Solving Nongray Boltzmann Transport Equation in Gallium Nitride

Ajit K. Vallabhaneni[1]
G. W. Woodruff School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, GA 30332
e-mail: ajitkvallabhaneni@gmail.com

Liang Chen
G. W. Woodruff School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, GA 30332

Man P. Gupta
G. W. Woodruff School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, GA 30332

Satish Kumar
G. W. Woodruff School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, GA 30332

Several studies have validated that diffusive Fourier model is inadequate to model thermal transport at submicron length scales. Hence, Boltzmann transport equation (BTE) is being utilized to improve thermal predictions in electronic devices, where ballistic effects dominate. In this work, we investigated the steady-state thermal transport in a gallium nitride (GaN) film using the BTE. The phonon properties of GaN for BTE simulations are calculated from first principles—density functional theory (DFT). Despite parallelization, solving the BTE is quite expensive and requires significant computational resources. Here, we propose two methods to accelerate the process of solving the BTE without significant loss of accuracy in temperature prediction. The first one is to use the Fourier model away from the hot-spot in the device where ballistic effects can be neglected and then couple it with a BTE model for the region close to hot-spot. The second method is to accelerate the BTE model itself by using an adaptive model which is faster to solve as BTE for phonon modes with low Knudsen number is replaced with a Fourier like equation. Both these methods involve choosing a cutoff parameter based on the phonon mean free path (mfp). For a GaN-based device considered in the present work, the first method decreases the computational time by about 70%, whereas the adaptive method reduces it by 60% compared to the case where full BTE is solved across the entire domain. Using both the methods together reduces the overall computational time by more than 85%. The methods proposed here are general and can be used for any material. These approaches are quite valuable for multiscale thermal modeling in solving device level problems at a faster pace without a significant loss of accuracy. [DOI: 10.1115/1.4036616]

1 Introduction

As the size of microelectronic devices reduces, the modeling of heat transfer becomes more challenging. This is mainly because of the ballistic nature of phonons, which arise when the size of the device approaches the same order as of the phonon mean free path (mfp). At such length scales, not all the phonons undergo scattering, and thus, the thermal transport deviate from the purely diffusive regime [1]. Hence, Fourier law-based models are found to be inadequate as they yield erroneous results for temperature at these length scales. Recently, it has also been demonstrated that certain phonon modes would not contribute to thermal transport process if the size of the heat source is smaller than their mean free path [2,3]. Any model that aims to model thermal transport in such devices should serve the dual purpose of taking into account these ballistic effects and should also recover the bulk diffusive behavior at larger length scales. The semiclassical BTE fits this description precisely and hence widely used to describe the transport of phonons at such mesoscopic length scales where wave effects can be neglected [4]. The description of the physical space of the geometry and the wave vector space of phonons of the material serve as inputs to solve the BTE model [5]. The information on phonon wave vector space (phonon dispersion and relaxation times) of any material is typically obtained from first-principles calculations [6]. Simplified approaches like gray and two fluid models were proposed earlier to solve the BTE [5]. Though, they can account for the ballistic and boundary scattering effects, they would not account for different polarizations and the entire phonon scattering in the domain and thereby fails to predict the temperature accurately. This has been clearly demonstrated by Narumanchi et al. in a silicon-based device by comparing the results against the BTE solved with full dispersion [7,8]. Similar BTE calculations with full dispersion were also performed by various research groups mostly in silicon and germanium [9–11].

Gallium nitride (GaN) based high electron mobility transistors (HEMTs) are strong candidates for the future high power applications due to the superior properties of AlGaN/GaN heterostructure such as high carrier saturation velocity, wide band gap, breakdown field, and thermal conductivity [12–14]. Heat dissipation is a huge challenge that needs to be solved to enhance the reliability and operation of these devices. It is important to understand the thermal transport mechanisms and estimate the maximum temperatures in these devices accurately. Hence, in the current study, we solved the nongray BTE model for a gallium nitride (GaN) semiconducting layer typically used in HEMTs.

The wide spread in phonon relaxation times (typically over three orders of magnitude) and the large number of equations ($\sim 10^8$ for GaN HEMT) that need to be solved simultaneously makes BTE computationally expensive. Researchers have suggested different techniques to parallelize [15,16] and to accelerate the convergence mostly in the context of radiation transport equation (RTE) [9,17–21] which is similar to BTE. For brevity, we discuss here only some of the studies relevant to the present work. Adams and Larsen proposed several techniques for solving the RTE iteratively using techniques like mesh rebalance, Chebychev acceleration, source extrapolation, synthetic acceleration [22]. But they all have convergence issues for high scattering ratio. Techniques like successive over-relaxation for the discrete ordinates method perform well only in the limit of high mesh-based optical thickness [18]. Mathur and Murthy proposed the COMET to accelerate convergence of the RTE by solving the discrete energy and intensity equations at each cell simultaneously [21]. Loy et al. extended this technique to solve the phonon BTE in silicon and demonstrated acceleration by a factor of 10–300 over the existing sequential methods [23]. They also proposed a hybrid BTE–Fourier solver in which the BTEs of optical phonons are replaced with a Fourier model like equation [24] in silicon and

[1]Corresponding author.

Contributed by the Heat Transfer Division of ASME for publication in the JOURNAL OF HEAT TRANSFER. Manuscript received December 16, 2015; final manuscript received April 24, 2017; published online June 6, 2017. Assoc. Editor: Alan McGaughey.

obtained speedup factors between 2 and 200 depending on the band lattice ratio and Knudsen number (Kn). Allu and Mazumder used a similar technique to treat phonons with low Kn with an isotropic P_1 approximation instead of BTE. This approach reduces the computational time [25] for both steady-state and transient problems in 2D and 3D geometries. This hybrid solution method (also referred as adaptive method here) can also provide fast and accurate solution for temperature profile in GaN-based devices, but such methods have not been employed yet for these devices in conjunction with the COMET algorithm. The size of the GaN HEMTs can be very large ($\sim$50 μm), larger than the conventional Si-based transistors used in complementary metal-oxide semiconductor (CMOS), and it is not computationally feasible to solve the BTE over the entire HEMT. There is a need to couple Fourier and BTE models by facilitating efficient exchange of information between the two domains for fast and accurate temperature predictions in GaN devices. These techniques will also allow efficient handling of materials interfaces where exact atomic structure and phonon transmission is not known but thermal boundary resistances (TBRs) are available from the experiments. Developing such model, further accelerating the computation in the BTE domain using adaptive method discussed above, and employment of the developed model on a GaN layer typically used in the HEMTs are the focus of the present work. Solving BTE involves traversing both physical and wave vector spaces. The algorithms used to solve BTE are broadly categorized into two groups: sequential method and COMET. In the sequential method, a phonon wave vector is chosen and solved for all the cells in the physical space before moving to the next phonon wave vector. In the COMET, the process is reversed, a cell in the physical space is chosen, and all the wave vectors in that cell are solved before moving to the next cell. Though, both methods yields identical results, and which method converges quickly is decided by the Knudsen number (ratio of phonon mfp to the length scale of the domain) [23]. In the current study, we adopted the COMET technique to solve the BTE in a GaN layer. As mentioned above, it is essential to solve BTE near the hot spot to capture the ballistic effects and recover the Fourier (diffusive) like behavior away from the hot-spot. In this context, we examined ways to couple the BTE and Fourier models to reduce the computational time. This technique will be quite useful in modeling the substrate along with the device to obtain temperature profile as will be discussed later in the manuscript.

In the current work, we considered a GaN device of size 50 μm (typical gate to gate distance) $\times$ 2 μm discretized with a uniform cell size of 0.1 μm. The *K*-space (reciprocal space) of GaN is also discretized using a MonkHorst-pack grid of size $16 \times 16 \times 4$ with 12 polarizations. Here, 12,288 BTEs need to be solved at each of the 10,000 cells in the physical space. These kind of calculations requires an enormous amount of computational and storage capacity to carry out iteratively to achieve convergence. The current work present two methods to significantly reduce the computational time required to predict temperature distribution in the GaN layer of HEMT without a significant loss of accuracy. The first one is to use Fourier model away from the hot spot in the channel region of device where ballistic effects can be neglected and then couple it with a BTE model for the region close to hot-spot. The heat flux and temperature information are exchanged across the shared boundary between the two models to ensure continuity in heat flux and temperature. This method can be easily extended to device scale simulations where the substrate can be treated with the Fourier model and the device (or just a portion of it) with BTE. Such multiscale model will enable the thermal simulations of large GaN-based electronic packages by applying the BTE model only close to the hot spot region, thus reducing the computational time significantly.

The second method is to accelerate the BTE model itself by using the adaptive method. This approach was earlier utilized by Loy et al., where the BTE was solved with the sequential method [24]. But, in the current work, we solved the BTE with COMET algorithm while employing the adaptive method. The COMET method is faster than the sequential method as demonstrated by Loy et al. in the case of silicon. In our work, we would like to accelerate it even further by using the adaptive method. The phonon properties of GaN (phonon dispersion, group velocities, and relaxation times) for BTE simulations are calculated from first principles—DFT. The details of our approach are presented in Sec. 2. To the best of our knowledge, this is the first time, the modeling techniques like the ones mentioned above are employed to solve BTE in GaN. We have observed that the overall time to achieve converged solution can be reduced by more than 85%. The methods presented here are quite general and are applicable to any material of interest.

2 Numerical Methodology

2.1 BTE for Phonon Transport. The steady-state nongray BTE of a phonon mode discretized using finite volume method (FVM) under relaxation time approximation (RTA) [4] is defined as

$$\sum_f e_f'' v.\nabla A_f = \frac{e_c^o - e_c''}{\tau_{\text{eff}}}\Delta V + S_{\text{vol}};\ \ e^o = \hbar\omega N = \frac{\hbar\omega}{\exp\left(\frac{\hbar\omega}{K_b T}\right) - 1} \tag{1}$$

Here, e'' is the energy of the phonon mode, and e^o is the equilibrium energy evaluated at the cell (denoted by c) centroid, v is the phonon group velocity, ω is the phonon frequency, ΔV represents the volume of each cell in the physical space, and A_f is the area vector associated with face f of a cell in the physical domain. τ_{eff} is the effective phonon relaxation time that includes contribution from different mechanisms like the intrinsic phonon–phonon scattering, defects, and impurities, and so on. S_{vol} is the external volumetric heat source which represents the Joule heating in the current case. N is the equilibrium Bose–Einstein distribution corresponding to the phonon mode with frequency ω, polarization p, and temperature T. Parameters like e'', ν, and τ_{eff} are all unique to a particular phonon mode, but not shown in Eq. (1) for brevity. Direct coupling of all these BTEs requires denser discretization of the phonon wave vector space that makes them extremely expensive to solve. Hence, RTA is employed to simplify the scattering term on the right-hand side of the BTE. In RTA, all the BTEs of a physical space cell are coupled through the lattice temperature to ensure energy conservation. In the current study, the BTE has been solved iteratively using COMET as mentioned above, where all the phonon modes within a physical space cell are solved together directly in a matrix form before moving to the next cell. The physical space cells are divided among multiple processors as part of the parallelization. More details on solving the BTE using COMET algorithm are presented in Ref. [23].

2.2 BTE–Fourier Coupling. Figure 1 shows the schematic of a GaN HEMT, where the lateral width of the device is 50 μm. It includes the GaN buffer layer (2 μm thick) on top of silicon substrate (330 μm thick). In this device, AlGaN/GaN heterostructure leads to the formation of a two-dimensional electron gas (2DEG) in GaN layer at the interface. This 2DEG region acts as the channel (6.5 μm $\times$ 6 nm) of the GaN HEMT where most of the power dissipation occurs due to Joule heating. As mentioned in Sec. 1, solving BTE is computationally challenging, especially when one has to use dense grids near the hot spot for better accuracy and capture the spatial temperature variation. But away from the hot spot, the temperature variation will recover the diffusive behavior and a Fourier model would be adequate to study the spatial variation. Hence, multiscale models that encapsulate BTE in the ballistic region and Fourier model in the diffusive region are the best approaches to study the thermal issues at a device level in an efficient manner. The right side of Fig. 1 shows the schematic of the

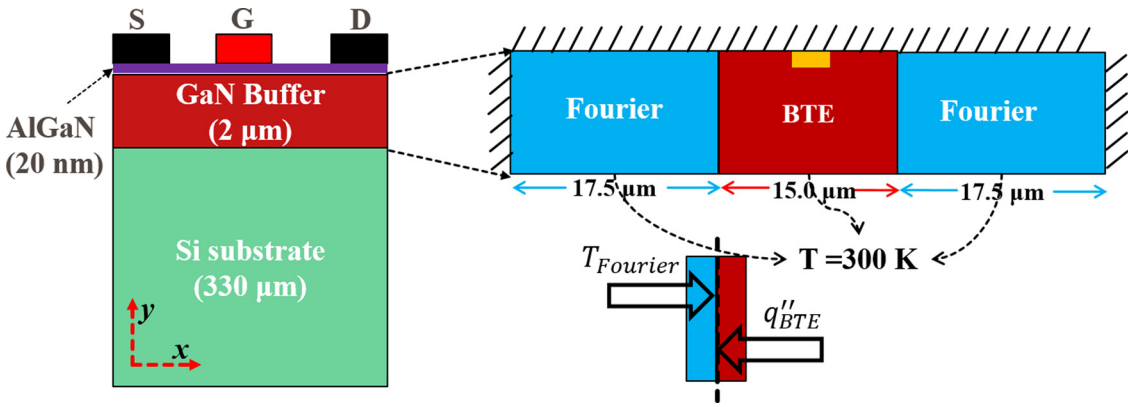

Fig. 1 (Left) block diagram of a GaN device with silicon substrate and AlGaN dielectric layer. (Right) schematic of the GaN buffer of size 50 μm × 2 μm divided into BTE and Fourier domain; the thick red line indicates the channel region where the energy is added. The exchange of temperature and heat flux information at the common interface is demonstrated in the bottom panel (see color figure online).

computational domain, i.e., GaN buffer region where we applied the multiscale approach by coupling Fourier and BTE models. The region highlighted in red is gate (G) electrode of the device below which hot spots typically appear due to the maximum heat generation [26]. Temperature prediction with high accuracy in this region is desired for efficient thermal management and device design. Hence, we solve BTE in the region close to the hot-spot, and on either sides of this region we employ diffusive Fourier model to resolve the lattice temperature distribution.

One of the main challenges in solving the multiscale model is the coupling between the BTE and Fourier models. If both domains are of the same material, then the continuity has to be maintained in both temperature and heat flux. If they are of different materials, there will be continuity in heat flux at the interface, but the thermal boundary resistance (TBR) leads to temperature drop at the interface [27]. We have considered two cases to demonstrate the effectiveness of our approach. In the first case, we have same material on both sides (GaN), and the two interfaces can be noticed from Fig. 1. At every interface, exchange of data on temperature and heat flux happens back and forth to ensure continuity. The information need to be exchanged to other domain is incorporated as the boundary conditions to that domain, e.g., either constant temperature (isothermal) or constant heat flux at the interface as demonstrated in the bottom panel of Fig. 1. Here, we calculate the wall temperature on the Fourier side of the interface and pass it to the BTE domain so that it can be used as the isothermal boundary condition and subsequently pass the interfacial heat flux from BTE to Fourier domain which in turn is used as the constant heat flux boundary condition. The size of the BTE domain is determined by the phonon mean free path (mfp) of the material of interest (GaN in our case). We have to ensure that the diffusive behavior is recovered at the end of the BTE domain. This happens when the domain size is larger than the maximum mfp. The maximum mfp in GaN is about 6 μm as shown in Fig. 3(*d*), and hence, we have decided to have a BTE domain of 15 μm encompassed by two Fourier domains of 17.5 μm on both sides. Choosing a BTE domain size close to the maximum phonon mfp could result in a temperature drop at the interface due to the ballistic effects which is erroneous. The bottom wall of the entire domain (it is in contact with substrate in actual device) has isothermal boundary conditions at 300 K, whereas the other three walls have adiabatic boundary conditions. The Joule heating is accounted for as a volumetric heat source close to the top edge of the GaN domain as shown in Fig. 1.

We have also demonstrated the use of this method by applying it across an interface of GaN and silicon. Silicon is widely used as a substrate in GaN electronic devices, and the corresponding TBR for GaN–Si interface calculated from diffusive mismatch model (DMM) is 3.25×10^{-9} m^2 K/W [11]. In our case, we have chosen a GaN domain whose length is 15 μm to avoid any ballistic effects. We have also chosen the size of silicon region to be 15 μm for convenience. The exchange of heat flux across the interface via the boundary conditions is same as mentioned above. But the temperature is modified based on the magnitude of TBR before applying on the other side of the interface as isothermal boundary condition. The effects of phonon transmission and reflection at the interface are indirectly accounted for by the TBR. We applied an energy source on top of GaN and let the heat pass through the interface by applying isothermal boundary condition on the bottom of the silicon substrate and adiabatic conditions on all other sides as shown in the inset of Fig. 5. Power distribution in the channel is obtained from hydrodynamic simulations of electrical transport using technology computer-aided design (TCAD) Sentaurus package at drain voltage (V_D) = 10 V and gate voltage (V_G) = 0 V [28].

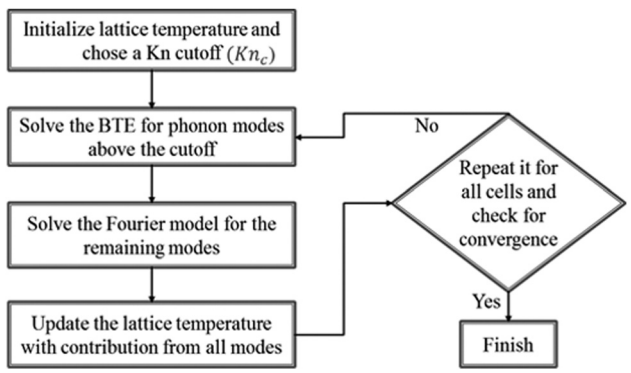

Fig. 2 Flowchart showing the algorithm used for the adaptive BTE model

2.3 Adaptive Method. The reciprocal (wave vector) space of GaN is discretized into 1024 ***K***-points for each phonon polarization. If one were to solve all the 12 phonon polarizations in GaN, it means that 12,288 BTEs are to be solved at each cell in the physical space cell. Although, the BTE–Fourier method mentioned in Sec. 2.2 is effective in reducing the BTE domain size, it is still quite expensive especially if one wants to use a denser physical mesh near the hot spot. Hence, techniques that can accelerate BTE directly are necessary to improve the efficiency. One such method is the adaptive method [24], in which the phonon modes are divided into two groups based on their Knudsen number (Kn). Kn of a phonon mode is defined as the ratio of phonon

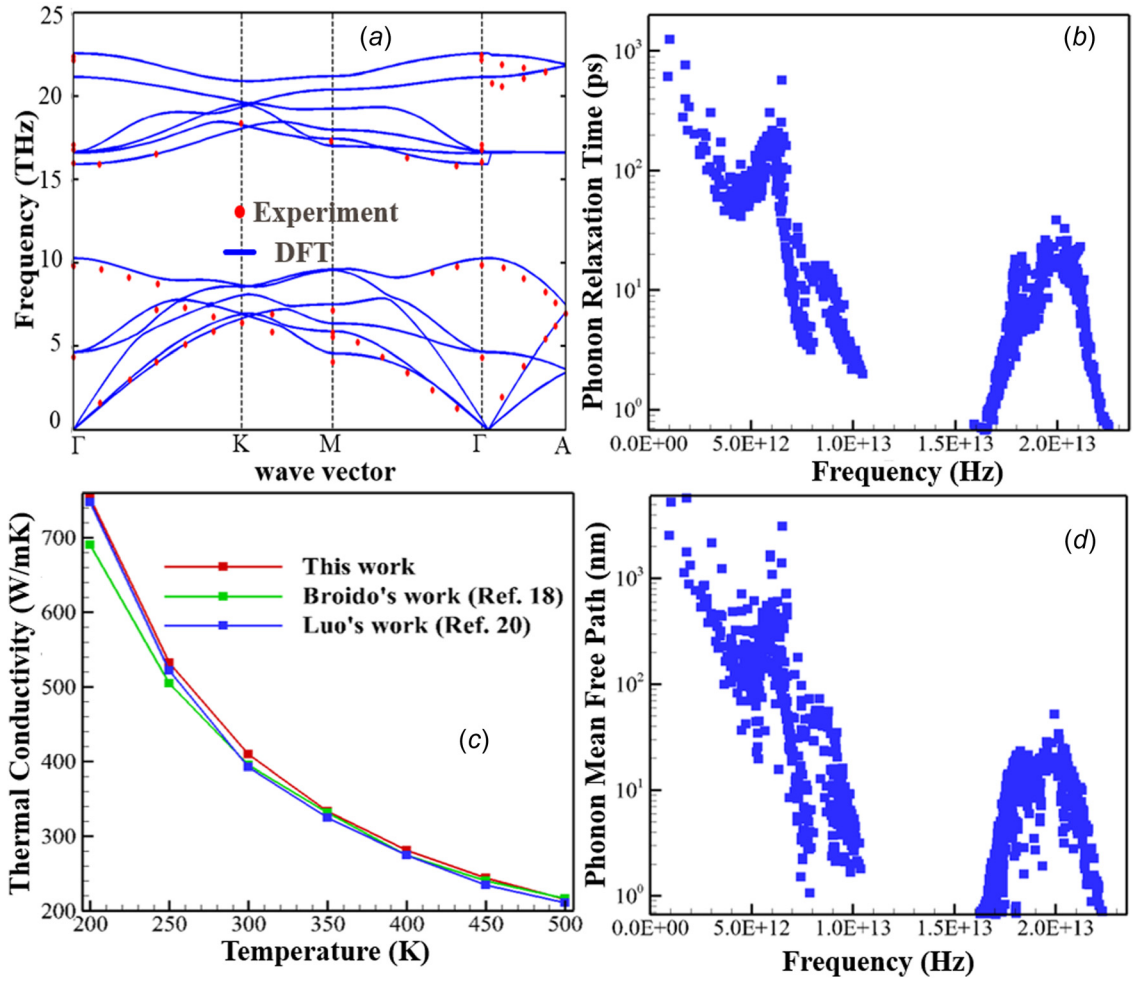

Fig. 3 (*a*) Phonon dispersion of GaN from DFT calculations. The red dots shows the comparison with experiments from Ref. [32], and the blue lines show DFT results; (*b*) phonon relaxation time (τ) in picoseconds as a function of frequency at room temperature; (*c*) thermal conductivity of GaN as a function of temperature; (*d*) phonon mean free path of GaN as a function of frequency at room temperature (see color figure online).

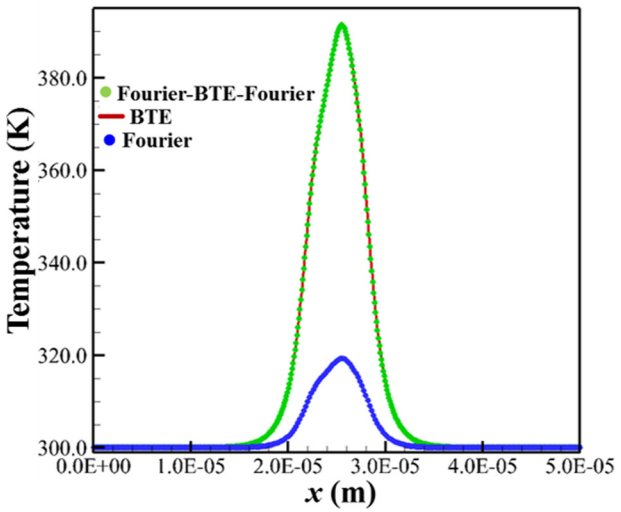

Fig. 4 Lattice temperature variation in the domain along the *x*-direction at the middle of the buffer (dotted line in the inset) at $y = 1.9 \times 10^{-6}$ μm. The Fourier model significantly underpredicts the hot-spot temperature.

mean free path (product of group velocity and relaxation time) to the length scale of the domain. Subsequently, BTE is only solved for those with high Kn which exhibit ballistic effects. The low Kn modes exhibit diffuse like behavior, and these can be solved simply using a Fourier like equation, but not exactly Fourier equation, instead of BTE. This method would be quite efficient for GaN since there are nine optical phonon polarizations. Optical phonons typically have low group velocities and short relaxation times, hence modeling them with Fourier instead of BTE would reduce the computational time significantly. The governing BTE with relaxation time approximation is given by Eq. (1), which can be integrated over all directions and converted into an equation representing the equivalent temperature T of that mode

$$\nabla \cdot (-\kappa \nabla T) = \frac{C}{\tau}(T_L - T) \tag{2}$$

This Eq. (2) is similar to a Fourier conduction equation of the phonon mode with an additional source term on the right-hand side, and hence, we call it as Fourier like equation. Here, κ, C, and τ are the thermal conductivity, specific heat, and relaxation times of a phonon mode, respectivley. T_L is the overall lattice temperature, and T is the corresponding temperature of the individual phonon mode. Each of the phonon modes will be characterized either by BTE or the Fourier based on their mfp. Solving the Fourier like equation instead of BTE is less compute intense because

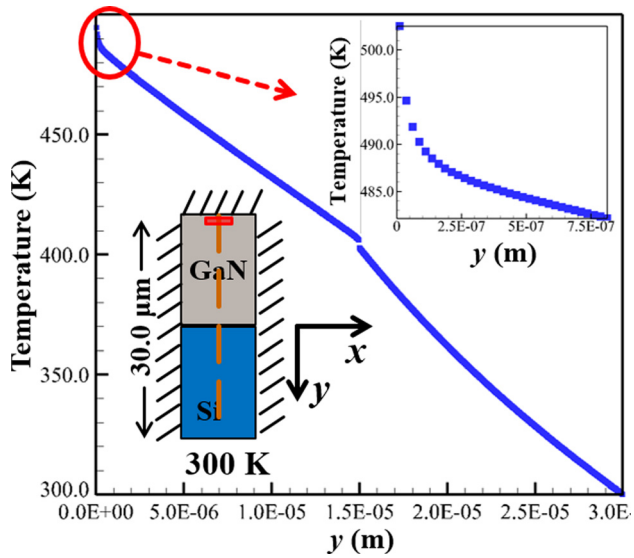

Fig. 5 Lattice temperature variation in the domain along the *y*-direction at the middle of the buffer ($x = 1.0 \times 10^{-6}\,\mu$m). Inset shows the schematic of the domain with the GaN–Si interface along with boundary conditions. The thick red box indicates the region over which energy is added. The right-hand side panel shows the nonlinear temperature variation near the source on GaN side (see color figure online).

the reciprocal space need not be resolved. The result of both these models is coupled together by the energy conservation condition which is used to compute the lattice temperature at the end of each iteration. The lattice temperature computed from the equation below determines the equilibrium energy of the phonons in both the models

$$\sum_{\omega,p}^{\text{BTE}} \frac{C_{\omega,p}(T_L - T_{\omega,p})}{\tau_{\omega,p}} + \sum_{\omega,p}^{\text{Four}} \frac{C_{\omega,p}(T_L - T_{\omega,p})}{\tau_{\omega,p}} = 0 \qquad (3)$$

The general algorithm used to solve the adaptive method is shown in the flowchart in Fig. 2. In this method, at each cell in the physical domain, all the Fourier modes are solved first, and the lattice temperature is updated. Subsequently, BTE is solved for the phonon modes with higher Knudsen number, and its contribution is updated to the lattice temperature before going to the next cell. The whole process is repeated until convergence. Knudsen number cutoff of 0.05 is chosen, unless specified differently, to divide the total number of phonon modes into Fourier and BTE modes.

2.4 Computational Details of GaN Phonon Properties. The phonon dispersion and relaxation times for GaN are calculated from the DFT [29] using Vienna ab initio simulation package (VASP) [30] following the methodology presented in Ref. [31] and briefly described next. The first principle calculations to obtain second-and third-order force constants are performed using the VASP package. The norm-conserving local-density approximation (LDA) pseudo potentials are used in these calculations, and the kinetic energy cut-off for wave-functions is 80 Rydberg. The harmonic (second-order) force constants are computed on a primitive unit cell using density functional perturbation theory (DFPT). The nonanalytical part of the dynamical matrices are considered by including the Born effective charge. The lattice constants chosen for the primitive hexagonal cell are $a = 0.3086$ nm and $c = 0.5023$ nm. These parameters were tuned so that the predicted phonon dispersion matches the experimental results [32] as shown in Fig. 3(*a*). The third-order force constants were computed using the finite difference method on a $3 \times 3 \times 3$ supercell. A displacement of 0.01 Å is systematically applied on selected atoms of the supercell, and the forces from DFT calculations are used to construct the third-order force constants. The lattice symmetry is considered using a Python tool in order to reduce the number of displacements. Generally, the relaxation time approximations (RTAs) are employed in many studies in which Fermi golden rule is used to calculate the phonon scattering rates. For GaN, the RTA predictions are quite close to the iterative solution of BTEs [31]. Based on the second- and third-order force constants from first principle calculations, an open source package (ShengBTE) has been employed to calculate the phonon relaxation times [33]. In this study, a $32 \times 32 \times 16$ q-point grid is used to obtain the converged phonon relaxation times for the pristine GaN crystal at any desired temperature. The corresponding phonon relaxation times are plotted in Fig. 3(*b*) as a function of frequency.

After obtaining the phonon relaxation times, we performed BTE calculations on a 50 -μm long thin GaN film under a constant temperature difference to estimate thermal conductivity. The obtained thermal conductivity is plotted as a function of temperature in Fig. 3(*c*) along with two other studies in the literature based on a similar approach [31,34]. It can be noticed from the plot that our values are in excellent agreement with them. After validating our DFT results, we explored the techniques to accelerate the BTE calculations. The phonon mfp is an important parameter used in optimizing these techniques, and its variation as a function of frequency, at room temperature, for GaN is plotted in Fig. 3(*d*). Though the DFT results are sensitive to input parameters like pseudo potentials used in the calculation, they would not have any significant impact on the conclusions of this work.

3 Results and Discussion

3.1 BTE–Fourier method. First, we modeled the entire domain using nongray BTE. We discretized the physical domain of $50\,\mu\text{m} \times 2\,\mu\text{m}$ with a uniform cell size of $0.1\,\mu\text{m} \times 0.1\,\mu\text{m}$ resulting in 10,000 cells and 1024 cells in the ***K***-space with 12 polarizations each. The nongray BTE is solved iteratively until convergence using 64 AMD Opteron "Abu Dhabi" @ 2.4 GHz processors. Then, in the second case, only 15 μm in the middle is modeled using BTE, and the remaining 35 μm with the Fourier model. The BTE and Fourier models are simulated in each iteration, and the boundary temperature and heat flux information are exchanged at every iteration as explained in Sec. 2.2. This simulation is performed until convergence on the same number of

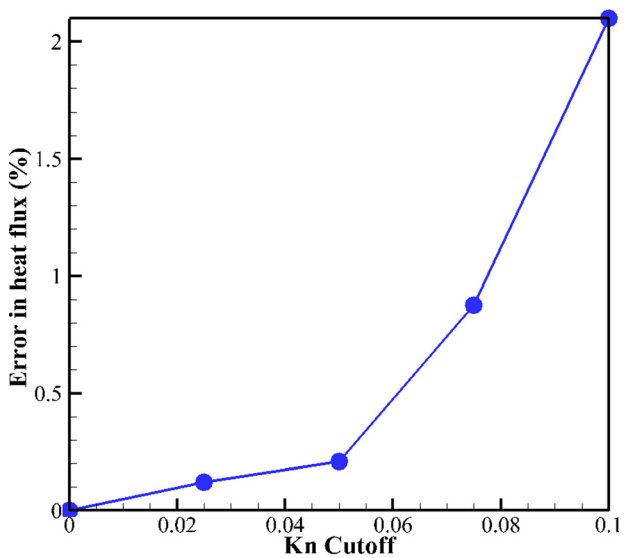

Fig. 6 Error in heat flux with respect to full BTE as a function of Kn cutoff

processors, similar to the full BTE case. It should be noted that running Fourier and BTE models just once in each iteration might give rise to issues with convergence. The convergence is sensitive to the rate of information exchange between BTE and Fourier models and sometimes could lead to overall instability. Hence, for every single iteration of the Fourier model, the BTE model is ran 10 times to ensure that sufficient heat flux is passed to the Fourier side and thereby facilitating convergence. Alternatively, an under-relaxation technique can be used as well to reduce the change in temperature of the Fourier domain at the common interface. The lattice temperature profile across the entire domain, at the middle along x direction, is plotted in Fig. 4 for both cases along with the Fourier model. The lattice temperature in BTE domain is the specific heat-averaged temperatures of all the individual phonon modes. The temperature variation is smooth across the interfaces which makes the temperature profiles in both cases identical. A major portion of the input energy from heat source leaves the domain through the bottom face which has isothermal boundary conditions. Heat flow in the lateral direction is comparatively low due to the high aspect ratio of the physical domain. Since the BTE domain is reduced by one-third, correspondingly, the time taken for each iteration in the second case also reduced significantly. The time taken to solve the Fourier equation once within the same iteration is negligible when compared to that of the BTE. Hence, the overall computational time is proportional to the number of cells in the BTE domain (N_B). In the current case, we used BTE for 15 μm which corresponds to $N_B = 3000$. It corresponds to a fraction (0.3) of the total cells in the domain. So, the computational time will be reduced by the factor N_B/N, where N is the total number of cells in the entire domain. Hence, in the current case, the total computational time is reduced by a factor of 0.3 which is approximately 70% less than the time taken for full BTE solution. It can be observed from Fig. 4 that the lattice temperature obtained from Fourier model (maximum is 321 K) is much less compared to BTE model (maximum is 392 K) in the central region near the gate. It is expected that the diffusive Fourier model cannot capture the ballistic effects.

For the case of domain with GaN–Si interface, a similar temperature profile is plotted in Fig. 5 at the center line along the GaN–Si interface, where the GaN region is modeled with BTE, and the silicon region is modeled with Fourier. Due to the TBR, a small temperature drop of about 3 K is observed at the interface. This drop is consistent with the temperature drop estimated by multiplying the applied heat flux with TBR ($\Delta T = q''*TBR$). The temperature dependence of phonon specific heat and relaxation times in GaN and thermal conductivity in silicon leads to the nonlinear temperature profile. It emphasizes the necessity to utilize BTE models to predict thermal transport in devices accurately. Both these cases demonstrate the effectiveness of our coupling algorithm in simulating thermal transport in devices.

3.2 Adaptive Method. After coupling BTE and Fourier models, we present the results using the adaptive method in this section. Our objective using this method is to reduce the computational time even further based on the algorithm explained

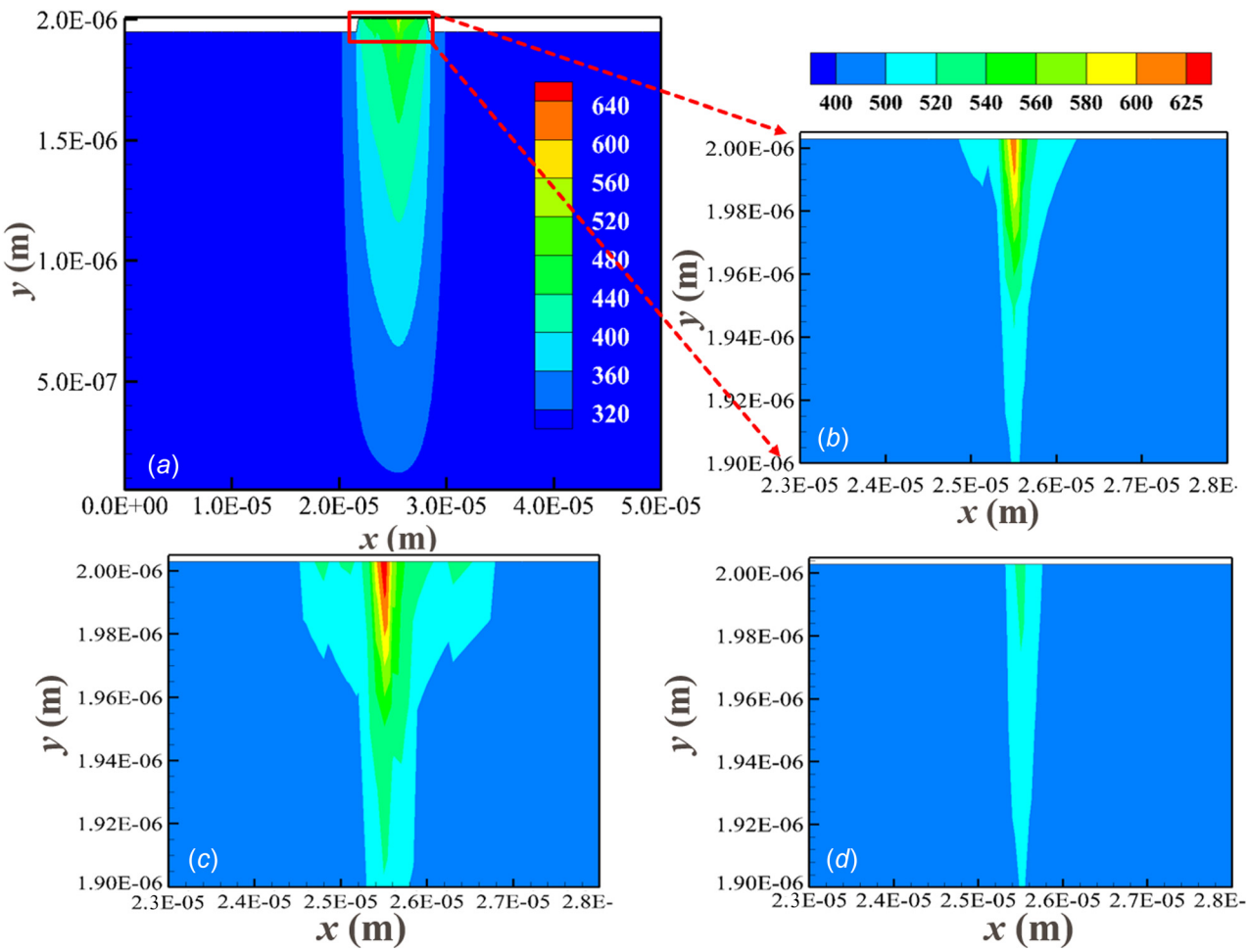

Fig. 7 (*a*) Spatial distribution of lattice temperature across the GaN Buffer layer using full BTE. Temperature distribution in the region highlighted in red from, (*b*) full BTE, (*c*) adaptive BTE, and (*d*) Fourier (the legend for temperature is same for the three cases (*b*), (*c*), and (*d*)) (see color figure online).

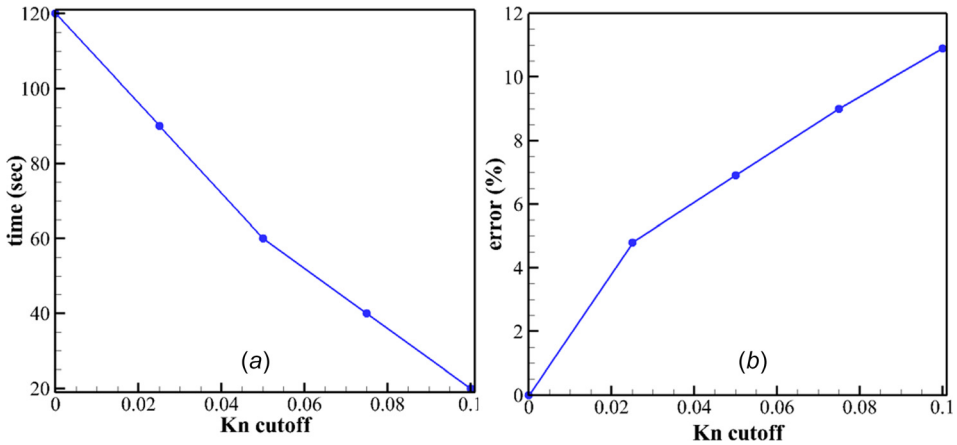

Fig. 8 Average time per iteration in seconds and error in maximum temperature as a function of Knudsen number cutoff for adaptive model (BTE is considered in only middle region of the domain, i.e., Fr + Adap. BTE + Fr model).

in Sec. 2. We first tested this methodology and the Kn cutoff on a simple one-dimensional heat transfer case where a 5-μm long block with adiabatic boundary conditions are applied on top and bottom faces. A temperature difference of 10 K (295 K, 305 K) is applied between the two ends, and the heat flux is obtained using adaptive BTE approach for different cutoffs and compared with the full BTE case. It can be noticed in Fig. 6 that the error in heat flux increases with increase in Kn cutoff as expected. For a Kn cutoff of 0.05, the heat flux from the adaptive BTE and full BTE differ by only 0.2% while yielding identical temperature profiles validating our methodology. Subsequently, we applied this methodology to reproduce the temperature profile of full BTE in two-dimensional GaN region. We used a Kn cutoff of 0.05 which means that phonon modes with mfp's less than 100 nm will be solved with the Fourier like equation described in Eq. (2), and the rest will be modeled using BTE. The discretization of the physical and reciprocal domains is same as in the earlier case. Out of the 12,288 modes, about 8400 modes fall in the earlier category, whereas the remaining falls in the later. It is the decrease in the total number of BTEs that needs to be solved will reduce the computational time significantly. One iteration of the combined BTE and Fourier modes takes about 153 s which is approximately 60% less than that of the full BTE case (360 s). We ran this simulation iteratively until convergence and obtained the temperature profile and compared it with that of the full BTE case. The results in Fig. 7 validate our approach as adaptive method yields similar result as the BTE in much less time.

Figure 7(*a*) shows the spatial distribution of lattice temperature for the full BTE case across the entire domain. Due to the isothermal boundary condition at the bottom and the high aspect ratio, heat transfer is mostly one dimensional. Though, the general variation in temperature looks similar in all cases, it is the magnitude that varies among the three. The temperature distribution in the vicinity of the channel is shown in Figs. 7(*b*)–7(*d*) for full BTE, adaptive BTE, and Fourier, respectively. The maximum temperature in the full BTE case is 621 K, whereas it is only 530 K in the Fourier case (Fig. 7(*d*)). This full BTE case serves as the benchmark for the computational time and the maximum temperature. The maximum temperature in the adaptive case for a Knudsen number cutoff of 0.05 is 649.6 K as shown in Fig. 7(*c*). This is 4.9% higher than the full BTE case, but the result can be obtained in 25% less time compared to the same. It should be noted that the temperature profile obtained using BTE–Fourier method is identical to that of the full BTE as shown in Fig. 4. But, using the adaptive method yielded temperature slightly higher than the desired. The corresponding heat flux from the bottom face after convergence is identical to that of the full BTE case for any cutoff as the principle of energy balance holds irrespective of the governing equation. Increasing the Kn cutoff would accelerate the simulation even more, but it would also increase the error. This is illustrated in Fig. 8 below where the computational time and the error are plotted as a function of the Kn cutoff for a case where both BTE–Fourier and adaptive method are applied. The Kn cutoff of zero in the figure below indicates the full BTE case. The cutoff is increased gradually from 0 to 0.1 (200 nm cutoff in phonon mean free path) in steps of 0.025, and the average computational time per iteration and error are calculated in each case. Increasing the cutoff implies that a lower number of modes are solved using the BTE and that would reduce the computational time. Hence, the computational time per iteration dropped from 120 s to 20 s for a Kn cutoff of 0.1. But the corresponding maximum temperature in the channel is higher than the full BTE case by more than 10%. The location of the source at the top edge along with reflective boundary conditions used in BTE could be the main reason for the error in the temperature estimation. One must chose the cutoff carefully depending on the desired level of accuracy and the available computational resources. Finally, we compared the average time taken per iteration in the four different cases discussed above on same number of processors in Fig. 9. The first case is the full

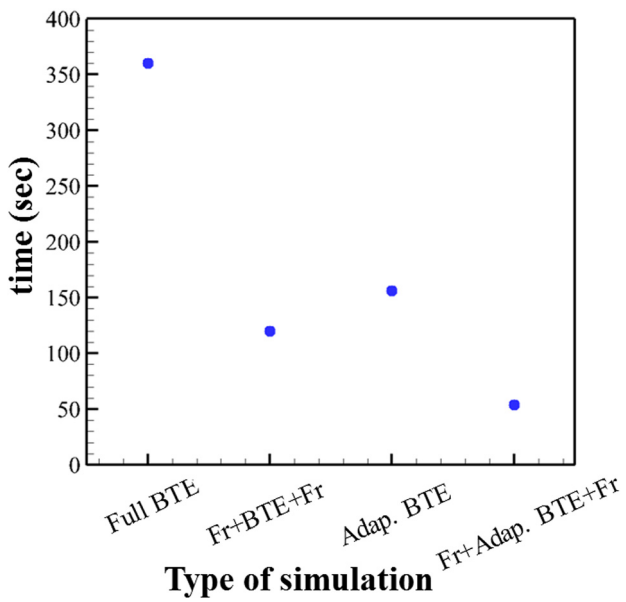

Fig. 9 Average time per iteration in seconds for the four methods considered

BTE case where only BTE is used across the entire domain, and in the second case is where we used Fourier model on both sides of the BTE domain (Fr + BTE + Fr). The third case is that of the adaptive BTE (Adap. BTE) where BTE is used across the entire domain but the phonon modes are separated into two categories based on their mfp. The fourth case is one in which we combine the second and third cases together (Fr + Adap. BTE + Fr). It can be noticed from the figure that the average time taken per iteration reduces significantly by 60–70% in either of the two methods and combining the two methods would give a reduction of more than 85%. The full BTE case takes 360 s per iteration, whereas the combined case takes only about 54 s per iteration as shown in Fig. 9.

4 Conclusions

In this work, we have studied different ways to accelerate the solving the nongray Boltzmann Transport equation in a 2D domain of GaN device. The first way is to model the domain away from the hot spots with a Fourier model and then couple it with BTE model for the region where ballistic effects must be considered. The reduction in computational time in this case is proportional to the number of cells in the BTE domain. This method can be used either within the same material or across different materials with a finite interfacial TBR. The second way is to replace the BTE for phonon modes with low mean free path with a Fourier like model to accelerate convergence. Both methods are successful in accelerating the process by yielding a reduction of 60–70% computational time. The coupled BTE–Fourier method yields the exact result as the full BTE, whereas the adaptive method yield a higher estimate in temperature which is dependent on the Kn cutoff used. Combining the two methods would lead to an overall computational time reduction of more than 85%. To the best of the author's knowledge, such acceleration techniques are applied to GaN for the first time. These techniques would be quite useful in modeling large scale electronic devices, such as multifinger GaN power amplifiers, much more rapidly without compromising on accuracy.

Acknowledgment

This work was supported in part by the DARPA Young Faculty Award No. N66001-14-1-4050. The authors would like to acknowledge the discussions with Dr. James Loy on the topic.

Nomenclature

C = volumetric specific heat capacity
e'' = phonon spectral energy density
$\hbar$ = reduced Planck's constant
K = phonon wavevector magnitude
Kn = Knudsen number
L = characteristic length
mfp = phonon mean free path
n = unit normal
N = phonon distribution function
q = heat flux vector
T = Temperature
ν_g = phonon group velocity
κ = thermal conductivity
τ = phonon relaxation time
ω = phonon frequency

References

[1] Chen, G., 2005, *Nanoscale Energy Transport and Conversion*, Oxford University Press, New York.
[2] Regner, K. T., Sellan, D. P., Su, Z., Amon, C. H., McGaughey, A. J., and Malen, J. A., 2013, "Broadband Phonon Mean Free Path Contributions to Thermal Conductivity Measured Using Frequency Domain Thermoreflectance," Nat. Commun., **4**, p. 1640.
[3] Hua, C., and Minnich, A. J., 2014, "Transport Regimes in Quasiballistic Heat Conduction," Phys. Rev. B, **89**(9), p. 094302.
[4] Hess, K., 1988, "Boltzmann Transport Equation," *The Physics of Submicron Semiconductor Devices*, Springer, New York, pp. 33–43.
[5] Ziman, J. M., 1960, *Electrons and Phonons: The Theory of Transport Phenomena in Solids*, Oxford University Press, New York.
[6] Maradudin, A. A., 1971, *Theory of Lattice Dynamics in the Harmonic Approximation*, Academic Press, Cambridge, MA.
[7] Narumanchi, S. V., Murthy, J. Y., and Amon, C. H., 2005, "Comparison of Different Phonon Transport Models for Predicting Heat Conduction in Silicon-on-Insulator Transistors," ASME J. Heat Transfer, **127**(7), pp. 713–723.
[8] Narumanchi, S. V., Murthy, J. Y., and Amon, C. H., 2006, "Boltzmann Transport Equation-Based Thermal Modeling Approaches for Hotspots in Microelectronics," Heat Mass Transfer, **42**(6), pp. 478–491.
[9] Ali, S. A., Kollu, G., Mazumder, S., Sadayappan, P., and Mittal, A., 2014, "Large-Scale Parallel Computation of the Phonon Boltzmann Transport Equation," Int. J. Therm. Sci., **86**, pp. 341–351.
[10] Ali, S. A., and Mazumder, S., 2015, "Phonon Heat Conduction in Multidimensional Heterostructures: Predictions Using the Boltzmann Transport Equation," ASME J. Heat Transfer, **137**(10), p. 102401.
[11] Singh, D., Murthy, J. Y., and Fisher, T. S., 2011, "Effect of Phonon Dispersion on Thermal Conduction Across Si/Ge Interfaces," ASME J. Heat Transfer, **133**(12), p. 122401.
[12] Mishra, U. K., Parikh, P., and Wu, Y.-F., 2002, "AlGaN/GaN HEMTs—An Overview of Device Operation and Applications," Proc. IEEE, **90**(6), pp. 1022–1031.
[13] Kuzmík, J., 2001, "Power Electronics on InAlN/(In) GaN: Prospect for a Record Performance," IEEE Electron. Device Lett., **22**(11), pp. 510–512.
[14] Wu, Y.-F., Kapolnek, D., Ibbetson, J. P., Parikh, P., Keller, B. P., and Mishra, U. K., 2001, "Very-High Power Density AlGaN/GaN HEMTs," IEEE Trans. Electron. Devices, **48**(3), pp. 586–590.
[15] Ni, C., and Murthy, J. Y., 2009, "Parallel Computation of the Phonon Boltzmann Transport Equation," Numer. Heat Transfer, Part B, **55**(6), pp. 435–456.
[16] Sawetprawichkul, A., Hsu, P.-F., and Mitra, K., 2002, "Parallel Computing of Three-Dimensional Monte Carlo Simulation of Transient Radiative Transfer in Participating Media," AIAA Paper No. 2002-2901.
[17] Chui, E., and Raithby, G., 1992, "Implicit Solution Scheme to Improve Convergence Rate in Radiative Transfer Problems," Numer. Heat Transfer, Part B, **22**(3), pp. 251–272.
[18] Fiveland, V., and Jessee, J., 1996, "Acceleration Schemes for the Discrete Ordinates Method," J. Thermophys. Heat Transfer, **10**(3), pp. 445–451.
[19] Mazumder, S., 2005, "A New Numerical Procedure for Coupling Radiation in Participating Media With Other Modes of Heat Transfer," ASME J. Heat Transfer, **127**(9), pp. 1037–1045.
[20] Mathur, S. R., and Murthy, J. Y., 2009, "An Acceleration Technique for the Computation of Participating Radiative Heat Transfer," ASME Paper No. IMECE2009-12923.
[21] Mathur, S., and Murthy, J., 1999, "Coupled Ordinates Method for Multigrid Acceleration of Radiation Calculations," J. Thermophys. Heat Transfer, **13**(4), pp. 467–473.
[22] Adams, M. L., and Larsen, E. W., 2002, "Fast Iterative Methods for Discrete-Ordinates Particle Transport Calculations," Prog. Nucl. Energy, **40**(1), pp. 3–159.
[23] Loy, J. M., Mathur, S. R., and Murthy, J. Y., 2015, "A Coupled Ordinates Method for Convergence Acceleration of the Phonon Boltzmann Transport Equation," ASME J. Heat Transfer, **137**(1), p. 012402.
[24] Loy, J. M., Murthy, J. Y., and Singh, D., 2013, "A Fast Hybrid Fourier–Boltzmann Transport Equation Solver for Nongray Phonon Transport," ASME J. Heat Transfer, **135**(1), p. 011008.
[25] Allu, P., and Mazumder, S., 2016, "Hybrid Ballistic–Diffusive Solution to the Frequency-Dependent Phonon Boltzmann Transport Equation," Int. J. Heat Mass Transfer, **100**, pp. 165–177.
[26] Baloch, K. H., Voskanian, N., Bronsgeest, M., and Cumings, J., 2012, "Remote Joule Heating by a Carbon Nanotube," Nat. Nanotechnol., **7**(5), pp. 316–319.
[27] Pollack, G. L., 1969, "Kapitza Resistance," Rev. Mod. Phys., **41**(1), p. 48.
[28] Zhang, Y., Sun, M., Liu, Z., Piedra, D., Lee, H.-S., Gao, F., Fujishima, T., and Palacios, T., 2013, "Electrothermal Simulation and Thermal Performance Study of GaN Vertical and Lateral Power Transistors," IEEE Trans. Electron Devices, **60**(7), pp. 2224–2230.
[29] Parr, R. G., and Yang, W., 1989, *Density-Functional Theory of Atoms and Molecules*, Oxford University Press, New York.
[30] Kresse, G., and Furthmüller, J., 1996, "Software VASP, Vienna (1999)," Phys. Rev. B, **54**(11), p. 169.
[31] Lindsay, L., Broido, D., and Reinecke, T., 2012, "Thermal Conductivity and Large Isotope Effect in GaN From First Principles," Phys. Rev. Lett., **109**(9), p. 095901.
[32] Ruf, T., Serrano, J., Cardona, M., Pavone, P., Pabst, M., Krisch, M., D'astuto, M., Suski, T., Grzegory, I., and Leszczynski, M., 2001, "Phonon Dispersion Curves in Wurtzite-Structure GaN Determined by Inelastic X-Ray Scattering," Phys. Rev. Lett., **86**(5), p. 906.
[33] Li, W., Carrete, J., Katcho, N. A., and Mingo, N., 2014, "ShengBTE: A Solver of the Boltzmann Transport Equation for Phonons," Comput. Phys. Commun., **185**(6), pp. 1747–1758.
[34] Wu, X., Lee, J., Varshney, V., Wohlwend, J. L., Roy, A. K., and Luo, T., 2016, "Thermal Conductivity of Wurtzite Zinc-Oxide From First-Principles Lattice Dynamics—A Comparative Study With Gallium Nitride," Sci. Rep., **6**(1), p. 22504.