Using Government drawings, specifications, or other data included in this document for any purpose other than Government procurement does not in any way obligate the U.S. Government. The fact that the Government formulated or supplied the drawings, specifications, or other data does not license the holder or any other person or corporation; or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

This report is the result of contracted fundamental research deemed exempt from public affairs security and policy review in accordance with SAF/AQR memorandum dated 10 Dec 08 and AFRL/CA policy clarification memorandum dated 16 Jan 09. This report is available to the general public, including foreign nationals.

Copies may be obtained from the Defense Technical Information Center (DTIC) (http://www.dtic.mil).

AFRL-RY-WP-TR-2019-0169 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

*//Signature//      //Signature//
CATALIN S. BADESCU ROSS W. DETTMER, Chief
Program Manager Electronic Devices Branch
Electronic Devices Branch Aerospace Components & Subsystems Division
Aerospace Components & Subsystems Division

//Signature//
JUSTIN W. CLEARY
Deputy (Acting)
Aerospace Components & Subsystems Division
Sensors Directorate

This report is published in the interest of scientific and technical information exchange, and its publication does not constitute the Government’s approval or disapproval of its ideas or findings.

*Disseminated copies will show “//Signature//” stamped or typed above the signature blocks.
REPORT DOCUMENTATION PAGE

1. REPORT DATE          2. REPORT TYPE          3. DATES COVERED
   (DD-MM-YY)            Final                        From - To
   September 2019        22 April 2015 – 21 February 2019

4. TITLE AND SUBTITLE
   ELECTRO-THERMAL SCALING PROPERTIES OF ELECTRON DEVICES:
   A MULTI-SCALE MONTE CARLO APPROACH

5a. CONTRACT NUMBER      5b. GRANT NUMBER
   FA8650-15-1-7525
   5c. PROGRAM ELEMENT NUMBER
   61101E
   5d. PROJECT NUMBER
   1000
   5e. TASK NUMBER
   N/A
   5f. WORK UNIT NUMBER
   Y14Q

6. AUTHOR(S)
   Marco Saraniti and John Albrecht

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
   Arizona State University
   660 S Mill Avenue, Suite 312
   Tempe, AZ 85281-3670

8. PERFORMING ORGANIZATION REPORT NUMBER
   AFRL-RY-WP-TR-2019-0169

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
   Air Force Research Laboratory
   Sensors Directorate
   Wright-Patterson Air Force Base, OH 45433-7320
   Air Force Materiel Command
   United States Air Force
   Defense Advanced Research Projects Agency
   (DARPA/MTO)
   675 North Randolph St.
   Arlington, VA 22203

10. SPONSORING/MONITORING AGENCY ACRONYM(S)
    AFRL/RYDD

11. SPONSORING/MONITORING AGENCY REPORT NUMBER(S)
    AFRL-RY-WP-TR-2019-0169

12. DISTRIBUTION/AVAILABILITY STATEMENT
    Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES
    This material is based on research sponsored by Air Force Research laboratory (AFRL) and the Defense Advanced Research Agency (DARPA) under agreement number FA8650-15-1-7525. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies of endorsements, either expressed or implied, of AFRL and DARPA or the U.S. Government. Report contains color.

14. ABSTRACT
    This report details the completion of a modeling approach for the development of the thermal management of solid-state devices and their enclosures. By the full inclusion of the phonon dynamics within the framework of charge transport, extremely efficient particle-based microscopic models in computer-aided design (CAD) tools for the design of electron devices has been implemented. To complement the microscopic models of the particle-based approach, we integrated them in a numerical framework based on a particle-flux paradigm in which electrons are treated as particles, while the phonon dynamics is represented as a heat flux. The two modeling approaches are a single multi-scale code that allows the extraction of novel scaling laws and performance metrics by modeling charge and heat transport as a property of the devices and the environment. Such code will allow for the first time, to the best of our knowledge, the identification of electro-thermal design metrics and scaling laws for the design of electron devices integrated with their passive heat management components.

15. SUBJECT TERMS
    radio frequency power transistors, thermal effects, phonon generation, computer-aided design

16. SECURITY CLASSIFICATION OF:
    a. REPORT Unclassified
    b. ABSTRACT Unclassified
    c. THIS PAGE Unclassified

17. LIMITATION OF ABSTRACT:
    SAR

18. NUMBER OF PAGES
    70

19a. NAME OF RESPONSIBLE PERSON (Monitor)
    Catalin Badescu

19b. TELEPHONE NUMBER (Include Area Code)
    N/A
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>List of Figures</td>
<td>ii</td>
</tr>
<tr>
<td>1. SUMMARY</td>
<td>1</td>
</tr>
<tr>
<td>2. INTRODUCTION</td>
<td>2</td>
</tr>
<tr>
<td>3. METHODS, ASSUMPTIONS, AND PROCEDURES</td>
<td>3</td>
</tr>
<tr>
<td>3.1 Energy Balance Equation for Phonons</td>
<td>3</td>
</tr>
<tr>
<td>3.2 Coupled Electron-Phonon Dynamics for Electro-Thermal Simulation</td>
<td>6</td>
</tr>
<tr>
<td>3.2.1 Thermal Conductivity</td>
<td>8</td>
</tr>
<tr>
<td>3.2.2 Isotope Scattering</td>
<td>9</td>
</tr>
<tr>
<td>4. RESULTS AND DISCUSSION</td>
<td>12</td>
</tr>
<tr>
<td>4.1 Energy Balance Equation for Phonons</td>
<td>12</td>
</tr>
<tr>
<td>4.1.2 Simulation Results</td>
<td>13</td>
</tr>
<tr>
<td>4.2 Robustness of the Solver</td>
<td>16</td>
</tr>
<tr>
<td>4.2.1 Simulation Results</td>
<td>16</td>
</tr>
<tr>
<td>4.2.2 Separate Thermal and Electrical Regions</td>
<td>17</td>
</tr>
<tr>
<td>4.3 Energy Balance Equation Averaging Procedure</td>
<td>19</td>
</tr>
<tr>
<td>4.3.1 Kirchhoff Transformation</td>
<td>20</td>
</tr>
<tr>
<td>4.3.2 Possible Effect of Quantum Confinement</td>
<td>21</td>
</tr>
<tr>
<td>4.4 Cooling Effect and Thermal Simulations on a GaN/AlGaN HEMT</td>
<td>23</td>
</tr>
<tr>
<td>4.5 Low Bias Simulations on a GaN/AlGaN HEMT</td>
<td>29</td>
</tr>
<tr>
<td>4.5.1 High Bias Simulations on a GaN/AlGaN HEMT</td>
<td>31</td>
</tr>
<tr>
<td>4.6 Merging of Separate Code Versions</td>
<td>32</td>
</tr>
<tr>
<td>4.6.1 Error Handling and Robustness Improvements</td>
<td>34</td>
</tr>
<tr>
<td>4.7 Fully Coupled Electrothermal Approach</td>
<td>35</td>
</tr>
<tr>
<td>4.7.1 Electrothermal Simulations of a GaN HEMT</td>
<td>36</td>
</tr>
<tr>
<td>4.8 Rejection Algorithm for Scattering</td>
<td>38</td>
</tr>
<tr>
<td>4.8.1 Phonon Scattering Rate</td>
<td>39</td>
</tr>
<tr>
<td>4.8.2 Electron-phonon Interaction</td>
<td>40</td>
</tr>
<tr>
<td>4.9 Electron-phonon Interaction</td>
<td>42</td>
</tr>
<tr>
<td>4.9.1 Hawkes Process for Phonon Absorption and Emission</td>
<td>44</td>
</tr>
<tr>
<td>4.10 Flux-based Approach and Convective Conditions</td>
<td>46</td>
</tr>
<tr>
<td>4.10.1 Iterative Approach for including Temperature Dependent Thermal Conductivity</td>
<td>46</td>
</tr>
<tr>
<td>4.10.2 Phonon Dynamics</td>
<td>48</td>
</tr>
<tr>
<td>4.11 Multi-finger GaN on SiC Package</td>
<td>50</td>
</tr>
<tr>
<td>4.11.1 Simulation of Interface Resistances</td>
<td>51</td>
</tr>
<tr>
<td>4.12 Phonon Dynamics</td>
<td>53</td>
</tr>
<tr>
<td>4.13 Flux-based Approach and Multiscale Simulation</td>
<td>55</td>
</tr>
<tr>
<td>4.13.1 GaAs MESFET Simulation</td>
<td>56</td>
</tr>
<tr>
<td>4.14 Particle-Flux Energy Balance Equation: T-gate and II-gate HEMT Results</td>
<td>57</td>
</tr>
<tr>
<td>5. CONCLUSIONS</td>
<td>61</td>
</tr>
<tr>
<td>6. REFERENCES</td>
<td>63</td>
</tr>
<tr>
<td>LIST OF ABBREVIATIONS, ACRONYMS, AND SYMBOLS</td>
<td>64</td>
</tr>
</tbody>
</table>
# List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1: Final Temperature as a Function of the Position within the Device where the left contact is set at 300K and the right at 350 K</td>
<td>6</td>
</tr>
<tr>
<td>Figure 2: Effective Temperature in a Silicon Thermal Resistor as Function of the Position</td>
<td>9</td>
</tr>
<tr>
<td>Figure 3: Thermal Conductivity as a Function of the Position</td>
<td>9</td>
</tr>
<tr>
<td>Figure 4: Comparison between Measured (lines) [14] and Simulated (symbols) Thermal Conductivity as a Function of Temperature in Multi-isotopic Natural ($Si_{nat}$) and Pure ($Si^{28}$) Silicon</td>
<td>11</td>
</tr>
<tr>
<td>Figure 5: Steady-state Temperature Distributions considering only Electron-phonon Scattering</td>
<td>13</td>
</tr>
<tr>
<td>Figure 6: Steady-state Temperature Distributions considering both Electron-phonon Scattering and Decay of Optical Phonons into Acoustic Phonons</td>
<td>14</td>
</tr>
<tr>
<td>Figure 7: Steady-state Temperature Distributions considering both Electron-phonon Scattering and Decay of Optical Phonons into Acoustic Phonons</td>
<td>15</td>
</tr>
<tr>
<td>Figure 8: Steady-state Temperature Distributions with a 10 ps Optical Phonon Decay Lifetime using the new Approach</td>
<td>17</td>
</tr>
<tr>
<td>Figure 9: Steady-state Temperature Distributions in the MESFET including a Region in which only Heat Transport is solved</td>
<td>18</td>
</tr>
<tr>
<td>Figure 10: Comparison of Temperature Maps for a Simple MESFET using the Algebraic Transformation (left) and that Computed and Stored in Discrete Arrays (right)</td>
<td>21</td>
</tr>
<tr>
<td>Figure 11: Example of the Artificial Cooling Effect seen in a GaN/AlGaN HEMT Structure</td>
<td>22</td>
</tr>
<tr>
<td>Figure 12: Electro-thermal Results on a GaN/AlGaN HEMT using the $J \cdot E$ Forcing Function</td>
<td>24</td>
</tr>
<tr>
<td>Figure 13: Electro-thermal Results on a GaN/AlGaN HEMT using the $J \cdot E$ Forcing Function</td>
<td>24</td>
</tr>
<tr>
<td>Figure 14: Electro-thermal Results on a GaN/AlGaN HEMT using the Phonon Emission Forcing Function</td>
<td>26</td>
</tr>
<tr>
<td>Figure 15: Electro-thermal Results on a GaN/AlGaN HEMT using the Phonon Emission Forcing Function</td>
<td>27</td>
</tr>
<tr>
<td>Figure 16: Electron Energies and associated Temperatures at Low Electric Fields in Bulk Simulations</td>
<td>28</td>
</tr>
<tr>
<td>Figure 17: Electron Energies and associated Temperatures at Low Electric Fields in Bulk Simulations</td>
<td>29</td>
</tr>
<tr>
<td>Figure 18: Acoustic and Optical Temperature Maps for a Device with a Gate Bias of $VG = -0.5$ V and a Zero Drain Bias</td>
<td>30</td>
</tr>
<tr>
<td>Figure 19: Previously seen Temperature Map obtained using only the x-components of the Velocity and Electric Field to compute the Forcing Function</td>
<td>31</td>
</tr>
<tr>
<td>Figure 20: Temperature Map using the Energies computed directly from the Phonon Absorption and Emission Scattering Events in the CMC</td>
<td>32</td>
</tr>
<tr>
<td>Figure 21: Difference in tabulated Electron-phonon Term of the Forcing Function for the Energy Balance Equation in the same GaAs MESFET Structure</td>
<td>34</td>
</tr>
<tr>
<td>Figure 22: Device Layout used in Simulations, from Altuntas et al [1]</td>
<td>36</td>
</tr>
<tr>
<td>Figure</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>------</td>
</tr>
<tr>
<td>Figure 23: Acoustic and Optical Temperature Maps obtained through the Coupled Electrothermal CMC Code</td>
<td>36</td>
</tr>
<tr>
<td>Figure 24: Drain Current for the Device under Study vs Drain Voltage (left) and Gate Voltage (right)</td>
<td>37</td>
</tr>
<tr>
<td>Figure 25: Profile of the Electron x-velocity (a) and Total Scattering (b) along the Channel</td>
<td>38</td>
</tr>
<tr>
<td>Figure 26: Rejection Algorithm Flow Chart</td>
<td>39</td>
</tr>
<tr>
<td>Figure 27: Total Energy lost by the Charge Ensemble and that gained by the Phonon Ensemble with Phonon Emission/Absorption treated as a Poisson Process</td>
<td>41</td>
</tr>
<tr>
<td>Figure 28: Total Energy lost by the Charge Ensemble and that gained by the Phonon Ensemble with Phonon Emission/Absorption treated as a Poisson Process</td>
<td>43</td>
</tr>
<tr>
<td>Figure 29: Total Energy lost by the Charge Ensemble and that gained by the Phonon Ensemble with Phonon Emission/Absorption modified using a Hawkes Process</td>
<td>45</td>
</tr>
<tr>
<td>Figure 30: Temperature along the Source Plane, i.e., at the Top of the Device</td>
<td>47</td>
</tr>
<tr>
<td>Figure 31: Electron Distribution initialized at 20 meV above the Conduction Band shown at t = 0.6 fs and t = 300 fs</td>
<td>48</td>
</tr>
<tr>
<td>Figure 32: Jensen-Shannon Divergence of the Correct Reference Distribution and the Simulation Distribution as a Function of Time</td>
<td>49</td>
</tr>
<tr>
<td>Figure 33: Quarter-model Simulation Layout for the 22 Finger GaN HEMT Structure</td>
<td>50</td>
</tr>
<tr>
<td>Figure 34: Comparison between the Iterative Approaches in the CMC with those obtained by Bagnall et al [2]</td>
<td>51</td>
</tr>
<tr>
<td>Figure 35: Simulation of a Single Interface Resistance with differing (constant) Thermal Conductivities on each Side (left) and Simulation of a 10 layered Structure similar to that of Hickson et al [3]</td>
<td>52</td>
</tr>
<tr>
<td>Figure 36: Hole (left) and Electron (right) Distributions at t = 0.1 fs</td>
<td>54</td>
</tr>
<tr>
<td>Figure 37: Hole (left) and Electron (right) Distributions at t = 0.1 fs</td>
<td>54</td>
</tr>
<tr>
<td>Figure 38: The Electron Distribution enlarged as to examine the Distance between the clearly Visible Peaks</td>
<td>55</td>
</tr>
<tr>
<td>Figure 39: Temperature Maps obtained for the GaAs MESFET Simulation</td>
<td>56</td>
</tr>
<tr>
<td>Figure 40: T-gate HEMT Layout under Study from [2]</td>
<td>57</td>
</tr>
<tr>
<td>Figure 41: Simulated (left) ID-V_DS (left) and ID-V_GS (right) for the Model of the Experimental T-gate Structure in Figure 40</td>
<td>58</td>
</tr>
<tr>
<td>Figure 42: Π-gate Architecture (left) as compared with the original T-gate (right)</td>
<td>59</td>
</tr>
</tbody>
</table>
1. SUMMARY

In terms of applications, we envision a dual purpose for the technology computer-aided design (TCAD) tools presented in this document: the extraction of scaling laws for specific families of electron devices and the utilization of those scaling laws for the design and optimization of specific devices. Most importantly, the extraction of scaling laws will include packaging and passive thermal management components coupled to the active devices. Furthermore, the determination of the relation between the whole electro-thermal setup and the power and frequency response of the devices will be performed in large-signal regime.

The second purpose of the proposed software tools is related to their capability of being effectively employed in the device design and optimization phases. The algorithmic approach used by our code is based on the Cellular Monte Carlo (CMC) method, which has been developed by principal investigator (PI) Saraniti in the last two decades. Such approach achieves a considerable speed-up as compared as traditional particle-based codes, by employing an extremely large, yet realistic, amount of fast computer storage (RAM). Such speed-up makes the simulation of electron and phonon scattering up to 50 times faster than a standard run-time algorithm for the selection and calculation of the state of a particle after scattering. While the CMC algorithm is still sensibly slower than the approaches based on the first few moments of the Boltzmann transport equation, the efficiency and robustness of our specific code makes it a valuable tool for the device designer.

Because of the recent introduction of several new families of high-power, high-frequency transistors based on multilayered structures and realized with wide gap semiconductors, the realization of non-heuristic, physics-based computer-aided design (CAD) tools of unprecedented accuracy is definitely needed by the industry, and it is not limited to the academic environment. In particular, the highly nonlinear nature of the microscopic models of heat generation and transport in such devices makes the implementation of such multi-scale nonlinear models a numerically arduous problem.

The broader strategic relevance of the proposed research program is based on the transformative technology that will be enabled by the modeling and design capabilities of the software tools described in this document. Referring to the field of applications described above, we are confident that our research program will allow the realization of extremely reliable electronic components both for ultra-low power portable electronic devices and space-based maintenance-free high power communication systems. Our integrated CAD tools will allow the design of thermally optimized active electronic components that will couple high power efficiency with extreme reliability. These properties will reduce the maintenance costs and the cooling requirements; therefore the weight and battery run time, of portable electronics. As a result, more functionality will be integrated in fewer, smaller, and extremely robust components.
2. INTRODUCTION

The crucial aspect that needs to be addressed in order to realize the transformational electronic technology mentioned above is the optimization of the interaction between charge carriers and lattice vibrations responsible for charge and heat transport in crystalline solids. Such interactions result in the generation and transport of heat that has to be extracted by the device package and/or managed internally. A crucial outcome of the study proposed here is the understanding of the behavior and performance of the devices in conditions of ideal heat extraction from the surface of the device. The presence of perfect radiators located at the surface of the device will be simulated, and a microscopic study of the mechanisms of heat generation, retention, and transport within the device will be performed. To the best of our knowledge, the dynamical path of heat carriers in the phase space has never been studied with the resolution and accuracy of particle-based simulation. The dynamics of such path originates from the hot spot and ends at the extraction surfaces at the periphery of the device, and is ultimately responsible for the local temperature of the device, its performance, and its reliability.

An aspect that plays a crucial role for the realization of electro-thermally optimized devices is that the generation of heat occurs in regions of the device that are much far, in terms of heat diffusivity, from the phonon extraction gates, and the slow nature of the heat transport processes drives the local temperature of the hot spots to dangerously high values. In other words, a transistor can run hot even in the presence of a “perfect” radiator contacted to the device. While the realization of effective cooling surfaces can definitely mitigate the effects of Joule’s heating, the basic issue of generating fewer phonons or effectively moving them from the hot spot to a radiator half a dozen microns far from the hot spot can be addressed only through the microscopic study of the heat generation and transport mechanisms.

In other words, the continuous exchange of energy and momentum between the electrons and the vibrational states of their medium determines all the functional characteristics of solid-state semiconductor devices. Furthermore, the characteristic response of the medium to charge and heat transport is characterized by greatly different characteristic times and lengths, as well as influencing the electron/phonon population as a whole. The present proposal is motivated by the urgency of two concurrent problems that need to be addressed simultaneously in relation to the broad issue of thermal management: 1) the state-of-the-art modeling capabilities of industrial and academic CAD tools are insufficient for the understanding of electro-thermal effects in semiconductor devices and, at the same time, 2) the further development of almost any family of semiconductor devices is currently stopped because of such insufficient understanding. In other, simply put words: we urgently need to solve this problem, and we don’t know how. We therefore conclude that any significant advancement of the capabilities of solid-state devices – including the optoelectronics ones – will have to explicitly account for, and exploit, the dynamics of phonons. We are describing here a novel approach to that problem, that has been seed-funded by the Defense Advanced Research Agency (DARPA) and has been demonstrated feasible, and we are requesting funds to integrate it in robust software and to initiate its application to specific device structures.
3. METHODS, ASSUMPTIONS, AND PROCEDURES

As it has been stressed above, a previous research thrust, sponsored by DARPA, demonstrated the feasibility of the particle-based CMC method for the simulation of heat generation and transport in electron devices. The research project provided the design for the production software, and tested many of the algorithmic components separately. During the feasibility study we decided to divide the work in two separate trusts to be integrated at the conclusion of the project. The two trusts, or sub-projects, extend the CMC modeling capabilities on the opposite end of the simulation space: the microscopic interactions between charge carriers (electrons) and heat carriers (phonons). We offer here a description of the preliminary results of the two sub-projects, currently funded by DARPA, aiming to the realization of self-consistent electro-thermal modeling of semiconductor devices. The first one is based on a representation of the process of heat generation and transport within the device though the solution of the energy-balance equation for phonons, while the second models the same process via a direct representation of phonons as particles. In both cases the thermal models will be self-consistently coupled with the CMC algorithm used to model charge transport.

We feel that the strict integration of the software produced by the two sub-projects in the same software tool will constitute an unprecedented link between the parameter-free physical models of the CMC and the macroscopic parameters relevant for the design of high-power and high frequency devices. Such integration will be enabled by the research proposed in this document.

3.1 Energy Balance Equation for Phonons

As mentioned above, in order to achieve efficient electro-thermal modeling capabilities in the existing CMC code, we plan to implement the models for heat generation and transport in two complementary ways. The first approach is based on an efficient technique to reach the electro-thermal steady-state within the device layout. Such technique is based on the solution of the Energy Balance Equation (EBE) for phonons (flux-based) coupled self-consistently with the particle-based electron dynamics. Once the electro-thermal steady-state conditions were reached, a population of phonons will replace the temperature map supplied by the EBE solver. The particle-based phonon dynamics simulation engine will be then used in order to solve transients and non-equilibrium heat transport as described in the next section.

We decided to obtain the flux-based analysis of heat generation and transport by solving the EBE for phonons rather than the Heat Transport Equation (HTE). This approach would allow a more accurate temperature map by supplying a separate solution for each phonon mode (or group of modes), and is more suited to self-consistent coupling with the electron dynamics. The approach we developed starts from the phonon Boltzmann Transport Equation (BTE) (see equation 6) to obtain the energy balance equation for each phonon mode $\mu$:

$$ \frac{\partial W_\mu}{\partial t} = -\nabla \cdot F_W + \frac{\partial W_\mu}{\partial t} \bigg|_{e-p} + \frac{\partial W_\mu}{\partial t} \bigg|_{p-p} \quad (1) $$

where $W_\mu(r, t) = \frac{1}{\Omega} \sum_k E_\mu(k) f_\mu(r, k, t)$ is the ensemble energy in the volume $\Omega$ of the reciprocal space, $F_\mu(r, t) = \frac{1}{\Omega} \sum_k v(k) E_\mu(k) f_\mu(r, k, t)$ is the energy flux, and the two partial derivatives of

Approved for public release; distribution is unlimited.
$W_{\mu}$ in the left hand side (LHS) of the equation represent the rate of change of the ensemble energy due to electron-phonon and phonon-phonon interaction, respectively.

At steady-state, the equation reads:

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

(2)

In this case, $F_\mu(\mathbf{r}, t)$ has been approximated with the well-known steady-state relation $F_\mu(\mathbf{r}) = -k_\mu(T, \mathbf{r}) \nabla T$, where $k$ is the (scalar) thermal conductivity. The energy balance equation derived above is indeed the classical heat transport equation; the novelty of the proposed approach is in the way the forcing function is computed within the CMC framework as well as the fact that one of such equations can be solved for each phonon mode.

The component of the forcing function $P_\mu(\mathbf{r}) = \frac{\partial W_\mu}{\partial t} \bigg|_{e-p} + \frac{\partial W_\mu}{\partial t} \bigg|_{p-p}$ due to electron-phonon scattering is represented by the $\frac{\partial W_\mu}{\partial t} \bigg|_{e-p}$ term, and can be extracted in real time by recording the phonon emissions and absorptions for each mode $\mu$ after each electron-phonon scattering event. The phonons decays and recombination events contributing to the term $\frac{\partial W_\mu}{\partial t} \bigg|_{p-p}$ in the forcing function $P_\mu(\mathbf{r})$ are approximated via relaxation time approximation. It should be noted that other approaches present in literature are based on a solution of the HTE and rely on one or more of three main assumptions: 1) the assumption that a specific energy decay path exist (electron $\rightarrow$ optical phonon $\rightarrow$ acoustic phonon), 2) the assumption that the acoustic phonon scattering is elastic, and 3) the assumption that the relaxation time approximation can be used also for determining the electron-phonon energy relaxation term $\frac{\partial W_\mu}{\partial t} \bigg|_{e-p}$. All these approximations are not needed by the proposed treatment. In particular, our approach will produce the statistical relevance of each energy path (see assumption 1) above) as an output, rather than assuming a specific one. Finally, we note that even the relaxation time approximation for the phonon-phonon scattering can be abandoned once a particle-based approach will be completed for the phonon dynamics. At that moment, the $\frac{\partial W_\mu}{\partial t} \bigg|_{p-p}$ term in the forcing function can be extracted and tabulated as a function of the temperature from a simple bulk simulation performed by using phonons as particles.

In principle, the nonlinear steady-state energy balance equation above can be solved via a 3D nonlinear finite element solver capable of handling the functional dependencies of $k_\mu(T, \mathbf{r})$. We decided not to pursue that approach for the following reasons: 1) such solver would have severe convergence problems and robustness issues due to the complexity of the geometry of the devices of interest, 2) the implementation of such solver is, per se, a formidable numerical problem, due to the necessity of generating an appropriate 2D or 3D grid for the devices, and 3) an extremely efficient finite-difference 2D and 3D multi-grid elliptical solver [1] is readily available within the existing code. For these reasons, we further manipulated the energy balance equation in order to re-write it as an elliptical partial differential equation (PDE).
The main issue with such manipulation is the dependency of $k_\mu$ from both the temperature and the position. We therefore assume that the thermal conductivity can be represented as a piecewise function of the position, in other words, $k_{\mu,C}(T)$ is a function of the temperature but is constant within each cell $C$ of the finite differences grid. We therefore express this restricted position dependency with the index $C$ rather than via a full functional dependence on the position vector $r$. Furthermore, we define an equivalent temperature $\theta_{\mu,C}(T)$ through the well-known Kirchhoff transformation [2]:

$$\theta_{\mu,C}(T) = T_0 + \frac{1}{k_{\mu,C}(T_0)} \int_{T_0}^{T} k_{\mu,C}(\tau) d\tau,$$

(3)

Where $T_0$ is a reference temperature. This allows us to rewrite the energy balance equation as follows:

$$\nabla^2 \theta_{\mu,C} = - \frac{p_\mu(r)}{k_{\mu,C}(T_0)},$$

(4)

which is a linear Poisson equation for the transformed temperature $\theta_{\mu,C}(T)$. This equation can be easily and efficiently solved with the existing multi-grid 2D and 3D solver available in the existing code.

However, a crucial issue is related to the Kirchhoff transformation above, which needs further discussion. Indeed, while $k_{\mu,C}(T)$ is safely assumed not to change with position within a finite difference cell, the different materials present in a device layout are expected to have a different functional dependency on temperature. In other words, we need to assign different functions $k_{\mu,C}(T)$ in different cells. To achieve that, the temperature and its normal derivative need to be continuous across cell boundaries even when the thermal conductivity is not. This implies that, in order to have a unique solution of the linearized (elliptical) form of the energy balance equation, the Kirchhoff transformation must be invariant for the continuity conditions expressed above. The invariance means that if $T$ and its normal derivative are continuous across cell boundaries, the equivalent temperature $\theta_{\mu,C}$ must have the same property. Unfortunately, while the Kirchhoff transformation is invariant for the continuity of the normal derivative of $T$, it is not invariant for the condition of continuity of $T$ across a boundary [3]. This means that if $T$ is continuous across a cell boundary, there is no guarantee that $\theta_{\mu,C}(T)$ is also continuous. This shouldn’t be a surprise: indeed the application of the Kirchhoff transformation does not eliminate the non-linearity from the equation; it just moves it from inside the cell to the boundary of the cell [4]. In passing, one should note that if we had chosen to use a non-linear finite element approach, we would have moved the issue of non-linearity from the analytical domain to the numerical one because of the convergence issues of finite elements nonlinear solvers in the presence of complex boundaries. So the overall level of difficulty would not have changed.
In order to address this rather important limitation we note that for many different semiconductor materials, we can express their thermal conductivity $k_{\mu,C}(T)$ through the same functional form $f(T)$:

$$k_{\mu,C}(T) = \alpha_C \cdot f(T),$$

where $\alpha_C$ is a constant that can vary in each cell. It is easy to verify that such condition translates in the needed invariance of the temperature continuity across boundaries. Fortunately, it turns out that this rather restrictive condition on the functional dependency can be successfully adopted for most semiconductors without significant loss of generality.

To ensure the approach is not flawed, we have solved the equation (4) with the forcing function set to zero. In this case we expect the solution to simply be linear in between two boundary “thermal contacts” where the temperature is set at a specific value. To verify our approach, a thermal GaAs resistor has been simulated where initially a right contact is set at 350 K, while a left contact and all intermediate cells are set to 300 K. The temperature map is first transformed into a Kirchhoff apparent temperature, then the PDE is solved, and finally the inverse transformation is applied to return to the normal temperature.

The resulting temperature map is shown in Figure 1. The result shows that our final solution is indeed linear. Furthermore, it also confirms our implementations for the thermal conductivity and Kirchhoff transformations.

![Figure 1: Final Temperature as a Function of the Position within the Device where the left Contact is set at 300K and the right at 350 K](image)

*The solution is linear in the region within the contacts as expected in the absence of a forcing function.*

### 3.2 Coupled Electron-Phonon Dynamics for Electro-Thermal Simulation

A second approach is presented here, which is not based on the numerical solution of the EBE, but rather on the stochastic particle-based solution of the solution of the BTE for phonons. In order to accurately simulate the heating, heat distribution and eventual thermal failure of semiconductor devices, it is critical to correctly understand and model at the microscopic scale
the complex electro-thermal coupling effects, the resistive-thermal relationship in the devices, and their thermal breakdown property. Existing models cannot depict electro-thermal phenomena accurately because the phonon heating effect is not fully taken into account. However, it is believed that the phonon dynamics plays a main role in self-heating of devices, especially for sub-100nm and nano-structures where the ballistic-phonon scattering causes micro-over-heating as critical device dimensions become comparable to phonon mean free path. To fully account for the electro-thermal behavior, we propose to include full particle-based phonon dynamics into the electro-thermal device-modeling framework.

In semiconductors, while most of the heat generation is due to the interaction between electrons and phonons, the heat conduction is mainly due to the motion of phonons. When the device dimensions are larger than the phonon mean free path, heat transport occurs in a diffusive regime that is effectively described by the EBE. However, when the device size is comparable with the effective mean free path of phonons, the heat transport regime becomes ballistic, phonons undergo fewer scattering events, and the thermal transport occurs in non-equilibrium conditions. Being, for example, the phonon effective mean free path in silicon approximately 300nm [5], it is obvious that the heat flux in many current devices occurs within this non-equilibrium ballistic limit and that the EBE is inadequate to modeling the heat flux. To address this problem and in full analogy with the procedure used by the CMC approach in simulating charge transport, we will propose a thrust to model the phonon dynamics by the stochastic solution of the phonon Boltzmann equation:

$$\frac{\partial f_p (r, q, t)}{\partial t} = -V_p \cdot \nabla f_p (r, q, t) + \left[ \frac{\partial f_p (r, q, t)}{\partial t} \right]_{coll}$$

where $f_p$ is the phonon distribution function expressed in terms of position $r$, momentum $q$ and time, $V_p$ is the phonon group velocity, and the second term on the right side of the equation is a collisional term accounting for the interactions of the phonons with other phonons, boundary conditions and electrons. The proposed technique is inspired by the work of Mazdumer and Majumdar [6], and is based on the representation of the phase-space evolution of a significant portion of the phonon population and on its coupling with the charge carrier particles (electron and holes). In analogy with what is done with charge carriers, the dynamics of the phonons will be modeled as a sequence of ballistic free-flights interrupted by stochastic momentum changing collisions due to phonon-phonon, phonon-electron, phonon-impurity, and phonon-boundary scattering. The phonon scattering probabilities of three-phonon Normal and Umklapp scattering [7,8], as well as the phonon-impurity scattering, will be included in the fully discretized Brillouin Zone (BZ) of the CMC simulative framework by using perturbative analysis. In other words, we propose a stochastic self-consistent solution of both phonon and electron BTEs. A crucial aspect of this work will be the model used for the phonon-electron interaction. Within the proposed framework, the BTE is expressed for electrons as follows:

$$\frac{\partial f_e (r, k, t)}{\partial t} = -V_e \cdot \nabla f_e (r, k, t) - \mathbf{F} \cdot \nabla_k f_e (r, k, t) + \left[ \frac{\partial f_e (r, k, t)}{\partial t} \right]_{coll}$$

Approved for public release; distribution is unlimited.
where the index $e$ identifies quantities related to charge carriers. Note the drift term $\mathbf{F} \cdot \nabla_k f_e(\mathbf{r}, \mathbf{k}, t)$, not present in the phonon BTE (equation 6), which accounts for the long-range electrostatic interaction expressed by the electrostatic force $\mathbf{F}$. The collisional term for charge carriers $\left[ \frac{\partial f_e(\mathbf{r}, \mathbf{k}, t)}{\partial t} \right]_{\text{coll}}$ accounts for the momentum-changing interactions of electrons and holes with their environment, including phonons. In particular, within the CMC framework, nonpolar phonon scattering is handled via perturbative approach in order to obtain the scattering rate from a region of band $n$ centered in the point $\mathbf{k}$ to a region $\Omega_{k'}$ in band $n'$ centered around the point $\mathbf{k'}$ in momentum space:

$$P_{nn'}(\mathbf{k}_n, \Omega_{k'}) = \frac{\pi}{\rho \omega_{nq}} \left| \Delta_{n,n'}(\mathbf{q}) \right|^2 \omega_{nq} \left| I(n, n', \mathbf{k}, \mathbf{k'}) \right|^2 D_n(E', \Omega_{k'})(n_{nq} + \frac{1}{2} \pm \frac{1}{2}) \quad (8)$$

where $\rho$ is the semiconductor density, $\mathbf{q}$ the phonon wave vector, $\Delta_{n,n'}(\mathbf{q})$ is the nonpolar matrix element, $I$ is the overlap integral between Bloch states, $D_n(E', \Omega_{k'})$ is the electronic density of states in $\Omega_{k'}$ at the final (after scattering) energy $E'$ in band $n'$, and $n_{nq}$ is the phonon occupation number usually computed at equilibrium and at the lattice temperature.

While this rate could be used “as is” for the calculations of the Joule term for the heat diffusion equation, a crucial modification is necessary in the non-equilibrium framework of the proposed particle representation for both phonons and charge carriers. The equilibrium phonon number will be replaced by the local value of the phonon population at the position where the scattering occurs. This value is supplied by the part of the algorithm that models the phonon dynamics as a solution of the phonon BTE. In addition, in this case, the scattering rates $P_{nn'}(\mathbf{k}_n, \Omega_{k'})$ will be tabulated for the estimated maximum value of $n_{nq}$ and a rejection technique has been successfully implemented to handle storage issues.

As a consequence of the scattering, a phonon creation/destruction mechanism has been also implemented for the three-phonon scattering mechanism.

It should be stressed that the inclusion of the full-band representation for both electron and phonon dispersions is essential for a correct quantitative evaluation of the energy exchange between the particles in the system.

### 3.2.1 Thermal Conductivity

A device similar to the one depicted in Figure 1 has been simulated by representing phonons as particles in the presence of a thermal gradient while considering both optical and acoustic modes. The thermal contacts of the device have been set to temperatures $T_n = 310$ and $T_i = 29$. 

Approved for public release; distribution is unlimited.
When the system reached steady state, the effective temperature was extracted from the local energy density by inverting the following relation:

\[
E_{\text{crystal}}(T, V) = \frac{v}{8\pi^3} \sum_p \int_{q \in BZ} \hbar \omega_{q,p} ((n_{q,p}) + \frac{1}{2}) d^3q = E_0 + E_{\text{pho}}
\] (9)

The results of the simulation are shown in Figures 2 and 3. The scattering algorithm successfully reproduces the expected linear behavior in the diffusive regime. Moreover, the thermal conductivity matches the experimental result along the device [9].

![Figure 2: Effective Temperature in a Silicon Thermal Resistor as Function of the Position](image)

Using the new scattering table with multiple Grüneisen parameters successfully reproduces the expected linear behavior in the diffusive regime.

![Figure 3: Thermal Conductivity as a Function of the Position](image)

The Monte-Carlo simulation is able to reproduce the thermal conductivity seen in experiment.

### 3.2.2 Isotope Scattering

The isotopes are variants of an element with the same number of protons and different number of neutrons, therefore different atomic weight. All the isotopes show (almost) identical chemical behavior, crystal structure, and electronic structure. However, the mass difference due to the presence of isotopes disrupts the periodicity of the crystal potential and results in a substantial rate of phonon scattering. The overall effect is a major reduction of the phonon thermal conductivity. Indeed, the thermal conductivity of the natural crystal \((k_{\text{nat}})\) and the isotopically pure crystal \((k_{\text{iso}})\) can differ up to one order of magnitude at low temperatures [10]. At room
temperature the effect is more ambiguous, even for a material studied as extensively as silicon the reported $k_{iso}$ range is between 110% \cite{11,12} and 160% \cite{13,14,15} of the value of $k_{nat}$.

Our CMC code contains an isotope scattering model derived by Srivastava \cite{16}:

$$P_{q\rightarrow\Omega q'} = \frac{V}{8\pi^3} \int_{\Omega q'} \frac{\pi \Gamma}{2\hbar N_0} n(n' + 1)EE'(e_q^* \cdot e_{q'})^2 \delta(E - E'),$$

where $\Omega q'$ is the reciprocal volume around the phonon state $q'$, $n$ is the Bose-Einstein distribution, $V$ is the crystal volume, $N_0$ is the number of Wigner-Seitz cell in the crystal, $e$ is the polarization vector, $E$ is the phonon energy, and $\Gamma$ is defined as:

$$\Gamma = \sum_i f_i \left( \frac{\Delta M_i}{\bar{M}} \right),$$

where $f_i$ is the fraction of atoms with mass $M_i$, and $\bar{M}$ the average mass. $E'$ can be assumed slowly varying over the whole momentum space, and therefore as constant in the small volume $\Omega q'$, while $(e_q^* \cdot e_{q'})^2$ can be approximated by the value 1/3. These assumptions reduce the relation for the probability of scattering in equation (10) to the following expression:

$$P_{q\rightarrow\Omega q'} = \frac{Vc\Gamma}{6\hbar} n(n' + 1)EE' \frac{1}{8\pi^3} \int_{\Omega q'} \delta(E - E')$$

$$= \frac{Vc\Gamma}{6\hbar} n(n' + 1)EE'Dos_{\Omega q'}(E),$$

where $V_c$ is the Wigner-Seitz cell volume and $Dos_{\Omega q'}$ is the density of phonon states with energy $E$ within the reciprocal volume $\Omega q'$.

Figure 4 shows a comparison of the thermal conductivity as obtained from our CMC simulation (symbols) and from experiment (lines) \cite{14}. Natural silicon ($Si_{nat}$) is composed by the three isotopes $Si^{28}$ (92.2%), $Si^{29}$ (4.6%), and $Si^{32}$ (3.1%), and has been simulated by activating the isotope scattering mechanism, while the pure mono-isotopic $Si^{28}$ has been simulated without isotope scattering. In both cases the agreement between simulation and measurement is excellent. Further analysis at higher temperature is being currently performed.
The isotope scattering plays a major role in the physics of thermal conductivity, and cannot be neglected while simulating real semiconductor devices. Moreover, our simulation technique offers a theoretical tool to investigate how different scattering mechanisms influence the thermal conductivity.

Some final considerations should be made concerning the performance of the CMC simulation tool. The method proposed here is based on the full self-consistent particle-based simulation of the charge and heat transport due to the combined dynamics of electrons and phonons. The time resolution of the algorithm is expressed in fractions of femtoseconds, while the timescale of the heat transfer from the electrons to the lattice is measured in fractions of picosecond [17] and the relaxation time of the phonon-driven thermal transport is approximately 80 picoseconds [18]. This means that a simulation should be carried out for about 100 picoseconds in order to reach the electro-thermal steady state and successfully extract the device parameters to be used by the fast circuitual simulators. By coupling the particle-based electron dynamics with the flux-based solution of EBE described in the previous section, we will be capable of efficiently reaching the electro-thermal steady-state condition. Subsequently, performing a particle-based simulation of both electron and phonons will continue the simulation. This multi-scale approach will allow studying the transient, non-linear electro-thermal properties of high-power and high frequency devices.

Figure 4: Comparison between Measured (lines) [14] and Simulated (symbols) Thermal Conductivity as a Function of Temperature in Multi-isotopic Natural (Si$^{nat}$) and Pure (Si$^{28}$) Silicon
4. RESULTS AND DISCUSSION

4.1 Energy Balance Equation for Phonons

We decided to obtain the flux-based analysis of heat generation and transport by solving the EBE for phonons rather than the HTE. This approach would allow a more accurate temperature map by supplying a separate solution for each phonon mode (or group of modes), and is more suited to self-consistent coupling with the electron dynamics. The approach we developed starts from the phonon BTE to obtain the energy balance equation for each phonon mode $\mu$:

$$\frac{\partial W_\mu}{\partial t} = -\nabla \cdot F_W + \frac{\partial W_\mu}{\partial t} |_{e-p} + \frac{\partial W_\mu}{\partial t} |_{p-p}$$

(1)

where $W_\mu (r, t) = \frac{1}{\Omega} \sum_k E_\mu (k) f_\mu (r, k, t)$ is the ensemble energy in the volume $\Omega$ of the reciprocal space, $F_\mu (r, t) = \frac{1}{\Omega} \sum_k v(k) E_\mu (k) f_\mu (r, k, t)$ is the energy flux, and the two partial derivatives of $W_\mu$ in the LHS of the equation represent the rate of change of the ensemble energy due to electron-phonon and phonon-phonon interaction, respectively.

At steady-state, the equation reads:

$$\nabla \cdot (k_\mu (T, r) \nabla T) = -\left( \frac{\partial W_\mu}{\partial t} |_{e-p} + \frac{\partial W_\mu}{\partial t} |_{p-p} \right) = -P_\mu (r).$$

(2)

In this case, $F_\mu (r, t)$ has been approximated with the well-known steady-state relation $F_\mu (r) = -k_\mu (T, r) \nabla T$, where $k_\mu$ is the (scalar) thermal conductivity. The energy balance equation derived above is indeed the classical heat transport equation; the novelty of the proposed approach is in the way the forcing function is computed within the Cellular Monte Carlo framework as well as the fact that one of such equations can be solved for each phonon mode.

The component of the forcing function $P_\mu (r) = \frac{\partial W_\mu}{\partial t} |_{e-p} + \frac{\partial W_\mu}{\partial t} |_{p-p}$ due to electron-phonon scattering is represented by the $\frac{\partial W_\mu}{\partial t} |_{e-p}$ term, and can be extracted in real time by recording the phonon emissions and absorptions for each mode $\mu$ after each electron-phonon scattering event. The phonon decay and recombination events contributing to the term $\frac{\partial W_\mu}{\partial t} |_{p-p}$ in the forcing function $P_\mu (r)$ are approximated via relaxation time approximation of the form:

$$\frac{\partial W_\mu}{\partial t} |_{p-p} = C_i \frac{T_j - T_i}{\tau_i}$$

(3)

where $i$ denotes the initial phonon mode, $j$ the final phonon mode, $C$ the volumetric heat capacity, $T$ the temperature, and $\tau$ the relaxation time for the event. For example, in the event of only optical ($o$) phonons decaying to acoustic ($a$) phonons the term would be

$$\frac{\partial W_\mu}{\partial t} |_{p-p} = C_o \frac{T_a - T_o}{\tau_o}$$

(4)
4.1.2 Simulation Results

The approach above has been tested on a 1.29 μm x 0.4 μm GaAs metal-semiconductor field-effect transistor (MESFET), a structure commonly used in the benchmark and validation of the CMC. At this time, a perfect heat sink is attached to the bottom of the device with a constant temperature of 300 K. The assumption was made that the optical modes account for 5% of the total thermal conductivity, and the acoustic modes the remaining 95%. Three cases for the forcing function in equation 2 have been considered; 1) exclusively the electron-phonon scattering, 2) inclusion of an energy decay path from optical to acoustic phonons, and finally 3) inclusion of an energy recombination path from acoustic to optical phonons. The resulting steady-state temperature maps for both the optical and acoustic modes are shown in Figures 5-7.

**Figure 5: Steady-state Temperature Distributions considering only Electron-phonon Scattering**

A peak optical mode temperature of 354.41 K is observed.

First, only the electron-phonon term was considered and steady-state temperature maps obtained. This serves as a useful reference case because it gives a lower bound on the acoustic mode temperature and an upper bound on the optical mode temperature for the other two cases. Another interesting observation from this simulation is that the value of the forcing function for the acoustic modes was comparable to that of the optical modes. This suggests that the large temperature difference is due solely to the discrepancy in the thermal conductivity between the two groups of modes.
In the second set of simulations, optical phonons were allowed to decay into acoustic phonons with a relaxation time of 10 ps as used in [1]. A decrease in the optical mode temperature ($\approx 25$ K) and a corresponding increase in the acoustic mode temperature ($\approx 1.5$ K) is indeed seen in Figure 6. The much larger change in optical mode temperature compared to acoustic mode temperature indicates that after the decay process energy is diffused away rather efficiently, once again due to the larger thermal conductivity.

![Figure 6: Steady-state Temperature Distributions considering both Electron-phonon Scattering and Decay of Optical Phonons into Acoustic Phonons](image)

A decrease in the peak optical temperature to 329.59 K is seen.

In the final set of simulations, energy was allowed to flow from the acoustic modes to the optical modes through a recombination process. This process has not been considered in any literature known to us, but rather was suggested by the phonon dynamics simulations being carried out under the project to occur more frequently than previously thought. A relaxation time for this acoustic-to-optical recombination process was chosen to be 30 ps, which represents a 3x larger likelihood of optical-to-acoustic phonon decay ($\tau_o = 10$ ps) as compared to acoustic-to-optical phonon recombination. We believe that this “two-way street” of energy flow is a more realistic picture than that of the specific energy decay path (electron-optical phonon-acoustic phonon) usually assumed in the literature [1] [2].

Approved for public release; distribution is unlimited.
The resulting steady-state temperature maps for this final case are shown in Figure 7. An increase in the peak optical mode temperature of ≈1.5 K is seen compared to the case of Figure 6, while the acoustic mode temperature is largely unaffected. The increase in optical mode temperature from the second case is expected with the addition of the recombination process, effectively reducing the net energy flow from optical to acoustic modes.

![Figure 7: Steady-state Temperature Distributions considering both Electron-phonon Scattering and Decay of Optical Phonons into Acoustic Phonons](image)

The peak optical temperature is now seen to be 331.12 K.

In summary, we have presented test cases demonstrating a fully coupled electro-thermal CMC approach using a flux-based solution of the energy balance equation for phonons. The resulting temperature maps within the device behave just as one would expect given the respective paths available for energy flow in each case.

The next step in this phase of the project was to attempt to extend the computational domain to include an electrical region (the device of interest) on the sub-micron scale and a region in which only thermal transport is of interest on the scale of many microns. The goal is to obtain a more realistic picture in which the heat sink cannot be attached directly to the device, but rather is located a great distance away.
Section 4.1 References:


4.2 Robustness of the Solver

A potential robustness issue in the manner which insulating regions were handled within the device was discovered since the July quarterly report. As a first approximation, it had been initially assumed that an insulator region was a perfect thermal insulator in which there was no heat generation and the temperature was not allowed to change from the initial value. The manner in which this was done imposed a pseudo-Dirichlet condition in these regions.

The intention was always to remove this constraint and allow for independent thermal properties to be specified; however convergence of the solver was impacted significantly. The iterative methods were modified in order to overcome this, switching to a nested iteration in which an electron-phonon term is found from the Monte Carlo (MC) averages and held constant while allowing the phonon-phonon term to be updated in a self-consistent manner. Once a steady-state solution has been found, a new MC electron-phonon term is found and the procedure repeated until two subsequent steady-state solutions have converged within a given tolerance.

4.2.1 Simulation Results

This modified approach has been tested on the 1.29 μm x 0.4 μm GaAs MESFET from the last report. The assumptions of a perfect heat sink at 300K attached to the bottom and a 95%/5% contribution to the bulk thermal conductivity from the acoustic and optical modes respectively have been retained. The result of a simulation using a 10ps phonon decay time from optical to acoustic modes seen in Figure 8 shows a temperature map very similar to that obtained in the previous approach. The advantage of this are expected to be seen in robustness and applicability to a wider range of device layouts. Work is still to be done on optimization of convergence threshold and simulation time, as well as analysis of alternate, continuous forms of the Kirchhoff Transformation as opposed to the current discrete tabulated values.
4.2.2 Separate Thermal and Electrical Regions

The CMC code has been modified to allow for large regions in which only the heat diffusion equation is solved. This thermal region is defined such that charge carriers are forbidden so that it is electrically inactive, meaning that there is region (the device) in which both electrical and thermal behaviors are found and a region in which only thermal properties are relevant. This allows for an increase in the spacing of the computational grid in the thermal region so that simulation time is not significantly impacted.

As a test, the 1.29 μm x 0.4 μm GaAs MESFET device was embedded inside an overall domain of 11.5 μm x 10.6 μm. The temperature maps inside the electrical region are shown in Figure 9. All previous assumptions remain the same, the only exception being that the 300K heat sink is now located at the bottom of the domain (~10 μm from the bottom of the electrical region).

The acoustic and optical temperatures are both seen to increase, as expected, due to a larger distance from the heat sink. Since we assume that the optical phonons have a smaller contribution to the thermal conductivity, we expect the hot spot to be more localized as compared to the acoustic temperature. This effect is also seen in Figure 9.
This modification should allow for a much more realistic representation of a real device, in which the heat sink cannot be attached directly to the bottom of the device but instead at the bottom of the “package”.

![Steady-state Temperature Distributions in the MESFET including a Region in which only Heat Transport is solved](image)

**Figure 9: Steady-state Temperature Distributions in the MESFET including a Region in which only Heat Transport is solved**

*The total dimensions are 11.5 μm x 10.6 μm*

To summarize, we have presented test cases demonstrating a fully coupled electro-thermal CMC approach using a flux-based solution of the energy balance equation for phonons. A potential robustness issue has been recognized and resolved, allowing for more general device layouts. The capability to quickly solve for a large scale thermal region in addition to the sub-micron device has been added to the CMC and tested, producing results as expected due simply to a greater distance between the thermal generation within the electrical region and the heat sink which dissipates this thermal energy.

The next step in this phase of the project was an optimization analysis of acceptable accuracy, convergence tolerance, and simulation time. Also, since the phonon-phonon term depends on the temperature values which are discrete due to the direct and inverse Kirchhoff transformation, the size of this discretization can lead to potential inaccuracies. However as the discretization is made smaller the memory needed to store the arrays increases, to the order of GB at 1e-6 K. Due to this, continuous algebraic forms for both the transformation and inverse transformation were explored.

Approved for public release; distribution is unlimited.
4.3 Energy Balance Equation Averaging Procedure

As detailed in previous sections, our approach is based on the EBE for phonons:

\[
\frac{\partial W_\mu}{\partial t} = -\nabla \cdot F_W + \frac{\partial W_\mu}{\partial t} \bigg|_{e-p} + \frac{\partial W_\mu}{\partial t} \bigg|_{p-p}
\]  

(1)

where \(W_\mu(r, t) = \frac{1}{\Omega} \sum_k E_\mu(k) f_\mu(r, k, t)\) is the ensemble energy in the volume \(\Omega\) of the reciprocal space, \(F_\mu(r, t) = \frac{1}{\Omega} \sum_k v(k)E_\mu(k) f_\mu(r, k, t)\) is the energy flux, and the two partial derivatives of \(W_\mu\) in the LHS of the equation represent the rate of change of the ensemble energy due to electron-phonon and phonon-phonon interaction, respectively.

At steady-state, the equation reads:

\[
\nabla \cdot \left( k_\mu(T, r) \nabla T \right) = - \left( \frac{\partial W_\mu}{\partial t} \bigg|_{e-p} + \frac{\partial W_\mu}{\partial t} \bigg|_{p-p} \right) = -P_\mu(r)
\]  

(2)

where \(F_\mu(r, t)\) has been approximated with the well-known steady-state relation \(F_\mu(r) = -k_\mu(T, r) \nabla T\), and \(k_\mu\) is the (scalar) thermal conductivity.

In the current CMC we have implemented two choices for the forcing function, i.e., the right hand side (RHS) of (2):

1. \(\vec{j} \cdot \vec{E}\), where the relation current density \(\vec{j} = qnv\) has been used.
2. Directly from the CMC electron-phonon scattering events.

In the first case, each cell in position space needs to keep an average of its electron carrier concentration, carrier velocity \((x, y, z)\), and electric field \((x, y, z)\), while in the latter averages of the electron-acoustic phonon and electron-optical phonon energies must be stored over a pre-defined length of time. This was previously done using an approximation for the exponential moving average (EMA) so that the individual values need not be stored in arrays in memory throughout the simulation.

The EMA was sufficient for the simpler device structures initially considered such as a \(n^+\)-\(n\)-\(n^+\) diode and a simple MESFET structure. However, when considering a more complicated high electron mobility transistor (HEMT) device the error made in the EMA approximation, i.e., the difference between the true average and that given by the EMA, could at times cause delayed convergence. Hence, the decision was made to switch to an exact simple moving average (SMA) in which the values are stored in circular arrays, eliminating the error made in the EMA approximation.
4.3.1 Kirchhoff Transformation

Starting from the LHS of equation (2), the temperature dependence of the thermal conductivity is separated out and lumped into a new “apparent” temperature variable through the use of the Kirchhoff Transformation [1]:

\[ \theta_\mu(T) = T_0 + \frac{1}{k_\mu(T_0)} \int_{T_0}^{T} k_\mu(\tau)d\tau \]  

(3)

where \( T_0 \) is a reference temperature, and \( k_\mu(T_0) \) the thermal conductivity of the phonon mode \( \mu \) at the reference temperature. This pulls the temperature dependence of \( \kappa \) into the newly defined \( \theta_\mu \), allowing us to rewrite the energy balance equation as follows:

\[ \nabla^2 \theta_\mu = -\frac{p_\mu(r)}{k_\mu(T_0)} \]  

(4)

which is a linear Poisson equation for the temperature \( \theta_\mu(T) \). This form of equation can be efficiently solved using preexisting solvers in the CMC. Initially the transformation was tabulated for discrete values at the start of a simulation and stored in memory to be used for both the Kirchhoff Transformation and inverse Kirchhoff Transformation.

A potential issue arises however in the phonon-phonon interaction term in equation 2, which is expressed through a relaxation time approximation:

\[ \frac{\partial w_\mu}{\partial t} \bigg|_{p-p} = C_i \frac{T_i-T_j}{\tau_i} \]  

(5)

Since the phonon-phonon term depends on the temperature values the size of this discretization can lead to potential inaccuracies. However as the discretization is made smaller the memory needed to store the arrays increases, e.g., to the order of GB at 1e-6 K. Due to this, a continuous algebraic form for both the transformation and inverse transformation has been implemented. We choose a power law fit for the temperature dependence of the thermal conductivity to be [2]:

\[ k_\mu(\tau) = \kappa_\mu(T_0) \cdot \left( \frac{T}{T_0} \right)^\alpha \]

Then the Kirchhoff Transformation can be written in algebraic form as:

\[ \theta_\mu = T_0 - \frac{T_0}{\alpha + 1} + \frac{T^{\alpha+1}}{T_0^\alpha(\alpha + 1)} \]

and the corresponding inverse transformation as:

\[ T = \left[ T_0^\alpha(\alpha + 1)\theta + T_0^\alpha(\alpha + 1)\frac{T_0}{\alpha + 1} - T_0^\alpha(\alpha + 1)T_0 \right]^{1/(\alpha+1)} \]
These two relations allow the transformations to be performed in a continuous algebraic manner removing the potential for errors being produced in the forcing function due to the temperature discretization. This algorithmic choice should, in general, improve both accuracy and the rate of convergence while allowing a significant reduction in runtime memory cost.

The resulting temperature map using both the algebraic and the discretely tabulated transformation are shown in Figure 10. The temperatures are very similar in both cases; however the algebraic case achieved convergence in eight self-consistent cycles while the discrete case converged after eleven self-consistent cycles. An even larger improvement is expected in more complicated device structures.

![Figure 10: Comparison of Temperature Maps for a Simple MESFET using the Algebraic Transformation (left) and that Computed and Stored in Discrete Arrays (right)](image)

4.3.2 Possible Effect of Quantum Confinement

The electro-thermal solver was then used to perform a simulation on a GaN/AlGaN HEMT. The resulting temperature map and forcing function using the $\mathbf{J} \cdot \mathbf{E}$ approach is shown in Figure 11, where a clearly unphysical cooling effect is seen.
This behavior was not seen in any of the previous devices considered, including other HEMT structures. Initially, this was thought to be the result of noise due to an issue with the averaging procedure used in the calculation of the forcing function. After attempts to alleviate this possible noise were unsuccessful, other potential causes had to be considered. The cause is currently suspected to be an issue of quantum confinement that is not properly captured in the semi-classical device simulator as the channel region is seen to be only a few nanometers in thickness. The highly localized nature of this effect is seen on the right hand side of Figure 11; where it should be noted that due to sign conventions in the CMC, a positive forcing function value corresponds to cooling and a negative value to heating.

![Figure 11: Example of the Artificial Cooling Effect seen in a GaN/AlGaN HEMT Structure](image)

*On the left is the temperature map and on the right is a map of the forcing function.*

When considering the phonon emission approach discussed in the previous section, an extremely large number of phonons were seen to be emitted in the same “trouble region” that produced this cooling effect using the $\mathbf{J} \cdot \mathbf{E}$ approach. This can still indicate that carriers in the simulation are not truly confined, but instead bouncing back and forth with a very high velocity and creating a large number of phonons due to scattering.

This effect needs to be handled in a robust way such that the CMC will be able to capture potential quantum effects in devices such as this, which is a non-trivial problem. A large part of the upcoming work during the next yearly quarter of the project will be concerning this issue.

In summary, some algorithmic changes have been made regarding the averaging procedure used for the thermal solver as well as the way in which the Kirchhoff transformation is computed. These are simply refinements which should improve the code without having an effect on its functionality. In addition, an interesting cooling effect has been encountered in a particular device structure currently believed to be an artifact due to non-captured effects of quantum confinement.
Section 4.3 References:


4.4 Cooling Effect and Thermal Simulations on a GaN/AlGaN HEMT

As detailed in previous reports, our approach is based on the EBE for phonons:

\[
\frac{\partial W_\mu}{\partial t} = -\nabla \cdot F_\mu + \frac{\partial W_\mu}{\partial t} \bigg|_{e-p} + \frac{\partial W_\mu}{\partial t} \bigg|_{p-p} \tag{1}
\]

where \( W_\mu(r, t) = \frac{1}{\Omega} \sum_k E_\mu(k) f_\mu(r, k, t) \) is the ensemble energy in the volume \( \Omega \) of the reciprocal space, \( F_\mu(r, t) = \frac{1}{\Omega} \sum k \nu(k) E_\mu(k) f_\mu(r, k, t) \) is the energy flux, and the two partial derivatives of \( W_\mu \) in the LHS of the equation represent the rate of change of the ensemble energy due to electron-phonon and phonon-phonon interaction, respectively.

At steady-state, the equation reads:

\[
\nabla \cdot \left( k_\mu(T, r) \nabla T \right) = -\left( \frac{\partial W_\mu}{\partial t} \bigg|_{e-p} + \frac{\partial W_\mu}{\partial t} \bigg|_{p-p} \right) = -P_\mu(r) \tag{2}
\]

where \( F_\mu(r, t) \) has been approximated with the well-known steady-state relation \( F_\mu(r) = -k_\mu(T, r) \nabla T \), and \( k_\mu \) is the (scalar) thermal conductivity.

In the current CMC we have implemented two choices for the forcing function, i.e., the RHS of (2):

1. \( \vec{J} \cdot \vec{E} \), where the relation current density \( \vec{J} = qnv \) has been used.
2. Directly from the CMC electron-phonon scattering events.

In the previous report, an artificial cooling effect was shown in a HEMT structure including a GaN/AlN interface. The polarization charge, including both piezoelectric and pyroelectric, at the GaN/AlN surface is much larger than that of an interface between GaN and an AlxGa1-xN alloy. For the structure considered this value was 0.1 \( C/m^2 \) while in the case of an alloy they are generally close to one order of magnitude smaller. This large value of polarization charge at the interface creates a very large electric field in the direction normal to the surface, in this case the \( y \)-direction, which is the component of the forcing function seen to be responsible for the cooling effect.
Due to this, the decision was made to simulate another GaN/AlGaN HEMT device that is a well understood layout used as one of the benchmarks for the CMC code. The value of the polarization charge at the GaN/AlGaN interface in this layout is \(0.022 \text{C/m}^2\). The resulting temperature map for a gate bias of both \(V_g = -0.5\) and \(V_g = -1.0 \text{ V}\), with the drain bias \(V_D = 15\text{V}\) in both cases is shown in Figure 12.

![Temperature map for different gate biases](image)

**Figure 12:** Electro-thermal Results on a GaN/AlGaN HEMT using the \(\vec{J} \cdot \vec{E}\) Forcing Function

*A higher temperature is seen in the case of \(V_g = -0.5\), owing to the larger current density for this bias point.*

The resulting temperature maps are as should be expected, and the cooling effect is not observed. Figure 13 shows the maximum temperature seen in the device as a function of the applied drain bias. The peak temperature is seen to increase with the applied bias, which is the expected result.

![Temperature vs drain bias](image)

**Figure 13:** Electro-thermal Results on a GaN/AlGaN HEMT using the \(\vec{J} \cdot \vec{E}\) Forcing Function

*Left: Current density vs drain bias. Right: maximum temperature vs drain bias. Peak temperature in the device is seen to increase with drain bias as should be expected.*
These results suggest that the previous cooling effect is due to the large polarization charge present in that particular layout, or potentially due to the manner in which the polarization charge is handled within the CMC code.

In addition, other approaches are being taken to better understand the cause of this cooling phenomena and to validate results. One of these is the implementation of a third forcing function for the energy balance equation based on the work of Majumdar [1], in which the temperature equations for the optical and acoustic phonon modes are given by:

\[
C_{op} \frac{\partial T_{op}}{\partial t} = \frac{3n k_B}{2} \left( \frac{T_e - T_{op}}{\tau_{el-op}} \right) + \frac{nm^*v^2}{2\tau_{el-op}} - C_{op} \frac{T_{op} - T_{ac}}{\tau_{op-ac}}
\]

\[
C_{ac} \frac{\partial T_{ac}}{\partial t} = \nabla \cdot (\kappa_{ac} \nabla T_{ac}) + \frac{3n k_B}{2} \left( \frac{T_e - T_{ac}}{\tau_{el-ac}} \right) + C_{op} \frac{T_{op} - T_{ac}}{\tau_{op-ac}}
\]

A second approach is to cross-check our results using the CMC code with those of the commercial Synopsys® Sentaurus simulator. Sentaurus has the capability to use a hydrodynamic model for heat generation which should be a good comparison to our approach in the CMC. In addition, this will allow the implementation of polarization charges in the CMC to be validated in order to eliminate potential causes for the cooling phenomenon.

In summary, a well understood GaN/AlGaN HEMT structure has been simulated with the CMC in which no cooling effect has been observed, and the resulting temperature maps show the expected behavior. Work is currently underway investigating the effects, and validity thereof, of the polarization charge within the CMC framework.

The novel approach, as to our knowledge, in the current work being done is that of the phonon emission forcing function. In this approach, energy conservation is used to track the energy either gained or lost by an electron due to a scattering event which creates either an optical or acoustic phonon along with the energy exchanged due to phonon decay (optical to acoustic) or recombination (acoustic to optical). The sum of this energy, within each cell defined in the simulation, is then used as the forcing function which generates a temperature map for both the acoustic and optical phonons. These temperatures correspond to the population of phonons given by the Bose-Einstein distribution. Electro-thermal simulations using this approach have been performed on the same HEMT layout previously used, and the resulting temperature maps are shown in Figure 14 and Figure 15 for different biasing.
Figure 14: Electro-thermal Results on a GaN/AlGaN HEMT using the Phonon Emission Forcing Function

At low biasing the temperatures are seen to be well above the expected near equilibrium result.
Figure 15: Electro-thermal Results on a GaN/AlGaN HEMT using the Phonon Emission Forcing Function

At a drain to source bias of 8 V, the temperatures increase only about 10 K from the initial low bias case.

It’s seen in Figure 14 that the initial acoustic temperature, which should correspond closely to the overall lattice temperature, is elevated at low bias above the near equilibrium (300K) condition that would be expected. To investigate this, bulk simulations were performed to check the electron energy at low electric fields. It was discovered that for the high doping levels present in the layout, the electron energies (or related electron temperatures) were elevated due to the size of the gridding being used in k-space. The grid in k-space is used to generate scattering tables, which are then used in a bulk material simulation. The outputs of this simulation show quantities such as the average electron energy and average drift velocity as a function of the applied electric field. The bulk simulation results are shown in Figure 16 for the average electron energy and the corresponding electron temperature.
In addition, the temperature does not increase as much as expected with bias as seen in Figure 15. This, however, can be tuned somewhat by adjusting the optical-acoustic phonon decay lifetime in the simulations which can also be seen in Figure 16 by comparison of the resulting temperatures for the 2 ps decay time and 1 ps decay time respectively.

Figure 16: Electron Energies and associated Temperatures at Low Electric Fields in Bulk Simulations

Ideally, the electron energy at very low fields should be nearly 40 mE with corresponding to a temperature near 300 K. Also, the energies are seen to decrease with a finer grid.

In summary, the instability in the phonon emission forcing function has been overcome but other issues have arisen in its place. Namely, in this particular layout, the low bias temperatures are seen to be far out of equilibrium, and the increase in temperature with bias is much smaller than anticipated. However, the low bias temperature is likely an effect of the gridding used in the computation of the material’s scattering rates.
Section 4.4 References:


4.5 Low Bias Simulations on a GaN/AlGaN HEMT

While the bulk of the work in the middle of the project was focused on the integration of the phonon dynamics code with this energy balance approach, some important results have been obtained as a continuation of the previous approaches.

In particular, the temperature maps inside of the GaN/AlGaN HEMT structure were seen to be considerably higher at low bias than what would be expected in a near-equilibrium situation. It was hypothesized that the elevated temperature of owed to the k-space gridding used in the generation of the material scattering tables, which in turn caused an elevated electron temperature at low electric fields. To this end, simulations were performed using two different tables; one with a coarser k-space grid resulting in a low-field electron temperature near 860 K and one with a finer k-space grid resulting in a low-field electron temperature near 355 K as seen in Figure 17.

![Figure 17: Electron Energies and associated Temperatures at Low Electric Fields in Bulk Simulations](image)

*Ideally, the electron energy at very low fields should be nearly 40 mEv corresponding to a temperature near 300 K. However, the memory needed to store the scattering tables at runtime quickly increases with finer gridding. The device simulation using the finer k-space grid, and lower temperature, shown here used just over 50 GB of RAM.*
The elevated temperatures were indeed due to the k-space gridding and the consequential error in electron energy at low bias (and electric field), then one would expect to see a decrease in the low-biased device’s temperature map between the two tables. However, the temperature maps between the two are virtually identical, as seen in Figure 18, with only a 1K difference in the peak acoustic temperature value.

Figure 18: Acoustic and Optical Temperature Maps for a Device with a Gate Bias of $V_G = -0.5$ V and a Zero Drain Bias

The coarse grid is shown on the left, while the finer grid maps are on the right. The two cases are seen to be virtually identical.

At this time, the focus has been on the relationship between the x-electric fields (the direction of electron transport in the layout) with the electron energies and resulting device temperature maps. There is a large electric field present in the y-direction (perpendicular to the GaN/AlGaN interface) even at low bias due to the large surface charge expected to be present, and thus included in the simulation, at the interface due to polarization effects. It is this field that causes electron confinement in the direction perpendicular to the interface, and it’s reasonable to suspect
this to be the culprit for the elevated temperatures at low bias conditions seen, after having ruled out the possibility of a simple gridding issue.

4.5.1 High Bias Simulations on a GaN/AlGaN HEMT

The value in a device layout such as this is in its potential high power applications. With this in mind, simulations were also performed at a drain-to-source bias of 30V to compare the result of the two forcing functions;

1. \( \vec{J} \cdot \vec{E} \), where the relation current density \( \vec{J} = qnv \) has been used.
2. Directly from the CMC electron-phonon scattering events.

A previous result using only the x-components of \( \vec{J} \cdot \vec{E} \) is shown in Figure 19. The characteristics to take note of are specifically the peak temperature and the location of the hot spot. It is known that in the use of this forcing function the hot spot is expected to be located near the edge of the gate as seen because of the high electric field at this location.

![Figure 19: Previously seen Temperature Map obtained using only the x-components of the Velocity and Electric Field to compute the Forcing Function](image)

The peak temperature is near 540 K and localized near the gate of the device.

In Figure 20, the 2D optical and acoustic temperature maps resulting from the forcing function computed directly from the phonon absorption and emission scattering events for comparison.

The hot spot is seen to shift significantly over towards the drain with respect to its location in the \( \vec{J} \cdot \vec{E} \) case, but it’s well known that the latter “anticipates” the hot spots since it does not account for the energy relaxation time of electrons. The concerning discrepancies between the two cases are the magnitude of the difference in the peak temperature (540K to 390K), and the spreading...
behavior of the temperature in the two cases. Whereas the hot spot is seen to be very localized near the gate in the former, it is generated in the channel region (as seen in the optical temperature map, which has a very low thermal conductivity) and spreads upwards such that the peak temperature is seen at the location furthest from the heat sink on the bottom of the device.

Figure 20: Temperature Map using the Energies computed directly from the Phonon Absorption and Emission Scattering Events in the CMC

The peak temperature seen to be much lower and the hot spot shifted over towards the drain. These effects are both expected, however.

In summary, the elevated near-equilibrium temperature maps have been explored, and although the cause has not been isolated, as of yet, the possibility of it arising from a k-space gridding issue has been ruled out. In addition, the high power state of a GaN/AlGaN HEMT has been simulated and compared using the two forcing functions and differences between them to be further explored highlighted.

4.6 Merging of Separate Code Versions

There have been two thermal simulation packages under development independently; the first being the coupled electro-thermal solver based on solving an energy balance equation to obtain near steady-state temperature maps, while the second is purely a phonon dynamics code. The
The ultimate goal is to use the energy balance equation solver as an input for the phonon dynamics code, in the aim of providing an accurate initial condition for the temperatures within the device.

The merging of the two has been largely successful, with the exception of one remaining issue being explored which causes a difference in the results obtained with the phonon emission forcing function. Repeating the summary from previous reports as a reminder, the EBE for phonons has the form:

$$\frac{\partial W_\mu}{\partial t} = -\nabla \cdot F_W + \left. \frac{\partial W_\mu}{\partial t} \right|_{e-p} + \left. \frac{\partial W_\mu}{\partial t} \right|_{p-p}$$

(1)

where $W_\mu (\mathbf{r}, t) = \frac{1}{\Omega} \sum_k E_\mu (\mathbf{k}) f_\mu (\mathbf{r}, \mathbf{k}, t)$ is the ensemble energy in the volume $\Omega$ of the reciprocal space, $F_\mu (\mathbf{r}, t) = \frac{1}{\Omega} \sum_k v(\mathbf{k}) E_\mu (\mathbf{k}) f_\mu (\mathbf{r}, \mathbf{k}, t)$ is the energy flux, and the two partial derivatives of $W_\mu$ in the LHS of the equation represent the rate of change of the ensemble energy due to electron-phonon and phonon-phonon interaction, respectively.

At steady-state, the equation reads:

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

(2)

where $F_\mu (\mathbf{r}, t)$ has been approximated with the well-known steady-state relation $F_\mu (\mathbf{r}) = -k_\mu (T, \mathbf{r}) \nabla T$, and $k_\mu$ is the (scalar) thermal conductivity.

When using the phonon emission forcing function, the RHS of equation (2), $-P_\mu (\mathbf{r})$, is computed directly from the CMC electron-phonon scattering events. For each scattering event in a cell, it is first decided whether the mechanism absorbs/emits either optical or acoustic phonons, then the difference in energy is kept track of accordingly for each and the values averaged for a long enough period of time to eliminate noise. The average value is then used to compute the electron-phonon term of the forcing function throughout the device.

The one issue remaining is a discrepancy between the results obtained using the pre-merging EBE solver and those seen after, shown in Figure 21. The forcing function values are decreased significantly, and hence the thermal solution differs. However, all electrical results appear to match those from the tests used to benchmark the code which are extremely reliable. The source for this is very likely simply an, as of yet, elusive bug within the code. Due to the large discrepancy between the two (nearly two orders of magnitude), it’s very unlikely to have a physical cause.
Figure 21: Difference in tabulated Electron-phonon Term of the Forcing Function for the Energy Balance Equation in the same GaAs MESFET Structure
Acoustic and optical results for the merged code are shown on the top, while the previous results are on the bottom set to the same scale.

4.6.1 Error Handling and Robustness Improvements

It was decided to undertake the work now to improve error handling and robustness in order to save a much larger amount of time in the future debugging device simulations, and in the case when “users” are running simulations without knowledge of the underlying code itself.

Examples include bounds checking on inputs, e.g., thermal conductivity or heat capacity, and bounds checking on the results of the thermal solver. The goal is to recognize obvious errors as soon as possible, and exit with an error message rather than leaving it fully up to the user to verify results and search for the possible issues.
For example, thermal conductivity values for the materials of interest range from \(10^0\) for an insulator to \(10^3\) in the case of diamond. Any input values significantly outside of this range are very likely mistakes and should be flagged as such. The time savings due to this error checking are extremely significant as simulation time for a complicated device layout such as the HEMTs of interest in this project can reach days or weeks.

We have focused on merging two separate versions of the code together as well as improving usability and robustness of the CMC code. The aim of this is to vastly reduce debugging time throughout the lifetime of the code itself. As soon as the issue with the forcing function was resolved, work began on full characterization of devices.

### 4.7 Fully Coupled Electrothermal Approach

Repeating the summary from previous reports as a reminder, the EBE for phonons has the form:

\[
\frac{\partial W_\mu}{\delta t} = -\nabla \cdot F_W + \left. \frac{\partial W_\mu}{\delta t} \right|_{e-p} + \left. \frac{\partial W_\mu}{\delta t} \right|_{p-p}
\]  

(1)

where \(W_\mu(r, t) = \frac{1}{\Omega} \sum_k E_\mu(k) f_\mu(r, k, t)\) is the ensemble energy in the volume \(\Omega\) of the reciprocal space, \(F_\mu(r, t) = \frac{1}{\Omega} \sum_k v(k) E_\mu(k) f_\mu(r, k, t)\) is the energy flux, and the two partial derivatives of \(W_\mu\) in the LHS of the equation represent the rate of change of the ensemble energy due to electron-phonon and phonon-phonon interaction, respectively.

At steady-state, the equation reads:

\[
\nabla \cdot \left( k_\mu(T, r) \nabla T \right) = -\left( \left. \frac{\partial W_\mu}{\delta t} \right|_{e-p} + \left. \frac{\partial W_\mu}{\delta t} \right|_{p-p} \right) = -P_\mu(r)
\]  

(2)

where \(F_\mu(r, t)\) has been approximated with the well-known steady-state relation \(F_\mu(r) = -k_\mu(T, r) \nabla T\), and \(k_\mu\) is the (scalar) thermal conductivity.

When using the phonon emission forcing function, the RHS of equation (2), \(-P_\mu(r)\), is computed directly from the CMC electron-phonon scattering events. For each scattering event in a cell, it is first decided whether the mechanism absorbs/emits either optical or acoustic phonons, then the difference in energy is kept track of accordingly for each and the values averaged for a long enough period of time to eliminate noise. The average value is then used to compute the electron-phonon term of the forcing function throughout the device. A relaxation time approximation is used for the phonon-phonon term in (2), and the resulting temperature map, for both optical and acoustic modes respectively, is obtained. The temperature is then updated in each computational cell, and used to adjust the scattering rates based on a rejection algorithm. Hence, the result is a fully coupled electrothermal simulation.
4.7.1 Electrothermal Simulations of a GaN HEMT

The device layout studied here was reported by Altuntas et al [1], and is shown in Figure 22. Effects due to both thread dislocation densities and polarization at the GaN/AlN interface are captured in the simulation. The resulting steady-state temperature maps are shown in Figure 23, with the peak acoustic temperature ~330 K and the peak optical temperature over 800 K. The biases used in Figure 23 correspond to an On-stress condition of $V_{GS} = 4.0 \, V$ and $V_{DS} = 14 \, V$, where high currents and electric fields are present in the device. It should be noted that the temperature peaks are seen to be shifted to the drain side of the gate rather than being aligned directly with the peak of the electric field.

![Device Layout](image)

**Figure 22:** Device Layout used in Simulations, from Altuntas et al [1]

![Temperature Maps](image)

**Figure 23:** Acoustic and Optical Temperature Maps obtained through the Coupled Electrothermal CMC Code
Figure 24 shows electrical characteristics of the device determined experimentally [1], with the CMC using a 300 K isothermal simulation, and with the fully coupled electrothermal CMC using the phonon emission forcing function. The isothermal simulation is seen to overestimate currents at higher bias points where self-heating effects are most notable, while the fully electrothermal result closely matches experiment.

Figure 24: Drain Current for the Device under Study vs Drain Voltage (left) and Gate Voltage (right)

Obtained experimentally [1], with a 300 K isothermal simulation, and with a fully coupled electrothermal simulation (denoted as SHF).

In addition, to verify that the coupled electrothermal code is working as expected, we expect to see an increase in the scattering rates and a corresponding decrease in velocity due to self-heating effects. This is indeed the case as seen in Figure 25, where the average carrier velocity and total scattering rate is compared between isothermal and electrothermal simulations. The scattering rates shift from roughly $2 \cdot 10^{13}$ s$^{-1}$ up to $4 \cdot 10^{13}$ s$^{-1}$ throughout the device, and the resulting decrease in electron velocity, and hence decrease in current, is seen in Figures 24 and 25.
For the most part, we have focused on electrical characterization of a previously published GaN based high electron mobility transistor. The CMC’s electrothermal solver has been verified to be functioning as expected through the increase in scattering, which is temperature dependent. The resulting electrical characteristics were seen to capture the decrease in current expected through self-heating, and currents obtained through electrothermal simulation closely match those reported experimentally.

Section 4.7 References:


4.8 Rejection Algorithm for Scattering

The capabilities of the original CMC algorithm [1] have been extended through the use of a rejection algorithm. This approach retains the simulation speed advantages of the CMC scheme while allowing the adaptation of the scattering rates to the real time local conditions. Figure 26 shows the rejection algorithm flow chart. After the CMC triggers an event with probability $P_{CMC}$, a new scattering probability $P_{loc}$ is computed from the value of local parameters. The rejection probability $P_{rej} = P_{loc}/P_{CMC}$ is then used in a stochastic procedure to decide whether the scattering event occurs or is rejected.
The scattering probability after the rejection is

\[ P_{q\rightarrow q'} = P_{\text{CMC}} \times \left( \frac{P_{\text{loc}}}{P_{\text{CMC}}} \right) = P_{\text{loc}} \]  

(1)

where \( q \) is the initial state, and \( q' \) the final state.

**Figure 26: Rejection Algorithm Flow Chart**

4.8.1 Phonon Scattering Rate

Perturbation theory is used to implement both phonon-phonon and phonon-defect interactions. The anharmonic decay and recombination scattering rate based on the Klemens model [2] is given by the following expression:

\[ P_{q\rightarrow q', q''} = \frac{2h}{M^3} \frac{c^2(q, q', q'')}{\omega \omega' \omega''} \delta(\omega, \omega', \omega'') \delta(q, q', q'') F_{\text{loc}} \]  

(2)
where \( F_{\text{loc}} = F(n'_{\text{loc}}, n''_{\text{loc}}) \) is a local population-dependent factor, \( n'_{\text{loc}} \) and \( n''_{\text{loc}} \) are the populations at \( \mathbf{q}' \) and \( \mathbf{q}'' \), respectively, \( M \) is the average mass, \( c^2(\mathbf{q}, \mathbf{q}', \mathbf{q}'') \) is the anharmonic coefficient obtained from the perturbation Hamiltonian, and \( \omega, \omega' \) and \( \omega'' \) are the angular frequencies corresponding to the states \( \mathbf{q}, \mathbf{q}' \), and \( \mathbf{q}'' \), respectively.

The phonon-defect scattering probability based on the work of Srivastava [3] is given by:

\[
P_{\mathbf{q} \rightarrow \mathbf{q}'} = \frac{V_w \Gamma}{6} \omega \omega' (\mathbf{e}_q' \cdot \mathbf{e}_q')^2 \delta(\omega - \omega') (n'_{\text{loc}} + 1)
\]

where \( \mathbf{e} \) is the polarization vector, \( V_w \) is the Wigner-Seitz cell volume, and \( \Gamma \) is an expression that depends on the phonon-defect type.

Computing the local scattering probability, \( P_{\text{loc}} \), requires estimating the local phonon populations \( n'_{\text{loc}} \) and \( n''_{\text{loc}} \) from the ensemble. This task requires defining a sampling volume of the particle position both in real space \( (V_{\text{loc}}) \) and in momentum space \( (\Omega_{\text{loc}}) \), then counting the number of simulated phonons \( \eta(\Omega_{\text{loc}}, V_{\text{loc}}) \) in the volume, and, finally, obtaining the phonon population using the following approximation:

\[
n_{\text{loc}} \approx \frac{\eta(\Omega_{\text{loc}}, V_{\text{loc}})}{N_{\text{state}}} = \eta(\Omega_{\text{loc}}, V_{\text{loc}}) \times \frac{8\pi^3}{\Omega_{\text{loc}} V_{\text{loc}}}
\]

where \( N_{\text{state}} \) is the number of states in the volume.

### 4.8.2 Electron-phonon Interaction

Electrons in a material interact with the atomic lattice through scattering events, each resulting in either the absorption or emission of phonons. From a simulation viewpoint, this is analogous to simple addition or removal of simulated phonons from the domain. However, this made more complex by the fact that, in order to lower the computational costs, a simulated particle represents a relatively arbitrary number of real particles. This number is what we call the weight of the simulated particle, which we will denote as \( W^c \) for charged particles (e.g., electrons) or \( W^p \) for phonons. Hence, when an electron-phonon scattering even occurs in the simulation, it is not one electron scattering with one phonon, but instead \( W^c \) electrons interacting with \( W^p \) phonons.

Ideally, we would use weights such that \( W^c = W^p \) for reasons of statistical accuracy. However, in a typical semiconductor, the density of phonons is orders of magnitude larger than that of charged particles, and this situation becomes unfeasible computationally within reasonable time and memory requirements.
We attempt here to model phonon emission at steady state as a Poisson process [4][5]:

\[ V(k, \lambda^e(q)) = \frac{(\lambda^e(q))^k e^{-\lambda^e(q)}}{k!}, \quad (5) \]

where \( \lambda^e(q) \) is the average number of emitted phonons into state \( q \) during the time interval. At steady state \( \lambda^e(q) \) is simply the expectation value of emission scattering events per unit time:

\[ \lambda^e(q) = \frac{E[\eta^e(q, t + dt) - \eta^e(q, t)]}{dt} = \eta^c \Gamma^{ce}(q) \quad (6) \]

In a typical CMC simulation, the expression for the number of phonons emitted in time \( dt \) due to charge scattering events is similar:

\[ \lambda^{ce}(q) = \frac{E[N^{ce}(q, t + dt) - N^{ce}(q, t)]}{dt} = \frac{\eta^c \Gamma^{ce}(q)}{W^e} = \frac{\lambda^e(q)}{W^e} \quad (7) \]

where the capital \( N \) denoted the number of simulated particles with a weight \( W^c \). Then dividing the expected number of real emitted phonons by the weight used for phonons in the simulation gives the number of phonons to add to the simulation:

\[ \lambda^{pe}(q) = \frac{\lambda^e(q)}{W^p}. \quad (8) \]

The result of a low temperature simulation using this model of a Poisson process is shown in Figure 27. It is clear that there is a large discrepancy between the energy lost by the charge carriers and that gained by the phonons, i.e., energy is not being conserved in the system between the two ensembles despite using relatively close weights for the charge carriers and phonons. Furthermore, this discrepancy accumulates during the transient regime and stays constant once steady-state conditions are reached.

**Figure 27: Total Energy lost by the Charge Ensemble and that gained by the Phonon Ensemble with Phonon Emission/Absorption treated as a Poisson Process**
In fact, this is clearly true looking at equations (7) and (8). Since \( W^c < W^p \), this means that \( \lambda^{ce} > \lambda^{pe} \) and this simple approach must be modified to account for this fact. This modification continued over the course of the project. We have been focused during this section of the work on electron-phonon interaction using a particle-particle simulation approach. The electron-phonon interaction was modeled at steady state using a Poisson process, the result of which produced a discrepancy in the total energy gained by the phonon ensemble and that lost by the charge ensemble.

Section 4.8 References:


4.9 Electron-phonon Interaction

Previously, we attempted to model phonon emission at steady state as a Poisson process [1] [2]:

\[
V(k, \lambda^e(q)) = \frac{\left(\lambda^e(q)\right)^k e^{-\lambda^e(q)}}{k!},
\]

where \( \lambda^e(q) \) is the average number of emitted phonons into state \( q \) during the time interval. In a typical CMC simulation, the expression for the number of phonons emitted in time \( dt \) due to charge scattering events is then:

\[
\lambda^{ce}(q) = \frac{E[N^{ce}(q, t + dt) - N^{ce}(q, t)]}{dt} = \frac{\eta_c}{W^c} \Gamma^{ce}(q) = \frac{\lambda^e(q)}{W^c},
\]

where the capital \( N \) denoted the number of simulated particles with a weight \( W^c \). Then dividing the expected number of real emitted phonons by the weight used for phonons in the simulation gives the number of phonons to add to the simulation:

\[
\lambda^{pe}(q) = \frac{\lambda^e(q)}{W^p}.
\]
However, a difference between the energy lost by the charge ensemble and that gained by the phonon ensemble accumulates through the transient regime, resulting in the loss of conservation of energy. Figure 28, seen in the last chapter, reproduces the plot of this phenomenon here.

Figure 28: Total Energy lost by the Charge Ensemble and that gained by the Phonon Ensemble with Phonon Emission/Absorption treated as a Poisson Process

It turns out that this behavior is due to the statistical variance of the energy gained by the phonon ensemble, $E^{pe}$, being (often, much) larger than that of the energy lost by the charge ensemble, $E^{ce}$. This can be seen by finding the expectation value of $E^{pe}$ and $E^{ce}$, and then finding the variance. The final result for the central variances is:

$$\text{Var}(dE^{ce}(t)) = \sum_{q \in BZ_1} (dt \hbar \omega_q W^c)^2 \lambda^{ce}(q)$$

$$\text{Var}(dE^{pe}(t)) = \sum_{q \in BZ_1} (dt \hbar \omega_q W^p)^2 \lambda^{pe}(q) = \sum_{q \in BZ_1} (dt \hbar \omega_q)^2 W^p W^c \lambda^{ce}(q).$$

Since $W^p > W^c$, it follows that $\text{Var}(dE^{pe}(t)) > \text{Var}(dE^{ce}(t))$; the same argument produces $\text{Var}(dE^{pa}(t)) > \text{Var}(dE^{ca}(t))$ in the case of phonon absorption. The larger noise in the energy gained by the phonon ensemble produces a statistical fluctuation on the transient regime, which then remains unaltered in the steady-state regime, as it should since the average values of $dE^{ce}$ and $dE^{pe}$ are identical at steady-state.

To account for this error without making arbitrary decisions on which state the remove or add a phonon, a model based on a Hawkes [3] process is employed.
4.9.1 Hawkes Process for Phonon Absorption and Emission

Essentially, the Hawkes process introduces memory into the Markovian Poisson process; it modifies the probability of a new event based on the preceding sequence of events. Given a history of events, \( H(t) \), the Hawkes process expresses the conditional probability of observing an event given \( H(t) \):

\[
\lim_{dt \to 0^+} P(M(t + dt) - M(t) = m | H(t)) = \begin{cases} 
  o(dt) & m > 1 \\
  \lambda(t, H(t))dt + o(dt) & m = 1, \\
  1 - \lambda(t, H(t))dt + o(dt) & m = 0
\end{cases} \quad (4)
\]

where \( \lambda(t, H(t)) \) is the intensity of the process and \( M(\tau) \) is the counting measure of the point process, representing the cumulative count of events. In general, \( \lambda(t, H(t)) \) for a Hawkes process can be written as \[4\] \[5\]:

\[
\lambda(t, H) = v \left( \zeta(t) + \sum_{t_i < t} h(t, t_i) \right), \quad (5)
\]

clearly showing the relationship between the history of events, the exciting function, \( h(t) \), and the overall intensity; each past event influences the intensity with the weight \( h \).

We are free to choose the form of the intensities as desired, in this case to minimize the error in conservation of energy between the charge carrier and phonon ensembles. We’ve found that a suitable choice to provide instantaneous response and limit the error in energy between the populations, while considering efficiency of the algorithm, is to use a log-linear model. This corresponds simply to making \( v \) an exponential function in equation (5).

We are able to implement the Hawkes process into the previous work modeling phonon absorption and emission as a Poisson process, and using the rejection technique with probabilities of removing and creating a simulated phonon given by:

\[
P_{\text{rej}}^a = \frac{\lambda_{\text{pa}}(q, t, H)}{\lambda_{\text{qa}}(q, t)} = \frac{W_c}{W_p} e^{-a \int dE(t)} \\
P_{\text{rej}}^e = \frac{\lambda_{\text{pe}}(q, t, H)}{\lambda_{\text{qe}}(q, t)} = \frac{W_c}{W_p} e^{a \int dE(t)} \quad (6)
\]

To test the validity of this approach, we repeat the simulation in Figure 28, but now using the Hawkes model with two different weights; first using \( \frac{W_c}{W_p} = 2.8 \times 10^{-3} \) and then with \( \frac{W_c}{W_p} = 5.6 \times 10^{-3} \).
We see that despite using weight ratios much smaller than in Figure 28, where absorption and emission was treated simply as a Poisson process, the energy lost by the charge carriers agrees much better with that gained by the phonon population. In the case of $\frac{W_c}{W_p} = 5.6 \times 10^{-3}$, the energies agree nearly exactly, and hence conservation of energy is maintained within the synchronous system of charge and phonon ensembles. This means that we now have the capability to realistically simulate electrothermal behavior using a full particle-particle approach.

We are focused on evaluating the electron-phonon interaction using a particle-particle simulation approach. Whereas before the electron-phonon interaction was modeled at steady state using simply a Poisson process, we now take into account a sophisticated bookkeeping scheme for the error in energy using a Hawkes Process. This approach allows us to use a much smaller number of simulated particles, reducing simulation time and memory costs, while at the same time minimizing the error in energy previously seen in our work.

Section 4.9 References:

4.10 Flux-based Approach and Convective Conditions

At steady-state, our flux-based EBE reads:

\[ \nabla \cdot \left( k_\mu(T, r) \nabla T \right) = - \left( \frac{\partial W_\mu}{\partial t} \right)_{e-p} + \left( \frac{\partial W_\mu}{\partial t} \right)_{p-p} = -P_\mu(r), \]  

(1)

where in this work we have been using the Kirchhoff Transformation to linearize the EBE into a linear Poisson-like equation while retaining the information on the temperature dependence of the thermal conductivity.

Boundary conditions imposing a prescribed temperature (Dirichlet) and a prescribed heat flux (Neumann) are linear under the Kirchhoff Transformation [1]. However, a boundary condition based on convective cooling is nonlinear and hence there is an error involved when using the Kirchhoff Transformation [2]. Convective boundary conditions are of the form:

\[ -k \nabla T = h(T - T_\infty), \]  

(2)

where the LHS is the heat flux approaching the boundary, \( T \) is the temperature at the boundary, \( T_\infty \) is the ambient cooling temperature, and \( h \) is the heat transport coefficient representing how efficient the cooling process is. In the case of \( h \to \infty \), we have that \( T = T_\infty \) which is in fact the prescribed temperature Dirichlet condition. We could also use the concept of a thermal resistance, \( R_{th} = \frac{1}{h} \), which intuitively represents a non-perfect cooling process.

This boundary condition can be generalized to an internal interface as well, where it may in fact be most useful in the case of heterostructures. The equation for an interface is simply:

\[ -k \nabla T = h(T_A - T_B), \]  

(3)

where \( T_A \) and \( T_B \) represent the temperature on the respective sides of the interface.

A number of approaches were taken to get around this difficulty, but in the end, an iterative process that abandons the Kirchhoff Transformation altogether has given the best results.

4.10.1 Iterative Approach for including Temperature Dependent Thermal Conductivity

The idea behind this approach comes from the fact that the finite difference grid in the CMC can handle the case of a spatially varying thermal conductivity, i.e., a thermal conductivity that changes in a piecewise fashion from cell to cell.

We first initialize the thermal conductivity values and temperatures in the simulation at 300K, and then run the electrical part of the CMC until we reach (electrical) steady state. At this point, averages are taken on the net scattering energies cell by cell to obtain the forcing function on the RHS of equation (1). We can then solve for a new temperature map, and use these values to
compute a new corresponding thermal conductivity for each cell in the device. We continue doing this in a self-consistent manner until convergence on the temperature is reached everywhere in the simulation.

As an initial test for this approach, we have simulated the two cases published and used the full finite-element solutions, as well as the Kirchhoff transformation solution obtained by [2], as a comparison to the CMC results. Figure 30 shows the temperature along the source plane, i.e., along the top layer of the simulated region, in both plots. The approach of Bagnall et al is to linearize the Kirchhoff Transformation around a temperature that they must estimate beforehand, to minimize the error made due to the non-linearity of the boundary conditions. In our approach no information is needed a priori, and the final solution is reached in a self-consistent manner. Since the transformation is not linearized around an estimated temperature at the convective interface, we should be able to simulate multiple interfaces at once whereas this would obviously be a problem in the approach of Bagnall, et al [2].

![Figure 30: Temperature along the Source Plane, i.e., at the Top of the Device](chart)

500 um x 100 um Silicon region, where the blue line circles are the results from [2], while the black dots are our results using the iterative approach (left) and 500 um x 5 um region of Silicon, where the red line and circle are taken from [2], while the black squares are again our proposed iterative approach in the CMC (right).

On the left is a 500 um x 100 um region with a prescribed heat flux of $1\times10^9 \frac{W}{m^2}$ in the center of the source plane, representative of a device generating heat inside of the much larger overall package. On the right is the same situation, except the y dimension has been reduced to 5 um. In both cases, the ambient temperature ($T_a$) is set to 350 K and the heat transport coefficient ($h$) is $5\times10^6 \frac{W}{m^2K}$. The Kirchhoff transformation result from [2] begins to show a significant error at this size dimension, which we expect to become even worse as the size dimension becomes even smaller. Hence, this suggests the Kirchhoff transformation may not be suitable for simulating a device alone where the size dimensions are likely to be sub-micron. However, the iterative approach shows results still exact with the full non-linear FEA [2]. This suggests our iterative approach is able to accurately capture the temperature dependent effects on size scales more
typical of individual devices, e.g., a single HEMT that has an imperfect surface between two layers of material.

4.10.2 Phonon Dynamics

This procedure is designed to test the codes capability of correctly thermalizing an electron population. An isolated cube of silicon is initialized with the phonon population at 300K, while the electron population is initialized with an energy of 20 meV above the conduction band. This represents a far out of equilibrium electron distribution, as it is a spike at a single energy value.

Figure 31: Electron Distribution initialized at 20 meV above the Conduction Band shown at $t = 0.6 \text{ fs}$ and $t = 300 \text{ fs}$

*The solid dots show the reference distribution, which the electron system is seen to evolve towards.*

Figure 31 compares the electron distribution at $t=0.6 \text{ fs}$ and $t=300 \text{ fs}$ with the reference 300K electron distribution. The figure clearly shows that at $t=0.6\text{fs}$ there has been a negligible amount of scattering and the electrons are still at their initial energy, far away from the equilibrium. At $t = 300 \text{ fs}$, the electron distribution has had time to thermalize, and the picture clearly shows the agreement with the reference distribution at thermal equilibrium.

In order to observe the evolution of the system, the distance between the reference distribution and the simulation distribution is measured via the Jensen-Shannon divergence, which plotted in Figure 32 as a function of time. The plot clearly show the transient regime, lasting 150 fs, followed by the steady state regime.
Figure 32: Jensen-Shannon Divergence of the Correct Reference Distribution and the Simulation Distribution as a Function of Time

The transient regime is the region before approximately $t = 150$ fs, at which point a steady-state has been reached and the divergence no longer decreases.

In particular, the distribution previously plotted, marked by the black square at $t = 300$ fs, indicates that the system has indeed reached steady state. After this point, the divergence is only measuring the noise level of the distribution.

We have proposed a new approach for including the effects of temperature-dependent thermal conductivity while simulating a convective condition. No information is needed a priori to linearize the transformation as elsewhere, and it has the advantages of being able to simulate much smaller size scales as well as multiple convective conditions at once. In addition, the ability of the phonon dynamics code to thermalize a far out of equilibrium electron population to the correct reference distribution was demonstrated.

Section 4.10 References:

4.11 Multi-finger GaN on SiC Package

Following the work of Garven and Calame [1], and Bagnal et al [2], we have reproduced a quarter-model simulation layout for the 22 finger GaN HEMT structure described in [1], [2]. It’s possible to reduce the problem to a quarter-model to simulate only 11 HEMT devices, and half their width, because of the symmetry of the layout. A top view of the simulated region is shown at the top of Figure 33 showing the 50 μm spacing between each 0.5 μm x 75 μm device, as well as a cross-section of the layout at the bottom showing the thickness of each material region. The values for the reference thermal conductivity in each material and the α value in the temperature dependence are taken from [2], $\kappa_{\text{GaN}}(T) = 150 \left( \frac{T}{300} \right)^{-1.4}$ and $\kappa_{\text{SiC}}(T) = 150 \left( \frac{T}{300} \right)^{-1.3}$.

Finally, the values for the heat transport coefficient and ambient temperature are $h = 3.27e5$ and $T_\infty = 293K$ respectively.

![Figure 33: Quarter-model Simulation Layout for the 22 Finger GaN HEMT Structure](image)

*Figure 33: Quarter-model Simulation Layout for the 22 Finger GaN HEMT Structure*

View of the layout structure showing the 0.5 μm x 75 μm dimensions of each individual HEMT, the 50 μm spacing between the fingers, and the size of the overall region simulated (top) and cross-section showing the material thickness for the GaN and SiC regions respectively, as well as the thermal conductivity values used (bottom).

First, we compared the temperature map in the source plane using a single value of $\alpha = -1.3$ for both the GaN and SiC materials to make a direct comparison with the result using the Kirchhoff transformation from [2]. On the left side of Figure 34, we see there are some minor discrepancies between our result and those using the Kirchhoff transformation and obtained with a full FEA, but the agreement is still relatively good. These differences are very likely due to small differences in geometry in our layout which is still to be refined and confirmed to be the cause.
Next, we used the exact $\alpha$ values shown above to compare the result in each case using the CMC. The results are shown on the right side of figure 34, where the red markers represent the temperature using the material specific values, and the solid black line the temperature using a single value of $\alpha = -1.3$. There is an increase in the overall peak temperature of $\approx 2 \degree C$, increasing from 175.44 $\degree C$ to 177.63 $\degree C$ in this layout when using the more accurate value $\alpha_{GaN} = -1.4$.

Figure 34: Comparison between the Iterative Approaches in the CMC with those obtained by Bagnall et al [2]

There are small differences in the peak temperatures at the location of the heat sources as well as in the valleys in between, which are suspected to likely be simply due to small differences in our layout geometry.

4.11.1 Simulation of Interface Resistances

Our iterative approach is also easily extended to simulate multiple interfaces in a single simulation. In the Kirchhoff transformation approach used by [2] the transformation is essentially linearized around an average temperature seen at the convective boundary, which must be known ahead of time. It’s not at all clear that this can generalize to multiple interfaces where a different temperature would be seen at each respective interface.

Following the work of Hickson et al [3], we obtained a finite difference discretization across an interface with a thermal resistance by using a ghost point approach. As our work focuses on steady-state results, our result is shown in the left side of Figure 35 as the solid black line overlaid on the $t = 2$ curve from [3]. The result of a simple two-layer structure with differing thermal conductivity on each side of an interface resistance shows very good agreement in the constant thermal conductivity case when compared to the result of Hickson et al.

The final step here is to generalize to multiple interfaces in a single simulation. On the right side of Figure 35 we show the result of simulating a 10-layer structure with a resistance imposed at each material interface, and Dirichlet contacts imposed at the left and right edges of 300K and 400K respectively. Four separate cases are shown; 1) a constant thermal conductivity denoted
by the black dashed line for a reference value, 2) a temperature dependent simulation using the iterative method with region specific values of $\alpha$ shown as the solid red line, 3) a Kirchhoff transformation simulation using $\alpha = -1.2$ represented by the solid blue line, and 4) a Kirchhoff transformation simulation using $\alpha = -1.3$ represented by the dashed blue line.

Figure 35: Simulation of a Single Interface Resistance with differing (constant) Thermal Conductivities on each Side (left) and Simulation of a 10 layered Structure similar to that of Hickson et al [3]

Left: the black line overlaid on the image is our steady-state result in the CMC, while the image is taken from a time-dependent simulation in [3]. Their $t = 2$ curve corresponds to steady-state.

Right: the resistance set at each interface is $R = 0.5 \frac{m^2 K}{W}$, while $\kappa_A = 10 \frac{W}{mK}$ and $\kappa_B = 1 \frac{W}{mK}$.

The Kirchhoff results are seen to get progressively worse away from the Dirichlet conditions at the two edges, further suggesting it to be inaccurate for multiple interfaces.

In the cases of 3) and 4), the value of $\alpha$ used in the Kirchhoff transformation does not show a large effect in this simple layout, however the result for both gets progressively worse further away from the Dirichlet boundary conditions suggesting a growing inaccuracy when simulating multiple convective interfaces using the Kirchhoff transformation. In contrast, the iterative approach incorporating temperature dependent thermal conductivities closely resembles the constant conductivity result, as should be expected.

We have demonstrated the ability to simulate fully 3D package-scale thermal transport with temperature dependent thermal conductivity in the presence of convective boundaries through comparison of a full finite element analysis (FEA) result for a multi-finger GaN-on-SiC HEMT layout. Our approach is also easily extended to multiple convective interfaces as shown through the simulation of a 10-layer structure imposing a thermal resistance at each material resistance. The relative complexity of the simulations suggests the iterative approach to be quite robust as well, with only small refinements needed to be completed.
Section 4.11 References:


4.12 Phonon Dynamics

We have tested the full coupled charge-phonon dynamics algorithm by expanding on our previous work simulating out of equilibrium electron thermalization. This time we consider, in a single simulation, the thermalization of both electrons and holes and focus on the transient regime of the thermalization process. The goal of this simulation is to observe the energy exchange between charges and phonons, when absorption and emission are modeled using a Hawkes process.

The electron and hole populations here have been initialized using a shifted Maxwellian distribution, while the phonon population is initialized at an equilibrium 100K distribution. In other words, they are initialized from equilibrium. The populations are then allowed to evolve in time through energy exchange processes, e.g., scattering between the charge carriers and phonons.

We first take a snapshot of the charge distributions at $t = 0.1$ fs, shown in Figure 36. While it is evident the charge populations are still near their initial distributions this early in the transient regime, it’s interesting to note that secondary peaks are already visible in both populations at energies below the initialized spikes. We next look at the energy distribution of the charge carriers after they have evolved for 30 fs, e.g., at $t = 30.1$ fs.
Figure 36: Hole (left) and Electron (right) Distributions at $t = 0.1 \text{ fs}$

At this point early in the transient regime, the carriers are near their initialized out of equilibrium distribution. This is apparent from the sharp peak nature of the distributions in energy.

We would expect the distance between the peaks visible in Figure 37 to in fact be approximately the average energy of an optical phonon. This is because the dominant energy relaxation process is suspected to be the emission of higher energy optical phonons, rather than that of lower energy acoustic phonons. The distance between these peaks indeed corresponds to 0.6 eV, corresponding to the average optical phonon energy in silicon. Hence, the results satisfy our initial expectations of the evolution for the charge populations. Additionally, this reinforces that the dominant energy relaxation process is in fact due to emission of optical phonons.

Figure 37: Hole (left) and Electron (right) Distributions at $t = 0.1 \text{ fs}$

At this point early in the transient regime, the carriers are near their initialized out of equilibrium distribution. This is apparent from the sharp peak nature of the distributions in energy.
The transient regime in the thermalization of far out of equilibrium charge and phonon populations has been studied. The absorption and emission scattering events are modeled as a Hawkes process as detailed in previous reports. Finally, we clearly see that the charge distributions show peaks in energy corresponding to the average optical phonon energy; this is expected as optical phonon emission is the dominant energy relaxation process for the out of equilibrium carriers.

4.13 Flux-based Approach and Multiscale Simulation

As detailed previously, we take advantage of the fact that the thermal conductivity is allowed to vary from one cell to another and solve a constant-conductivity problem in an iterative manner. In each iteration, the thermal conductivity of each cell is updated according to the local temperature from the previous solution. In other words, we solve a system of the form:

\[ \nabla^2 T_{\mu,c}^i = -\frac{p_\mu(\vec{r})}{\kappa_{\mu,c}(T^{i-1})}. \]  \hspace{1cm} (1)

This approach overcomes all the restrictions due to the previously used Kirchhoff Transformation approach, e.g., using the same functional form everywhere for the thermal conductivity and certain boundary conditions becoming nonlinear under the transformation.

While the numerical solution of Poisson’s Equation in the context of devices imposes stability constraints on the maximum grid size and time step that can be used, the energy balance equation is macroscopic in nature and is not subject to these limitations. Hence, in the thermal solver the grid can be expanded to arbitrary sizes without worry of stability, so long as the expansion is relatively slow. In practice the expansion factor between cells can be roughly 2, meaning that you can nearly double the size dimension. This represents a geometric series which grows very rapidly, and so doesn’t present a great limitation with the overall number of cells.
For these multiscale simulations, we essentially define separate electrically active regions and regions where only thermal transport is assumed to be significant. In the active regions the full CMC is ran in the standard manner and the heat generation rates computed for the forcing function of (1). In the non-active regions however, the more computationally expensive CMC does not need to be ran at all as the heat generation here is assumed to be zero, and (1) is solved generally only a handful of times before reaching convergence. In this way the overhead of simulating large size scales is drastically reduced.

4.13.1 GaAs MESFET Simulation

Here we simulate a simple 1.28 μm x 0.392 μm GaAs MESFET, first only on the device scale and subsequently including a large substrate region underneath which is not electrically active. We use the MESFET device because it is part of a benchmarking procedure used often in the CMC to ensure reliability and accuracy anytime significant changes are made in the code.

For the device scale simulation, a perfect 300K heat sink is attached to directly to the bottom of the device, whereas for the larger scale we attach a 46 μm substrate between the bottom of the device and the 300K heat sink. The respective temperature maps obtained are shown in Figure 39.

Figure 39: Temperature Maps obtained for the GaAs MESFET Simulation
Left: simulated temperature map for the device only with a 300K heat sink attached directly to the bottom. Middle: simulated temperature map for the device with a 46 μm substrate separating the device from the heat sink. Right: Multiscale simulation temperature map zoomed in around the active region for comparison. The solid black line represents the end of the electrically active device.
When the 46 μm substrate is included in the simulation, the temperature peaks reach roughly 405K and 435K for the acoustic and optical modes respectively. This represents greater than a 100K increase for each of the respective peak temperatures seen in the device only simulation on the far left. We would expect the temperature to increase with the introduction of the large substrate region, simply because the further away the heat sink is from the heat source the less efficient it will be able to dissipate heat away. Hence, this result suggests the capability to expand the particle-flux electrothermal CMC solver into multi-scale simulations.

We are expanding the capability of the CMC simulation package to include multi-scale thermal simulations. This is possible by specifying regions in which the electrical characteristics need not be solved, but instead only the thermal energy balance equation is of importance. We have also performed electrothermal simulations on an ideal GaAs MESFET device on both the device scale as well as including a large substrate region.

### 4.14 Particle-Flux Energy Balance Equation: T-gate and II-gate HEMT Results

Here we first study the GaN/AlGaN HEMT device presented by Altuntas et al [2], the layout of which is shown in Figure 40. The first step in simulating an experimental device is calibration of the direct current (DC) operating curve to the low-bias linear region. This calibration will lead to an over-estimation of the saturation in an isothermal simulation because the effects of self-heating are neglected. However, with this calibration complete we can then perform a full electro-thermal simulation on the device and our simulated result closely follows that seen experimentally, as seen in Figure 41.

![Figure 40: T-gate HEMT Layout under Study from [2]](image)

*The source-to-gate length $L_{SG} = 500 \text{ nm}$, the gate length $L_G = 75 \text{ nm}$, and the gate-to-drain length $L_{GD} = 925 \text{ nm}$.*
The simulated (left) $I_D-V_{DS}$ (left) and $I_D-V_{GS}$ (right) for the Model of the Experimental T-gate Structure in Figure 40

Through the study of the electron distribution function, we have seen that changing the gate architecture to a Π-gate layout reduces the hot electron population by 19% - 41% as compared with a T-gate layout [1]. Here we take the current device under study from Figure 40 and split the T-gate into two gate stems with $L_{G1} = 50$ nm and $L_{G2} = 25$ nm such that the total gate length is still the original $L_G = 75$ nm. However, we add a $d = 50$ nm spacer region in between the two stems, while the dimensions of $L_{SG}$ and $L_{GB}$ are held constant, resulting in a device with a total length increased by $d$. The temperature maps obtained with both the Π-gate and T-gate layouts are shown in Figure 42, with the Π-gate results presented on the left side and the T-gate results presented on the right. The top plot is a contour of the dielectric constant simply because it provides a convenient illustration of the gate layouts of the respective devices.
Figure 42: ΙΙ-gate Architecture (left) as compared with the original T-gate (right)

Top: dielectric constant plot used as an illustration of the respective gate layouts. Middle: acoustic mode temperature maps throughout the device. Bottom: optical mode temperature maps throughout the device.

For the acoustic mode temperature, our electro-thermal solver shows approximately a 10K decrease in the peak as well as an improved temperature distribution near the top of the device where the region showing the hottest temperatures has been reduced in size. For the optical mode temperature, we see a 100K decrease in the peak temperature and a slight change in the shape of the distribution, but not as pronounced as in the acoustic mode case due to the low thermal conductivity of the optical modes.

We have studied a realistic T-gate GaN/AlGaN HEMT found in literature from [2]. For this device, the particle-flux electro-thermal CMC package accurately reproduces the $I - V$ characteristic observed experimentally. In addition, we show a ΙΙ-gate architecture improves the simulated temperature maps seen inside the device as suggested through the study of the electron distribution function (EDF) [1].
Section 4.14 References:


5. CONCLUSIONS

The CMC’s electrothermal solver has been verified to be functioning as expected through the increase in scattering, which is temperature dependent. The resulting electrical characteristics were seen to capture the decrease in current expected through self-heating, and currents obtained through electrothermal simulation closely match those reported experimentally. We have presented test cases demonstrating a fully coupled electro-thermal CMC approach using a flux-based solution of the energy balance equation for phonons. The resulting temperature maps within the device behave just as one would expect given the respective paths available for energy flow in each case.

The capabilities of the original CMC algorithm [1] have been extended through the use of a rejection algorithm. This approach retains the simulation speed advantages of the CMC scheme while allowing the adaptation of the scattering rates to the real time local conditions. Figure 26 showed the rejection algorithm flow chart. After the CMC triggers an event with probability $P_{CMC}$, a new scattering probability $P_{loc}$ is computed from the value of local parameters. The rejection probability $P_{rej} = P_{loc} / P_{CMC}$ is then used in a stochastic procedure to decide whether the scattering event occurs or is rejected.

We are focused on evaluating the electron-phonon interaction using a particle-particle simulation approach. Whereas before the electron-phonon interaction was modeled at steady state using simply a Poisson process, we now take into account a sophisticated bookkeeping scheme for the error in energy using a Hawkes Process. This approach allows us to use a much smaller number of simulated particles, reducing simulation time and memory costs, while at the same time minimizing the error in energy previously seen in our work.

We have proposed a new approach for including the effects of temperature-dependent thermal conductivity while simulating a convective condition. No information is needed a priori to linearize the transformation as elsewhere, and it has the advantages of being able to simulate much smaller size scales as well as multiple convective conditions at once. In addition, the ability of the phonon dynamics code to thermalize a far out of equilibrium electron population to the correct reference distribution was demonstrated.

We have demonstrated the ability to simulate fully 3D package-scale thermal transport with temperature dependent thermal conductivity in the presence of convective boundaries through comparison of a full FEA result for a multi-finger GaN-on-SiC HEMT layout. Our approach is also easily extended to multiple convective interfaces as shown through the simulation of a 10-layer structure imposing a thermal resistance at each material resistance. The relative complexity of the simulations suggests the iterative approach to be quite robust as well, with only small refinements needed to be completed.

The transient regime in the thermalization of far out of equilibrium charge and phonon populations has been studied. The absorption and emission scattering events are modeled as a Hawkes process as detailed in previous reports. Finally, we clearly see that the charge distributions show peaks in energy corresponding to the average optical phonon energy; this is
expected as optical phonon emission is the dominant energy relaxation process for the out of equilibrium carriers.

For multiscale simulations, we essentially define separate electrically active regions and regions where only thermal transport is assumed to be significant. In the active regions the full CMC is ran in the standard manner and the heat generation rates computed for the forcing function. In the non-active regions however, the more computationally expensive CMC does not need to be ran at all as the heat generation here is assumed to be zero, and is solved generally only a handful of times before reaching convergence. In this way the overhead of simulating large size scales is drastically reduced.

We are expanding the capability of the CMC simulation package to include multi-scale thermal simulations. This is possible by specifying regions in which the electrical characteristics need not be solved, but instead only the thermal energy balance equation is of importance. We have also performed electrothermal simulations on an ideal GaAs MESFET device on both the device scale as well as including a large substrate region.

We have studied a realistic T-gate GaN/AlGaN HEMT found in the literature. For this device, the particle-flux electro-thermal CMC package accurately reproduces the $I - V$ characteristic observed experimentally. In addition, we show a Π-gate architecture improves the simulated temperature maps seen inside the device as suggested through the study of the EDF.
6. REFERENCES

# LIST OF ABBREVIATIONS, ACRONYMS, AND SYMBOLS

<table>
<thead>
<tr>
<th>ACRONYM</th>
<th>DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>AFRL</td>
<td>Air Force Research Laboratory</td>
</tr>
<tr>
<td>BTE</td>
<td>Boltzmann Transport Equation</td>
</tr>
<tr>
<td>BZ</td>
<td>Brillouin Zone</td>
</tr>
<tr>
<td>CAD</td>
<td>Computer-Aided Design</td>
</tr>
<tr>
<td>CMC</td>
<td>Cellular Monte Carlo</td>
</tr>
<tr>
<td>DARPA</td>
<td>Defense Advanced Research Agency</td>
</tr>
<tr>
<td>DC</td>
<td>Direct Current</td>
</tr>
<tr>
<td>EBE</td>
<td>Energy Balance Equation</td>
</tr>
<tr>
<td>EDF</td>
<td>Electron Distribution Function</td>
</tr>
<tr>
<td>EMA</td>
<td>Exponential Moving Average</td>
</tr>
<tr>
<td>FEA</td>
<td>Finite Element Analysis</td>
</tr>
<tr>
<td>HEMT</td>
<td>High Electron Mobility Transistor</td>
</tr>
<tr>
<td>HTE</td>
<td>Heat Transfer Equation</td>
</tr>
<tr>
<td>LHS</td>
<td>Left Hand Side</td>
</tr>
<tr>
<td>MC</td>
<td>Monte Carlo</td>
</tr>
<tr>
<td>MESFET</td>
<td>Metal-Semiconductor Field-Effect Transistor</td>
</tr>
<tr>
<td>PDE</td>
<td>Partial Differential Equation</td>
</tr>
<tr>
<td>PI</td>
<td>Principal Investigator</td>
</tr>
<tr>
<td>RHS</td>
<td>Right Hand Side</td>
</tr>
<tr>
<td>SMA</td>
<td>Simple Moving Average</td>
</tr>
<tr>
<td>TCAD</td>
<td>Technology Computer-Aided Design</td>
</tr>
</tbody>
</table>