermal computer plasma," *Journal of Computational Physics*, vol. 8, no. 1, pp. 19–44, 1971.

[63] 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, pp. 141–150, 1996.

[64] E. K. de Rivas, "On the use of nonuniform grids in finite-difference equations," *Journal of Computational Physics*, vol. 10, no. 2, pp. 202–210, 1972.

[65] J. D. Jackson, "Classical electrodynamics," 1999.

[66] S. Balay, S. Abhyankar, M. F. Adams, S. Benson, J. Brown, P. Brune, K. Buschelman, E. M. Constantinescu, L. Dalcin, A. Dener, V. Eijkhout, J. Faibussowitsch, W. D. Gropp, V. Hapla, T. Isaac, P. Jolivet, D. Karpeev, D. Kaushik, M. G. Knepley, F. Kong, S. Kruger, D. A. May, L. C. McInnes, R. T. Mills, L. Mitchell, T. Munson, J. E. Roman, K. Rupp, P. Sanan, J. Sarich, B. F. Smith, S. Zampini, H. Zhang, H. Zhang, and J. Zhang, "PETSc Web page," https://petsc.org/, 2023. [Online]. Available: https://petsc.org/

[67] L. N. Trefethen and D. Bau, *Numerical linear algebra.* Siam, 2022, vol. 181.

[68] M. Ferronato, "Preconditioning for sparse linear systems at the dawn of the 21st century: History, current developments, and future perspectives," *International Scholarly Research Notices*, vol. 2012, 2012.

[69] D. Young, "The numerical solution of elliptic and parabolic partial differential equations," in *Modern Mathematics for the Engineer.* McGraw-Hill Book Company, Inc., 1961, pp. 373–419.

[70] N. Li, Y. Saad, and E. Chow, "Crout versions of ILU for general sparse matrices," *SIAM Journal on Scientific Computing*, vol. 25, no. 2, pp. 716–728, 2003.

[71] Y. Saad, "Multilevel ILU with reorderings for diagonal dominance," *SIAM Journal on Scientific Computing*, vol. 27, no. 3, pp. 1032–1057, 2005.

[72] N. Islam, M. F. P. Mohamed, M. F. A. J. Khan, S. Falina, H. Kawarada, and M. Syamsul, "Reliability, applications and challenges of GaN HEMT technology for modern power devices: A review," *Crystals*, vol. 12, no. 11, p. 1581, 2022.

[73] J. L. Hudgins, G. S. Simin, E. Santi, and M. A. Khan, "An assessment of wide bandgap semiconductors for power devices," *IEEE Transactions on power electronics*, vol. 18, no. 3, pp. 907–914, 2003.

[74] R. Gaska, J. Yang, A. Osinsky, M. A. Khan, and M. S. Shur, "Novel high power AlGaN/GaN HFETs on SiC substrates," in *International Electron Devices Meeting. IEDM Technical Digest.* IEEE, 1997, pp. 565–568.

[75] A. Chini, D. Buttari, R. Coffie, S. Heikman, S. Keller, and U. Mishra, "12W/mm power density AlGaN/GaN HEMTs on sapphire substrate," *Electronics Letters*, vol. 40, no. 1, p. 1, 2004.

[76] J. Johnson, E. Piner, A. Vescan, R. Therrien, P. Rajagopal, J. Roberts, J. Brown, S. Singhal, and K. Linthicum, "12 W/mm AlGaN-GaN HFETs on silicon substrates," *IEEE Electron Device Letters*, vol. 25, no. 7, pp. 459–461, 2004.

[77] L. Shen, S. Heikman, B. Moran, R. Coffie, N.-Q. Zhang, D. Buttari, I. Smorchkova, S. Keller, S. DenBaars, and U. Mishra, "AlGaN/AlN/GaN high-power microwave HEMT," *IEEE Electron Device Letters*, vol. 22, no. 10, pp. 457–459, 2001.

[78] M. Micovic, P. Hashimoto, M. Hu, I. Milosavljevic, J. Duvall, P. J. Willadsen, W.-S. Wong, A. M. Conway, A. Kurdoghlian, P. W. Deelman *et al.*, "GaN double heterojunction field effect transistor for microwave and millimeterwave power applications," in *IEDM Technical Digest. IEEE International Electron Devices Meeting, 2004.* IEEE, 2004, pp. 807–810.

[79] J. Zúñiga-Pérez, V. Consonni, L. Lymperakis, X. Kong, A. Trampert, S. Fernández-Garrido, O. Brandt, H. Renevier, S. Keller, K. Hestroffer *et al.*, "Polarity in GaN and ZnO: Theory, measurement, growth, and devices," *Applied Physics Reviews*, vol. 3, no. 4, 2016.

[80] M. Meneghini, G. Meneghesso, and E. Zanoni, "Power GaN Devices," *Cham: Springer International Publishing*, 2017.

[81] O. Ambacher, J. Majewski, C. Miskys, A. Link, M. Hermann, M. Eickhoff, M. Stutzmann, F. Bernardini, V. Fiorentini, V. Tilak, B. Schaff, and L. F. Eastman, "Pyroelectric properties of Al(In)GaN/GaN hetero- and quantum well structures," *Journal of Physics: Condensed Matter*, vol. 14, no. 13, p. 3399, March 2002.

[82] M. Asif Khan, A. Bhattarai, J. Kuznia, and D. Olson, "High electron mobility transistor based on a GaN-Al x Ga1- x N heterojunction," *Applied Physics Letters*, vol. 63, no. 9, pp. 1214–1215, 1993.

[83] X.-G. He, D.-G. Zhao, and D.-S. Jiang, "Formation of two-dimensional electron gas at AlGaN/GaN heterostructure and the derivation of its sheet density expression," *Chinese Physics B*, vol. 24, no. 6, p. 067301, 2015.

[84] G. Meneghesso, M. Meneghini, and E. Zanoni, "Breakdown mechanisms in AlGaN/GaN HEMTs: an overview," *Japanese Journal of Applied Physics*, vol. 53, no. 10, p. 100211, 2014.

[85] Y. Puzyrev, B. Tuttle, R. Schrimpf, D. Fleetwood, and S. Pantelides, "Theory of hot-carrier-induced phenomena in GaN high-electron-mobility transistors," *Applied Physics Letters*, vol. 96, no. 5, 2010.

[86] M. Meneghini, G. Meneghesso, and E. Zanoni, "Analysis of the reliability of AlGaN/GaN HEMTs submitted to on-state stress based on electroluminescence investigation," *IEEE Transactions on Device and Materials Reliability*, vol. 13, no. 2, pp. 357–361, 2013.

[87] J. Joh, J. A. Del Alamo, and J. Jimenez, "A simple current collapse measurement technique for GaN high-electron mobility transistors," *IEEE Electron Device Letters*, vol. 29, no. 7, pp. 665–667, 2008.

[88] F. Gao, B. Lu, L. Li, S. Kaun, J. S. Speck, C. Thompson, T. Palacios *et al.*, "Role of oxygen in the OFF-state degradation of AlGaN/GaN high electron mobility transistors," *Applied Physics Letters*, vol. 99, no. 22, 2011.

[89] M. Bouya, D. Carisetti, N. Malbert, N. Labat, P. Perdu, J.-C. Clément, M. Bonnet, and G. Pataut, "Study of passivation defects by electroluminescence in AlGaN/GaN HEMTS on SiC," *Microelectronics Reliability*, vol. 47, no. 9-11, pp. 1630–1633, 2007.

[90] M. Meneghini, A. Stocco, M. Bertin, D. Marcon, A. Chini, G. Meneghesso, and E. Zanoni, "Time-dependent degradation of AlGaN/GaN high electron mobility transistors under reverse bias," *Applied Physics Letters*, vol. 100, no. 3, 2012.

[91] A. Amerasekera, M.-C. Chang, J. A. Seitchik, A. Chatterjee, K. Mayaram, and J.-H. Chern, "Self-heating effects in basic semiconductor structures," *IEEE Transactions on Electron Devices*, vol. 40, no. 10, pp. 1836–1844, 1993.

[92] U. Lindefelt, "Heat generation in semiconductor devices," *Journal of Applied Physics*, vol. 75, no. 2, pp. 942–957, 1994.

[93] A. Nigam, T. N. Bhat, S. Rajamani, S. B. Dolmanan, S. Tripathy, and M. Kumar, "Effect of self-heating on electrical characteristics of AlGaN/GaN HEMT on Si (111) substrate," *AIP Advances*, vol. 7, no. 8, p. 085015, 2017.

[94] S. Liu, J. Wang, Y. Lu, D. Huang, C. Huang, W. Hsieh, J. Lee, Y. Tsai, J. Shih, Y.-H. Lee *et al.*, "Self-heating effect in FinFETs and its impact on devices reliability characterization," in *2014 IEEE International Reliability Physics Symposium*. IEEE, 2014, pp. 4A–4.

[95] M. Sabry, W. Fikry, K. Salam, M. Awad, and A. Nasser, "A lumped transient thermal model for self-heating in MOSFETs," *Microelectronics Journal*, vol. 32, no. 10-11, pp. 847–853, 2001.

[96] I. Hwang, J. Kim, S. Chong, H. Choi, S. Hwang, J. Oh, J. K. Shin, and U. Chung, "Impact of channel hot electrons on current collapse in AlGaN/GaN HEMTs," *IEEE Electron Device Letters*, vol. 34, no. 12, pp. 1494–1496, 2013.

[97] D. Bisi, M. Meneghini, C. D. Santi, A. Chini, M. Dammann, P. Brückner, M. Mikulla, G. Meneghesso, and E. Zanoni, "Deep-level characterization in GaN HEMTs-part I: Advantages and limitations of drain current transient measurements," *IEEE Transactions on Electron Devices*, vol. 60, no. 10, pp. 3166–3175, 2013.

[98] F. Gao, B. Lu, L. Li, S. Kaun, J. S. Speck, C. V. Thompson, and T. Palacios, "Role of oxygen in the off-state degradation of AlGaN/GaN high electron mobility transistors," *Applied Physics Letters*, vol. 99, no. 22, p. 223506, 2011.

[99] Y. Puzyrev, S. Mukherjee, J. Chen, T. Roy, M. Silvestri, R. D. Schrimpf, D. M. Fleetwood, J. Singh, J. M. Hinckley, A. Paccagnella, and S. T. Pantelides, "Gate bias dependence of defect-mediated hot-carrier degradation in GaN HEMTs," *IEEE Transactions on Electron Devices*, vol. 61, no. 5, pp. 1316–1320, 2014.

[100] N. Moultif, O. Latry, E. Joubert, M. Ndiaye, C. Moreau, J. F. Goupy, and P. Carton, "Reliability assessment of AlGaN/GaN HEMTs on the SiC substrate under the RF stress," *IEEE Transactions on Power Electronics*, vol. 36, no. 7, pp. 7442–7450, 2021.

[101] D. J. Cheney, E. A. Douglas, L. Liu, C. F. Lo, Y. Y. Xi, B. P. Gila, F. Ren, D. Horton, M. E. Law, D. J. Smith, and S. J. Pearton, "Reliability studies of AlGaN/GaN high electron mobility transistors," *Semiconductor Science and Technology*, vol. 28, no. 7, p. 074019, June 2013.

[102] G. Meneghesso, M. Meneghini, A. Stocco, D. Bisi, C. de Santi, I. Rossetto, A. Zanandrea, F. Rampazzo, and E. Zanoni, "Degradation of AlGaN/GaN HEMT devices: Role of reverse-bias and hot electron stress," *Microelectronic Engineering*, vol. 109, pp. 257–261, 2013, insulating Films on Semiconductors 2013.

[103] H. Rao and G. Bosman, "Hot-electron induced defect generation in AlGaN/GaN high electron mobility transistors," *Solid-State Electronics*, vol. 79, pp. 11–13, 2013.

[104] A. D. Latorre-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, pp. 4263–4270, 2018.

[105] K. Merrill and M. Saraniti, "Nonlinear Electro-Thermal Monte Carlo Device Simulation," *Journal of Heat Transfer*, vol. 142, no. 2, 12 2019, 022106.

[106] 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, no. 2, pp. 908–916, 2019.

[107] R. C. Fitch, D. E. Walker, A. J. Green, S. Tetlak, J. K. Gillespie., R. D. Gilbert, K. A. Sutherlin, W. D. Gouty, J. P. Theimer, G. D. Via, K. D. Chabak, and G. H. Jessen, "Implementation of High-Power-Density X -Band AlGaN/GaN High Electron Mobility Transistors in a Millimeter-Wave Monolithic Microwave Integrated Circuit Process," *IEEE Electron Device Letters*, vol. 36, no. 10, pp. 1004–1007, 2015.

[108] K. Sehra, V. Kumari, V. Nath, M. Gupta, and M. Saxena, "Optimization of asymmetric Π-gate HEMT for improved reliability frequency applications," in *2019 IEEE 9th International Nanoelectronics Conferences (INEC)*, 2019, pp. 1–4.

[109] K. Sehra, V. Kumari, V. Nath, M. Gupta, D. S. Rawal, and M. Saxena, "Comparison of linearity and intermodulation distortion metrics for T and Π -gate HEMT," in *2019 International Conference on Electrical, Electronics and Computer Engineering (UPCON)*, 2019, pp. 1–6.

[110] K. Sehra, V. Kumari, M. Gupta, M. Mishra, D. S. Rawal, and M. Saxena, "Optimization of Π–gate AlGaN/AlN/GaN HEMTs for low noise and high gain applications," *Silicon*, pp. 1–12, 2020.

[111] K. Sehra, V. Kumari, M. Gupta, M. Mishra, D. Rawal, and M. Saxena, "TCAD investigation of gate-lag measurements on conventional and Π-gate AlGaN/GaN HEMTs," in *2020 IEEE 20th International Conference on Nanotechnology (IEEE-NANO)*.   IEEE, 2020, pp. 128–133.

[112] K. Sehra, Chanchal, A. Anand, V. Kumari, Reeta, M. Gupta, M. Mishra, D. S. Rawal, and M. Saxena, "Efficacy of Π-gate in RF power performance of thin GaN Buffer AlGaN/GaN HEMTs," *IEEE Transactions on Electron Devices*, vol. 70, no. 5, pp. 2612–2615, 2023.

[113] F. A. Marino, N. Faralli, T. Palacios, D. Ferry, S. Goodnick, and M. Saraniti, "Effects of Threading Dislocations on AlGaN/GaN High-Electron Mobility Transistors," *IEEE Transactions on Electron Devices*, vol. 57, no. 1, pp. 353–360, 2010.

[114] Y. Sun, X. Shen, J. Wang, D. Zhao, G. Feng, Y. Fu, S. Zhang, Z. Zhang, Z. Feng, Y. Bai *et al.*, "Thermal annealing behaviour of Ni/Au on *n*-GaN Schottky contacts," *Journal of Physics D: Applied Physics*, vol. 35, no. 20, p. 2648, 2002.

[115] D. Kotchetkov, J. Zou, A. A. Balandin, D. I. Florescu, and F. H. Pollak, "Effect of dislocations on thermal conductivity of GaN layers," *Applied Physics Letters*, vol. 79, no. 26, pp. 4316–4318, 2001.

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

[117] 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, 2013.

[118] V. Palankovski and S. Selberherr, "Thermal models for semiconductor device simulation," in *HITEN 99. Third European Conference on High Temperature Electronics.(IEEE Cat. No. 99EX372)*. IEEE, 1999, pp. 25–28.

[119] F. Bonani and G. Ghione, "On the application of the Kirchhoff transformation to the steady-state thermal analysis of semiconductor devices with temperature-dependent and piecewise inhomogeneous thermal conductivity," *Solid-State Electronics*, vol. 38, no. 7, pp. 1409–1412, 1995.

[120] F. Schwierz and J. J. Liou, *Modern microwave transistors: theory, design, and performance*. New Jersey: Wiley-interscience, 2003.

[121] P. Tasker and B. Hughes, "Importance of source and drain resistance to the maximum f_T of millimeter-wave MODFETs," *IEEE Electron Device Letters*, vol. 10, no. 7, pp. 291–293, 1989.

[122] Y. Zhang and Q.-Z. Yin, "Carbon and other light element contents in the earth's core based on first-principles molecular dynamics," *Proceedings of the National Academy of Sciences*, vol. 109, no. 48, pp. 19 579–19 583, 2012.

[123] N. Govindaraju, "Thermal Conductivity Analysis of Diamond Films," Ph.D. dissertation, North Carolina State University, 2010.

[124] J. Robertson, "Electronic structure of diamond-like carbon," *Diamond and Related Materials*, vol. 6, no. 2-4, pp. 212–218, 1997.

[125] C. E. Nebel, "General properties of diamond," in *Nanodiamonds*. Elsevier, 2017, pp. 1–24.

[126] O. Madelung, *Semiconductors: group IV elements and III-V compounds*. Springer Science & Business Media, 2012.

[127] F. Bundy, H. T. Hall, H. Strong, and R. Wentorfjun, "Man-made diamonds," *nature*, vol. 176, no. 4471, pp. 51–55, 1955.

[128] J. C. Angus, H. A. Will, and W. S. Stanko, "Growth of diamond seed crystals by vapor deposition," *Journal of Applied Physics*, vol. 39, no. 6, pp. 2915–2922, 1968.

[129] K. Kurihara, K. Sasaki, M. Kawarada, and N. Koshino, "High rate synthesis of diamond by dc plasma jet chemical vapor deposition," *Applied physics letters*, vol. 52, no. 6, pp. 437–438, 1988.

[130] A. Bolshakov, V. Ralchenko, G. Shu, B. Dai, V. Y. Yurov, E. Bushuev, A. Khomich, A. Altakhov, E. Ashkinazi, I. Antonova *et al.*, "Single crystal diamond growth by mpcvd at subatmospheric pressures," *Materials Today Communications*, vol. 25, p. 101635, 2020.

[131] R. Haubner and B. Lux, "Diamond growth by hot-filament chemical vapor deposition: state of the art," *Diamond and Related Materials*, vol. 2, no. 9, pp. 1277–1294, 1993.

[132] T. Anthony, J. Fleischer, J. Olson, and D. G. Cahill, "The thermal conductivity of isotopically enriched polycrystalline diamond films," *Journal of applied physics*, vol. 69, no. 12, pp. 8122–8125, 1991.

[133] A. Ward, D. Broido, D. A. Stewart, and G. Deinzer, "Ab initio theory of the lattice thermal conductivity in diamond," *Physical Review B*, vol. 80, no. 12, p. 125203, 2009.

[134] T. R. Anthony and W. F. Banholzer, "Properties of diamond with varying isotopic composition," *Diamond and Related Materials*, vol. 1, no. 5-6, pp. 717–726, 1992.

[135] H. Umezawa, M. Nagase, Y. Kato, and S.-i. Shikata, "High temperature application of diamond power device," *Diamond and related materials*, vol. 24, pp. 201–205, 2012.

[136] O. S. Chaudhary, M. Denaï, S. S. Refaat, and G. Pissanidis, "Technology and applications of wide bandgap semiconductor materials: Current state and future trends," *Energies*, vol. 16, no. 18, p. 6689, 2023.

[137] B. J. Baliga, "Power semiconductor device figure of merit for high-frequency applications," *IEEE Electron Device Letters*, vol. 10, no. 10, pp. 455–457, 1989.

[138] A. Q. Huang, "New unipolar switching power device figures of merit," *IEEE Electron Device Letters*, vol. 25, no. 5, pp. 298–301, 2004.

[139] J. Shealy, V. Kaper, V. Tilak, T. Prunty, J. Smart, B. Green, and L. Eastman, "An AlGaN/GaN high-electron-mobility transistor with an AlN sub-buffer layer," *Journal of Physics: Condensed Matter*, vol. 14, no. 13, p. 3499, 2002.

[140] A. Bar-Cohen, J. Maurer, and D. Altman, "Embedded cooling for wide bandgap power amplifiers: a review," *Journal of Electronic Packaging*, vol. 141, no. 4, p. 040803, 2019.

[141] J. G. Felbinger, M. V. S. Chandra, Y. Sun, L. F. Eastman, J. Wasserbauer, F. Faili, D. Babic, D. Francis, and F. Ejeckam, "Comparison of GaN HEMTs on Diamond and SiC Substrates," *IEEE Electron Device Letters*, vol. 28, no. 11, pp. 948–950, 2007.

[142] P.-C. Chao, K. Chu, C. Creamer, J. Diaz, T. Yurovchak, M. Shur, R. Kallaher, C. McGray, G. D. Via, and J. D. Blevins, "Low-Temperature Bonded GaN-on-Diamond HEMTs With 11 W/mm Output Power at 10 GHz," *IEEE Transactions on Electron Devices*, vol. 62, no. 11, pp. 3658–3664, 2015.

[143] L. Sang, "Diamond as the heat spreader for the thermal dissipation of GaN-based electronic devices," *Functional Diamond*, vol. 1, no. 1, pp. 174–188, 2022.

[144] M.-C. Lu, "Development of Diamond Device-Level Heat Spreader for the Advancement of GaN HEMT Power and RF Electronics," *IEEE Transactions on Device and Materials Reliability*, 2023.

[145] P.-N. Volpe, P. Muret, J. Pernot, F. Omnès, T. Teraji, F. Jomard, D. Planson, P. Brosselard, N. Dheilly, B. Vergne *et al.*, "High breakdown voltage Schottky diodes synthesized on p-type CVD diamond layer," *physica status solidi (a)*, vol. 207, no. 9, pp. 2088–2092, 2010.

[146] V. Bormashov, S. Terentiev, S. Buga, S. Tarelkin, A. Volkov, D. Teteruk, N. Kornilov, M. Kuznetsov, and V. Blank, "Thin large area vertical Schottky barrier diamond diodes with low on-resistance made by ion-beam assisted lift-off technique," *Diamond and Related Materials*, vol. 75, pp. 78–84, 2017.

[147] P.-N. Volpe, P. Muret, J. Pernot, F. Omnès, T. Teraji, Y. Koide, F. Jomard, D. Planson, P. Brosselard, N. Dheilly *et al.*, "Extreme dielectric strength in boron doped homoepitaxial diamond," *Applied Physics Letters*, vol. 97, no. 22, 2010.

[148] H. Umezawa, Y. Kato, and S. ichi Shikata, "1 On-Resistance Diamond Vertical-Schottky Barrier Diode Operated at 250 ℃," *Applied Physics Express*, vol. 6, no. 1, p. 011302, dec 2012. [Online]. Available: https://dx.doi.org/10.7567/APEX.6.011302

[149] H. Umezawa and S.-i. Shikata, "Leakage current analysis of diamond Schottky barrier diodes operated at high temperature," *Japanese Journal of Applied Physics*, vol. 53, no. 4S, p. 04EP04, 2014.

[150] C. Masante, N. Rouger, and J. Pernot, "Recent progress in deep-depletion diamond metal–oxide–semiconductor field-effect transistors," *Journal of Physics D: Applied Physics*, vol. 54, no. 23, p. 233002, 2021.

[151] N. Donato, N. Rouger, J. Pernot, G. Longobardi, and F. Udrea, "Diamond power devices: state of the art, modelling, figures of merit and future perspective," *Journal of Physics D: Applied Physics*, vol. 53, no. 9, p. 093001, 2019.

[152] H. Kato, K. Oyama, T. Makino, M. Ogura, D. Takeuchi, and S. Yamasaki, "Diamond bipolar junction transistor device with phosphorus-doped diamond base layer," *Diamond and related materials*, vol. 27, pp. 19–22, 2012.

[153] N. Oi, M. Inaba, S. Okubo, I. Tsuyuzaki, T. Kageura, S. Onoda, A. Hiraiwa, and H. Kawarada, "Vertical-type two-dimensional hole gas diamond metal oxide semiconductor field-effect transistors," *Scientific reports*, vol. 8, no. 1, p. 10660, 2018.

[154] J. Callaway, "Model for lattice thermal conductivity at low temperatures," *Physical Review*, vol. 113, no. 4, p. 1046, 1959.

[155] P. Torres, A. Torelló, J. Bafaluy, J. Camacho, X. Cartoixà, and F. Alvarez, "First principles kinetic-collective thermal conductivity of semiconductors," *Physical Review B*, vol. 95, no. 16, p. 165407, 2017.

[156] A. Inyushkin, A. Taldenkov, V. Ralchenko, A. Bolshakov, A. Koliadin, and A. Katrusha, "Thermal conductivity of high purity synthetic single crystal diamonds," *Physical Review B*, vol. 97, no. 14, p. 144305, 2018.

[157] N. Novikov, A. Podoba, S. Shmegera, A. Witek, A. Zaitsev, A. Denisenko, W. Fahrner, and M. Werner, "Influence of isotopic content on diamond thermal conductivity," *Diamond and related materials*, vol. 8, no. 8-9, pp. 1602–1606, 1999.

[158] M. Asen-Palmer, K. Bartkowski, E. Gmelin, M. Cardona, A. Zhernov, A. Inyushkin, A. Taldenkov, V. Ozhogin, K. M. Itoh, and E. Haller, "Thermal conductivity of germanium crystals with different isotopic compositions," *Physical review B*, vol. 56, no. 15, p. 9431, 1997.

[159] D. Morelli, J. Heremans, and G. Slack, "Estimation of the isotope effect on the lattice thermal conductivity of group IV and group III-V semiconductors," *Physical Review B*, vol. 66, no. 19, p. 195304, 2002.

[160] Y.-J. Han and P. Klemens, "Anharmonic thermal resistivity of dielectric crystals at low temperatures," *Physical Review B*, vol. 48, no. 9, p. 6033, 1993.

[161] Y.-J. Han and H.-B. Chae, "Theoretical analysis of the thermal conductivity of diamond in a two-step model," *Physical Review B*, vol. 52, no. 1, p. 27, 1995.

[162] A. Sparavigna, "Influence of isotope scattering on the thermal conductivity of diamond," *Physical Review B*, vol. 65, no. 6, p. 064305, 2002.

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

[164] W. Kohn and L. J. Sham, "Self-consistent equations including exchange and correlation effects," *Physical review*, vol. 140, no. 4A, p. A1133, 1965.

[165] A. V. Inyushkin, A. N. Taldenkov, V. G. Ralchenko, A. P. Bolshakov, and A. V. Khomich, "Isotope effect in thermal conductivity of polycrystalline CVD-Diamond: Experiment and theory," *Crystals*, vol. 11, no. 4, p. 322, 2021.

[166] K. T. Regner, J. P. Freedman, and J. A. Malen, "Advances in studying phonon mean free path dependent contributions to thermal conductivity," *Nanoscale and Microscale Thermophysical Engineering*, vol. 19, no. 3, pp. 183–205, 2015.

[167] W. Li, N. Mingo, L. Lindsay, D. A. Broido, D. A. Stewart, and N. A. Katcho, "Thermal conductivity of diamond nanowires from first principles," *Physical Review B*, vol. 85, no. 19, p. 195436, 2012.

[168] X. Qian, J. Zhou, and G. Chen, "Phonon-engineered extreme thermal conductivity materials," *Nature Materials*, vol. 20, no. 9, pp. 1188–1202, 2021.

[169] S. Li, Q. Zheng, Y. Lv, X. Liu, X. Wang, P. Y. Huang, D. G. Cahill, and B. Lv, "High thermal conductivity in cubic boron arsenide crystals," *Science*, vol. 361, no. 6402, pp. 579–581, 2018.

[170] B. Lv, Y. Lan, X. Wang, Q. Zhang, Y. Hu, A. J. Jacobson, D. Broido, G. Chen, Z. Ren, and C.-W. Chu, "Experimental study of the proposed super-thermal-conductor: BAs," *Applied Physics Letters*, vol. 106, no. 7, 2015.

[171] L. Lindsay, D. Broido, and T. Reinecke, "First-principles determination of ultrahigh thermal conductivity of boron arsenide: a competitor for diamond?" *Physical review letters*, vol. 111, no. 2, p. 025901, 2013.

[172] D. A. Broido, M. Malorny, G. Birner, N. Mingo, and D. Stewart, "Intrinsic lattice thermal conductivity of semiconductors from first principles," *Applied Physics Letters*, vol. 91, no. 23, 2007.

[173] F. F. Sabatti, S. M. Goodnick, and M. Saraniti, "Simulation of phonon transport in semiconductors using a population-dependent many-body cellular Monte Carlo approach," *Journal of Heat Transfer*, vol. 139, no. 3, p. 032002, 2017.

[174] G. P. Srivastava, *The physics of phonons.* CRC press, 2022.

[175] P. Klemens, "Anharmonic attenuation of localized lattice vibrations," *Physical Review*, vol. 122, no. 2, p. 443, 1961.

[176] P. Klemens, "Thermal conductivity and lattice vibrational modes," in *Solid state physics.* Elsevier, 1958, vol. 7, pp. 1–98.

[177] P. Klemens, "The scattering of low-frequency lattice waves by static imperfections," *Proceedings of the Physical Society. Section A*, vol. 68, no. 12, p. 1113, 1955.

[178] T. Feng, L. Lindsay, and X. Ruan, "Four-phonon scattering significantly reduces intrinsic thermal conductivity of solids," *Physical Review B*, vol. 96, no. 16, p. 161201, 2017.