

Sub-Continuum Simulations of Heat Conduction in Silicon-on-Insulator Transistors

Per G. Sverdrup

Department of Mechanical Engineering,
Stanford University,
Stanford, CA 94305

Y. Sungtaek Ju

Associate Mem. ASME
IBM Corporation,
Almaden Research Center

Kenneth E. Goodson

Associate Professor,
Associate Mem. ASME
Department of Mechanical Engineering,
Stanford University,
Stanford, CA 94305

The temperature rise in sub-micrometer silicon devices is predicted at present by solving the heat diffusion equation based on the Fourier law. The accuracy of this approach needs to be carefully examined for semiconductor devices in which the channel length is comparable with or smaller than the phonon mean free path. The phonon mean free path in silicon at room temperature is near 300 nm and exceeds the channel length of contemporary transistors. This work numerically integrates the two-dimensional phonon Boltzmann transport equation (BTE) within the silicon region of a silicon-on-insulator (SOI) transistor. The BTE is solved together with the classical heat diffusion equation in the silicon dioxide layer beneath the transistor. The predicted peak temperature rise is nearly 160 percent larger than a prediction using the heat diffusion equation for the entire domain. The disparity results both from phonon-boundary scattering and from the small dimensions of the region of strongest electron-phonon energy transfer. This work clearly shows the importance of sub-continuum heat conduction in modern transistors and will facilitate the development of simpler calculation strategies, which are appropriate for commercial device simulators. [DOI: 10.1115/1.1337651]

Keywords: Computational; Electronics, Heat Transfer, Numerical Methods, Temperature

1 Introduction

Each successive generation of transistors is scaled to smaller dimensions to improve speed and functionality. This scaling is currently yielding transistors with channels of length near 200 nm. This is comparable to the effective mean free path of phonons, which are quanta of lattice vibrations and dominant heat carriers in silicon, near room temperature [1,2]. The large electric fields near the drain side of transistors create regions of highest electron-phonon energy transfer in a transistor with characteristic dimensions of approximately 100 nm [3,4]. These regions dominate heat generation within the device and decrease in width with device scaling. Silicon-on-insulator (SOI) devices are fabricated in a thin silicon layer that is electrically isolated by a buried oxide layer as shown in Fig. 1(a). The isolation provided by the buried oxide reduces the effective electrical capacitance of the channel region and allows for greater transistor speed. However, SOI devices are particularly susceptible to thermal failure because of the low thermal conductivity of the underlying silicon dioxide layer. Cooling is enhanced by lateral thermal conduction in a silicon layer of thickness as low as 50 nm [5]. Because dimensions are comparable with the phonon mean free path, thermal simulations must consider the impact of phonon scattering on material boundaries and the small dimensions of the heated region on the temperature distribution. These sub-continuum thermal conduction phenomena cannot be directly calculated using the heat-conduction equation based on the Fourier law.

Phonon transport in semiconductor devices is modeled at present using diffusion theory based on the Fourier law (e.g., [4,6,7]). However, past work has illustrated the impact of non-local, or sub-continuum, phonon conduction in relatively simple geometries, including thin films, superlattices, and polycrystalline materials [1,8,9]. Sondheimer [10] analytically solved the BTE for electron transport in thin films, yielding a reduced effective electrical conductivity accounting for boundary scattering. This result

can also be used to calculate a reduced effective phonon thermal conductivity accounting for interface scattering. This approach does not account for the severe departure from equilibrium in the phonon system that occurs when the phonon mean free path, Λ , is comparable to or exceeds the characteristic heat source dimension [11]. There have been no simulations that resolve non-local phonon transport in practical device geometries.

The present work integrates the two-dimensional BTE to simulate heat transport in a practical silicon-on-insulator (SOI) device geometry schematically drawn in Fig. 1(b). This type of device consists of two distinct regions: the silicon device layer and the buried oxide layer in which energy carriers have effective mean free paths of approximately 300 nm [2] and 1 nm, respectively, at room temperature. Although the phonon transport theory and therefore the stated mean free path of 1 nm are not strictly appropriate for amorphous materials (e.g., [12]), the important conclusion gained from this scaling remains valid. The phonon BTE is solved in the silicon layer to accurately model heat transport, but the simpler heat diffusion equation is used in the amorphous buried oxide region because the characteristic lengthscale of conduction is much smaller than the film thickness. The two distinct computational regions are coupled through interface conditions that account for differences in material properties. Predictions are compared to the temperature distribution obtained using the heat diffusion equation with bulk thermal properties in the commercially available semiconductor device simulator, MEDICI [13]. The electrical behavior of the SOI device is simulated in both cases using MEDICI, which solves the Poisson equation together with number conservation equations for electrons and holes. The electrical simulations are used to extract the spatial distribution of heat generation for incorporation into the phonon heat conduction simulations. The goal is to determine the impact of the small dimensions of the heat source and the small film thickness on the temperature distribution within the device. This will help with the interpretation of transistor failure data, e.g., due to electrostatic discharge (ESD) [14]. Furthermore, the simulations developed

Contributed by the Heat Transfer Division for publication in the JOURNAL OF HEAT TRANSFER. Manuscript received by the Heat Transfer Division September 23, 1999; revision received June 25, 2000. Associate Editor: A. Majumdar.

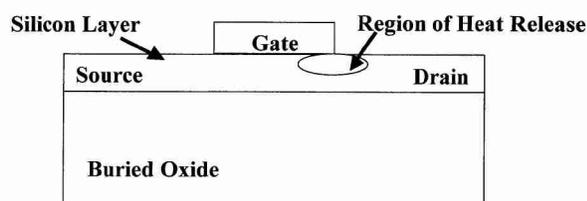
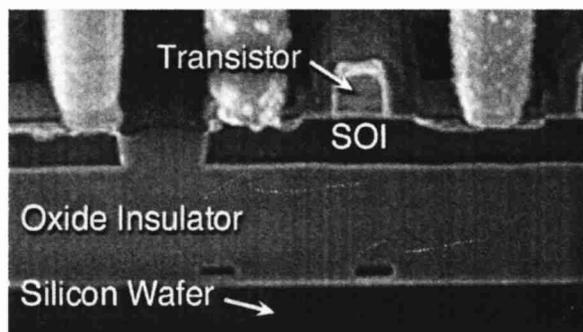


Fig. 1 (a) Scanning electron micrograph of an IBM SOI transistor, which shows the thin silicon device layer separated from the bulk silicon substrate by a buried oxide layer (Courtesy of IBM Corporation); (b) schematic representation of a SOI device used in these simulations. The region of heat release corresponds to the region of strong electron-phonon scattering. The simulated device is a NMOS transistor with source/drain doping of arsenic at $1.5 \times 10^{20} \text{ cm}^{-3}$, phosphorous source/drain extensions, and a background boron doping concentration of $1 \times 10^{17} \text{ cm}^{-3}$.

here will facilitate the development of simpler modeling strategies that can be readily incorporated into MEDICI and similar simulation packages for industrial use.

2 Governing Equations and Solution Approach

This manuscript demonstrates the impact of sub-continuum phonon transport on the temperature distribution in a practical device. The vehicle for demonstrating this is the SOI transistor with channel length of 400 nm shown in Fig. 1(b). Because of the computational demands presented by the phonon BTE, this initial study does not bi-directionally couple the electrical and thermal behavior in the device. Electrical transport and the spatial distribution of energy transfer to the phonon system are simulated using the standard industrial approach, which solves the Poisson equation and coupled conservation equations for electrons and holes. The phonon heat conduction simulations solve the BTE in the silicon layer, in which the mean free path is comparable to the layer thickness and the dimensions of the heat source. The continuum heat diffusion equation is solved in the underlying oxide.

2.1 Electrical Transport and Rate of Heat Generation in the Phonon System. The Poisson and electron and hole conservation equations are integrated using a commercial device simulator (MEDICI), which incorporates the distributions of impurities in the device based on a standard processing history. This approach is not the most detailed available for electron transport. More complicated simulations integrate the electron and hole Boltzmann equations [15] or take the hydrodynamic approach [3,16] to model electron energy and momentum transfer. In these simulations, energy transfer from highly energetic electrons, which are accelerated by the electric field in the device, to the heat conducting phonon system is calculated as the dot product of the electric current density and the electric field. This heat generation term is only approximately modeled in this manner in macro-

scopic devices, and this approach can yield significant errors in sub-micron devices. Due to the ballistic nature of electron transport, the peak in the dot product of current density and electric field profile can be displaced slightly from the peak in the transfer rate of electron energy to phonon energy. This shift in the heat source distribution could be expected to effect the peak temperature and temperature distribution within the device.

However, for two reasons detailed in the following paragraph, the use of the heat generation profile calculated from the dot product of current density and electric field is appropriate for this manuscript. Detailed simulation of hot electron effects in GaAs FET structures using the hydrodynamic electron transport equations [17] and Monte Carlo techniques [18] have demonstrated that the peak in the heat generation profile can be shifted from the peak in the electric field by as much as 100 nm. This work on GaAs found that the displacement of the maximum heating rate from the peak in the electric field is caused by the finite rate at which electrons leaving the high field region emit optical phonons. This shift can be approximated by multiplying the electron velocity (10^5 m/s) with the electron-optical phonon scattering rate, $\sim 1 \text{ ps}$ [17], to yield a shift in the heat source of 100 nm. In silicon, the relaxation time for electron-optical phonon scattering is approximately 0.3 ps [19]. The electron drift velocity is approximately 10^5 m/s in the saturation regime in silicon. The product of these two numbers yields $\Lambda_{\text{electron-optical phonon}} \sim 30 \text{ nm}$. This can be compared to the 400 nm channel length of the transistor and the phonon-phonon mean free path for heat transport, which is estimated to be 300 nm at room temperature. Based on this approximate estimate, the shift in the peak of the heated region in the transistors ($\sim 30 \text{ nm}$) may be a factor of ten smaller than the channel length (400 nm) and the phonon mean free path ($\sim 300 \text{ nm}$). While the impact on the lattice temperature distribution in the device may not be entirely negligible, the effect is indeed anticipated to be small and not expected to significantly alter the basic characteristics (peak and general shape) of the lattice temperature distribution obtained in the analysis.

The displacement of the heat source due to ballistic transport by the electrons is one among many important effects, including recombination of electrical carriers, that are neglected when the heat source is approximated as the dot product of current density and electric field. While the MEDICI simulations do not capture the precise location of the heated region, the magnitude of the heat generation calculated by the dot product of the current density and the electric field is a reasonable approximation to the actual heat generation rate in the steady state [20].

2.2 Phonon BTE in the Silicon Region. The phonon BTE is integrated in the silicon device layer because Λ exceeds the layer thickness and the heat source dimensions. The transient BTE can be written in the relaxation time approximation as

$$\frac{\partial f}{\partial t} + \vec{v} \cdot \nabla f = \frac{f_{\text{eq}} - f}{\tau_{\text{phonon}}} + q_{\text{electron-phonon}}, \quad (1)$$

where $f(x, y, t, \omega, \mu, \eta)$ is the number of phonons in a given state described by the spatial coordinates x and y , the time t , the phonon frequency ω , and the direction cosines $\mu = \cos \theta$ and $\eta = \sin \theta \cos \phi$, which project the directions of phonon transport onto the x - y plane using the polar and azimuthal angles, θ and ϕ . In these simulations, the phonon velocity is assumed to be isotropic and is given by v . The first term on the right of Eq. (1) accounts for phonon scattering in the relaxation time approximation using the Planck distribution function,

$$f_{\text{eq}} = \frac{1}{\exp\left(\frac{\hbar \omega}{k_b T}\right) - 1} \quad (2)$$

and the phonon scattering rate, τ_{phonon} . The mean free path is calculated using

$$\Lambda = v \tau_{\text{phonon}} \quad (3)$$

The second term on the right hand side of Eq. (1) is the heat generation term, which is modeled as the dot product of current density and electric field. The transfer of energy from the electrons to the silicon lattice occurs on a timescale of about 0.3 ps through interactions between electrons and optical phonons [19], while the transport of thermal energy by phonons has an effective relaxation time of approximately 70–80 ps [21], yielding a difference of two orders of magnitude. Consequently, the energy transfer from the electrons to phonons is assumed to occur instantaneously when compared to heat transport time scales. This allows direct incorporation of the heat generation term from the device simulator into the BTE solver without solving coupled phonon and electron equations.

To be rigorous, phonon transport simulations should incorporate the frequency dependence of the phonon mean free path, Λ , and account for interactions between phonons of different frequencies. This requires solution of the BTE for many different frequencies, which has been performed for one-dimensional transport by several authors [9,11,22]. The present study focuses rather on the complexity of two dimensional transport in a practical device structure, and aims to illustrate the key phenomena through calculations neglecting the frequency dependence of phonon properties while capturing the essential heat transport physics. The present calculations are based on a simplified phonon energy transport model [2,21], where phonons are grouped into propagation and reservoir modes based on the dispersion relationships of the phonon modes [23]. The reservoir mode phonons have very small relaxation times or group velocities, and are assumed to remain near equilibrium. The propagation mode accounts for the entire heat transport, and is characterized using a single effective mean free path under the gray medium approximation. The distinction between the two modes is particularly relevant for covalently bonded semiconductors, where neglecting phonon dispersion can lead to significant error in estimating an effective carrier mean free path [24]. Based on the experimentally measured thermal conductivities of silicon films with thickness below 300 nm, Ju and Goodson [2] proposed that the longitudinal acoustic phonons are dominant heat carriers in silicon near room temperature and above. The present study assumes that the longitudinal acoustic phonons constitute the propagation mode, which yields a value for the effective mean free path consistent with the experimental data. The mean free path and phonon scattering rate used in this work are effective values for bulk silicon which account for scattering between all types of phonon modes (propagating-propagating and propagating-reservoir) through Matthiessen's Rule. The effective reduction in thermal conductivity caused by boundary scattering occurs when phonons travel near boundaries, without any additional parameters. The mean free path and relaxation time are extracted from the bulk thermal conductivity using the longitudinal acoustic phonon or propagating mode heat capacity and velocity in the following equation:

$$k = \frac{1}{3} C_{\text{prop}} v_{\text{prop}} \Lambda = \frac{1}{3} C_{\text{prop}} v_{\text{prop}}^2 \tau_{\text{phonon}} \quad (4)$$

The BTE can be transformed into an equation that describes the phonon energy density by introducing the following quantity [2,21]

$$e'' = \int \hbar \omega [f - f_{\text{eq}}(T_{\text{ref}})] D(\omega) d\omega, \quad (5)$$

where e'' is the excess energy of phonons per unit volume and per unit solid angle, $\hbar \omega$ is the phonon energy, $f_{\text{eq}}(T_{\text{ref}})$ is the frequency dependent distribution function at a reference temperature, and $D(\omega)$ is the density of states per unit volume. The transformed BTE becomes

$$\frac{\partial e''}{\partial t} + \vec{v} \cdot \nabla e'' = \frac{e''_{\text{eq}} - e''}{\tau_{\text{phonon}}} + q'_{\text{electron-phonon}} \quad (6)$$

in which the parameters are frequency independent.

The phonon energy per unit volume per unit solid angle, e'' , depends on the two dimensional x and y grid location as well as the direction of phonon transport within the sphere of solid angles. For calculations of the heat flux and total energy per unit volume at a given location, the Discrete Ordinates Method (DOM), originally developed for neutron and radiative transport, is used for integration over all solid angles. The solid angle sphere is divided into discrete directions of phonon transport distinguished by sets of μ and η values with weights, w_i , corresponding to solid angle fractions on the unit sphere [25–27]. The weights satisfy

$$\sum_i w_i = 4\pi. \quad (7)$$

This work uses the level symmetric hybrid (LSH) scheme [26] for angular discretization. Level symmetric techniques use the same set of direction cosine values in each coordinate direction ($\mu_1 = \eta_1, \mu_2 = \eta_2, \dots$). In addition, the directions of phonon transport are chosen such that any particular direction has corresponding directions with 90 deg rotations about any axis ($\mu_n = \eta_n, \mu_n = -\eta_n, -\mu_n = \eta_n, -\mu_n = -\eta_n$). This simplifies the implementation of boundary conditions and removes directional biasing. The LSH method calculates directions and weights based on constraint equations such that moment equations of the phonon intensity are satisfied [26].

In order to obtain the temperature distribution, Eq. (6) is integrated forward in time in each of the discrete directions. The MacCormack time integration method, which is frequently used for solving hyperbolic equations [28], is used for explicit time advancement in order to capture the sharp features associated with the source term and to resolve the wavelike nature of solutions to the BTE. At each time step, e''^{m+1} is calculated from e''^m and e''_{eq}^m . Then, e''_{eq}^{m+1} is calculated before the solution can proceed. The equilibrium energy per unit volume per unit solid angle, e''_{eq} , is assumed to be the average value of e'' over all solid angles that would be obtained if the phonon system was allowed to relax to a uniform state. The equilibrium value of phonon energy per unit volume per unit solid angle is calculated as shown in Eq. (8).

$$e''_{\text{eq}} = \frac{1}{4\pi} \int_{\Omega=4\pi} e'' d\Omega \cong \frac{1}{4\pi} \sum_i e''_i w_i. \quad (8)$$

If the temperatures of the system are large compared to the Debye temperature or if the temperature variations within the system are small, the phonon heat capacity per unit volume, C , can be assumed to be constant. In this case, integration of the equilibrium energy density per unit volume per unit solid angle over all solid angles results in

$$e'_{\text{eq}}(T) = C(T - T_{\text{ref}}). \quad (9)$$

In this study, Eq. (9) is used to set appropriate values for the boundary conditions and to extract the final equilibrium temperature of the system using the frequency averaged heat capacity.

In order to capture the essential physics with the gray medium approximation, the heat capacity of the dominant longitudinal acoustic phonon heat carriers must be properly calculated using the phonon dispersion relation. The temperature dependence of the heat capacity of longitudinal acoustic phonons is evaluated using [21]

$$C = \frac{k_b^4}{2\pi^2 \hbar^2} T^3 \int_0^{\theta_D/T} \frac{1}{v_{ph}^2 v_{gr}} \frac{x^4 e^x}{(e^x - 1)} dx, \quad (10)$$

where θ_D is the Debye temperature for longitudinal acoustic phonons, v_{ph} and v_{gr} are the phase and group velocities of the longitudinal acoustic phonons, respectively, and x is a nondimen-

sional frequency given by $x = (\hbar\omega)/(k_bT)$. Equation (10) is a direct extension of the dispersion model for thermal conduction in silicon developed by Holland [29] who provides the parameters used here.

2.3 Heat Diffusion Equation in the Silicon Dioxide Region

The heat diffusion equation is valid in the amorphous buried oxide layer, and is written as

$$\frac{1}{\alpha_t} \frac{\partial T}{\partial t} = \nabla^2 T, \quad (11)$$

where α_t is the thermal diffusivity. Parabolic equations of this type have severe time step restrictions when integrated using explicit methods, resulting in time steps that are much smaller than required for stability by the MacCormack time integration of the BTE. Consequently, the Alternating Direction Implicit scheme [30] is used for time integration of the diffusion equation such that the time step can be made equal to that used for the silicon region.

3 Modeling the Silicon/Oxide Interface

In order to couple the silicon and oxide solution domains, it is necessary to calculate the flux of energy through the interface between the two materials at each point along the interface for every time step. When phonons strike an interface between two dissimilar materials, a fraction of the phonons will be transmitted through the interface with the remainder being reflected. The acoustic mismatch model (AMM) and the diffuse mismatch model (DMM) predict the boundary resistance between dissimilar materials [31,32]. The acoustic mismatch model assumes that all phonons have a specular interaction with the interface. This assumption is appropriate at low temperatures, for which the phonon wavelengths are much larger than the characteristic length scales of the surface roughness (e.g., [33]). For room temperature device operation, the dominant transport phonons have a much shorter wavelength than low temperature phonons, and the diffuse scattering of the DMM is more appropriate. The DMM [32] calculates the phonon transmission probabilities by using the principle of detailed balance and the densities of phonon states in each material. Written in a form appropriate for calculations using a single phonon mode, this yields

$$\alpha_i(\omega) = \frac{v_{3-i}^2}{\sum_i v_i^{-2}}, \quad (12)$$

where i is the material type (1 or 2) and v is the speed of the phonon. For small temperature differences across the interface, α is frequency independent [33]. Equation (12) does not rigorously account for the dispersion relationships of phonons and is most appropriate for low temperatures. This equation yields values of the transmission coefficient equal to 0.52 from silicon to oxide and 0.48 from oxide to silicon, using phonon velocities equal to 4240 m/s and 4100 m/s for silicon (high temperature longitudinal acoustic phonons) and oxide, respectively [29,34].

Both the DMM and the AMM theories have larger discrepancies with the experimental data near room temperature than at low temperatures. One explanation for this discrepancy is that these theories neglect nonlinear phonon dispersion and the phonon cutoff frequency, ω_d . Lattice vibrations cannot exist with wavelengths smaller than the separation between neighboring atoms, resulting in a phonon frequency, ω_d , above which phonons cannot exist [35]. The cutoff frequency is related to a characteristic temperature known as the Debye temperature, θ_d , by

$$\theta_d = \frac{\hbar\omega_d}{k_b}. \quad (13)$$

If the materials on opposing sides of the interface have different Debye temperatures, the maximum frequency that can be transmitted across the interface is determined by the material with the

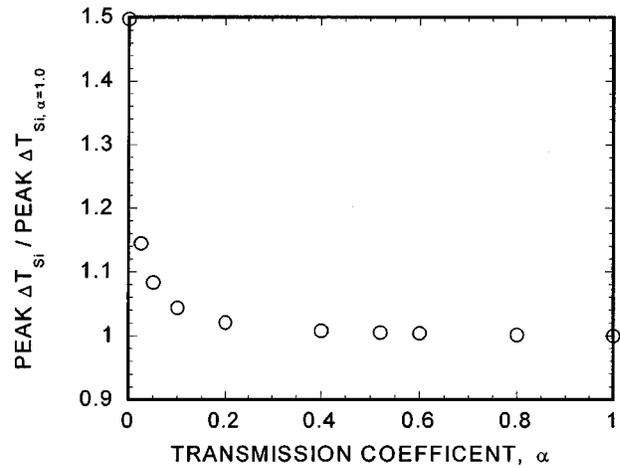


Fig. 2 Impact of the energy transmission coefficient at the silicon/silicon dioxide interface on the peak device temperature rise

lower Debye temperature if only elastic scattering events are considered. For the ensemble of phonons at all energy levels, this results in a smaller transmission coefficient than is predicted by the DMM. The Debye temperatures of interest for a SOI device are 570 K [29] and 492 K [34] for silicon (longitudinal acoustic phonons) and amorphous silicon dioxide, respectively. At temperatures significantly below the Debye temperature, phonons are not occupying modes near the cutoff frequency and the DMM is valid. In this case, the device temperature exceeds $0.5\theta_d$ and phonons with frequencies greater than $\omega_{d,oxide}$ cannot transmit between silicon and oxide if only elastic collisions are allowed.

In the graybody approximation used here, the effect of the cutoff frequency might be considered by reducing the transmission coefficient for transport from the material with the higher Debye frequency. The reduction would depend on the temperature near the interface because this influences the fraction of phonons above the lower Debye frequency. However, this approach might incur important errors due to inelastic scattering at the interface and nonlinearity in the phonon dispersion relationships near the Brillouin zone edge. For this reason, the present calculations use values of the transmission coefficients calculated using Eq. (12) without modification. In practice for this work, the precise value chosen for the transmission coefficients is not critical. This is shown in Fig. 2, which plots the peak temperature rise for a given transmission coefficient normalized by the peak temperature rise obtained by assuming that all of the incident energy on the silicon/oxide interface can pass through into the oxide ($\alpha = 1$). This figure shows that the peak device temperature rise changes by less than two percent resulting from varying the transmission coefficient for materials with similar acoustic properties ($\alpha \sim 0.2-0.8$). The thinness of the silicon layer results in many phonon boundary scattering events at the silicon/oxide interface as the heat is transported along the silicon layer. This means that a large fraction of the heat in the phonon system can pass into the oxide, yielding results insensitive to the transmission coefficient for the geometry and boundary conditions used in this analysis.

The interface condition can now be specified with these coefficients and an energy balance at the interface depicted in Fig. 3. The resulting Eqs. (14) and (15), are listed below.

$$q = G - J = G - (\varepsilon E_b + \rho G) \quad (14)$$

$$J = \varepsilon E_b + \rho G. \quad (15)$$

The parameter q is the flux of energy into the oxide, and G is the irradiation or incident flux on the interface from the silicon side. The radiosity, J , is the flux of energy into the silicon and

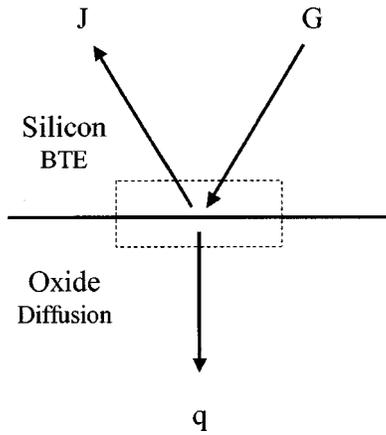


Fig. 3 Energy balance at the silicon/oxide interface

consists of phonon emission at the oxide surface temperature and the fraction of G that is reflected back into the silicon. The quantity E_b is analogous to the blackbody emissive power from radiation theory, but differs slightly due to the complexity of the dispersion relationships of a solid. It can best be defined as the energy flux that would be emitted into a silicon region by a surface maintained at the temperature of the neighboring oxide and is given by

$$E_b = \nu_p \sum_{\eta^+} \eta e_{\text{eq}}''(T_{\text{ox},i}^{n+1}) w = \frac{\nu_p}{4\pi} C_p(T_{\text{ox},i}^{n+1} - T_{\text{ref}}) \sum_{\eta^+} \eta w. \quad (16)$$

The parameter α , defined by Eq. (12), is the fraction of G that passes into the oxide, $\rho = 1 - \alpha$ is the fraction of G reflected back into the silicon, and $\varepsilon = \alpha$ is the emissivity of the interface.

The numerical integration procedure is as follows. First, the solution in the silicon layer is explicitly advanced from time step n to $n+1$, assuming that the entire system is initially in equilibrium at a uniform temperature. This calculation yields the distribution of phonon energy per unit volume and allows for calculation of the irradiation, G , of the interface at each grid point with integration over all solid angles

$$G = \nu \sum_{\eta^-} \eta e'' w, \quad (17)$$

where ν is the longitudinal acoustic phonon velocity, which is assumed to be a constant. Next, the oxide must be advanced to time $n+1$ using the flux, q , as a boundary condition. The Alternating Direction Implicit method is used for solution of the oxide temperature distribution, and this requires knowledge of the interface flux at time $n+1$ given by

$$q_i = \nu \left(\alpha \sum_{\eta^-} \eta e'' w - \varepsilon \sum_{\eta^+} \eta e_{\text{eq}}''(T_{\text{ox},j}^{n+1}) w \right), \quad (18)$$

where the diffuse blackbody radiation, E_b , leaving the surface is written in terms of the equilibrium energy per unit volume per unit solid angle at the oxide temperature at the time step $n+1$, T_{ox}^{n+1} . However, T_{ox}^{n+1} is an unknown that is obtained by solution of a system of equations. Iteration must be used with guesses for the oxide temperature in Eq. (18) until the following condition is satisfied:

$$q_i = -k_{\text{ox}} \nabla T_{\text{boundary}, t=n+1} \quad (19)$$

which states that the flux given by Eq. (18) is equal to the flux given by the diffusion equation in the oxide.

The solution is now updated to the time step $n+1$, and the solution can proceed with another solution of the BTE in the

silicon layer. The boundary condition for the BTE at the silicon/oxide interface must be changed to yield the correct energy flux from the boundary, J . This is accomplished by resetting the values of e'' leaving the boundary into the silicon. Diffuse mismatch theory assumes that all scattering events at the boundary are diffuse, resulting in a uniform distribution of phonon intensity over the hemisphere of solid angles leaving the interface. For diffuse scattering, the phonon intensity leaving the surface in each direction is given by

$$I_{\eta^+} = \frac{J}{\pi}, \quad (20)$$

where the intensity is

$$I = \nu e''. \quad (21)$$

Using these relations, e'' in all directions leaving the interface in the positive direction is

$$e''_{\eta^+} = \frac{1}{\pi} \left(\varepsilon \sum_{\eta^+} \eta e_{\text{eq}}''(T_{\text{ox}}^{n+1}) w + \rho \sum_{\eta^-} \eta e'' w \right). \quad (22)$$

An adiabatic condition is maintained at the top of the silicon device, which yields

$$e''_{\eta^-} = \frac{1}{\pi} \sum_{\eta^+} e''_{\eta^+} w_{\eta^+} \eta. \quad (23)$$

The bottom of the simulation domain is assumed to be at a uniform temperature of 300 K because of the high thermal conductivity silicon substrate that is beneath the silicon dioxide layer. The left and right boundaries are also at a uniform temperature of 300 K at a distance from the heat source of the order of the thermal decay length in the silicon layer. These boundary conditions minimize the impact of the boundaries in the heat source region and effectively illustrate the importance of boundary scattering and small heat source size on the temperature distribution.

4 Numerical Methods and Code Verification

A cartesian grid is used for the discretization of the x - y plane in the silicon and silicon dioxide region. The minimum feature size of the heat generation region is of the order of $0.01 \mu\text{m} = 10 \text{ nm}$. This is more than two orders of magnitude smaller than the domain width. Consequently, a non-uniform or stretched grid must be used in order to minimize computation time while accurately modeling the physics in the heat source region. The grid is centered around the peak in the heat generation profile and is increased in size using a geometric series with each consecutive grid spacing increasing by

$$dx_n = dx_{\text{initial}} r^{n-1} \quad (24)$$

in which r is the grid stretching parameter. In these simulations r is 1.04.

This code uses the explicit MacCormack time advancement scheme in order to capture the wave like nature of the solution. The time stepping criterion can be written in the familiar form used for advection type simulations in fluid mechanics

$$\nu \frac{dt}{dx} = cfl \quad (25)$$

in which ν is the carrier velocity, dx is the minimum grid size, dt is the time step, and cfl is a number which relates the grid size and time step. For the BTE, the cfl number must be less than 0.5 for numerical stability.

The numerical solution of a transport equation such as the BTE also requires angular discretization within the sphere of solid angles to account for transport out of the two dimensional simulation plane. In these simulations, the S4, S6, and S8 discrete ordinate approximations are used, corresponding to 12, 24, and 40 angular directions, respectively. The run times and relative accu-

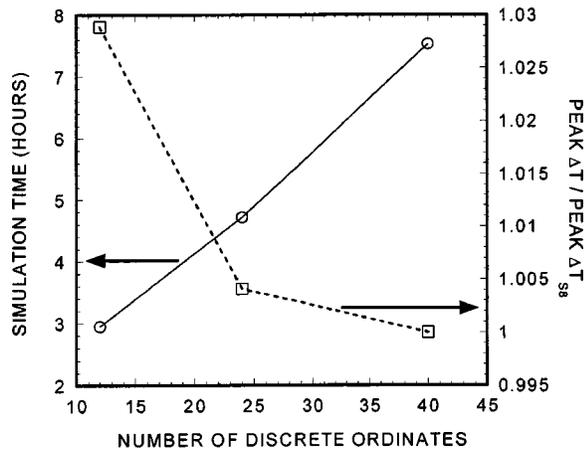


Fig. 4 Numerical characteristics of discrete ordinates approximations

racy are plotted in Fig. 4 and indicate that the S6 approximation is sufficient for reasonable numerical accuracy and run time.

The BTE code was numerically validated at both of its limits: the diffusion limit in which the phonon mean free path, Λ , is much smaller than the domain size and the near ballistic regime in which Λ is greater than the domain size. Excellent agreement is observed between the coupled BTE/diffusion solver and the solution obtained using a diffusion solver throughout the domain. The near ballistic limit can be verified by making comparisons with analytical solutions from radiation theory. Heaslet and Warming [36] calculated the ratio of the heat flux between two parallel plates compared to the heat flux from blackbody emission at the two wall temperatures when the mean free path of the radiation exceeds the plate separation. This is analogous to comparing the heat flux from the BTE to the heat flux from diffusion theory. Excellent agreement is obtained between the BTE code and the analytical solution for all mean free path to film thickness ratios.

5 Results and Discussion

In order to compare the temperature fields predicted by device simulators and the BTE code in a relevant device geometry, a model for a SOI device was generated using the semiconductor process simulator, TSUPREM4 [13]. This simulator models the processing steps required to fabricate transistors, including oxidation, diffusion, etching, and ion implantation. The result is a SOI device structure in which the silicon layer thickness, d_{Si} , is 72 nm, the buried oxide layer thickness, d_{ox} , is 243 nm, and the

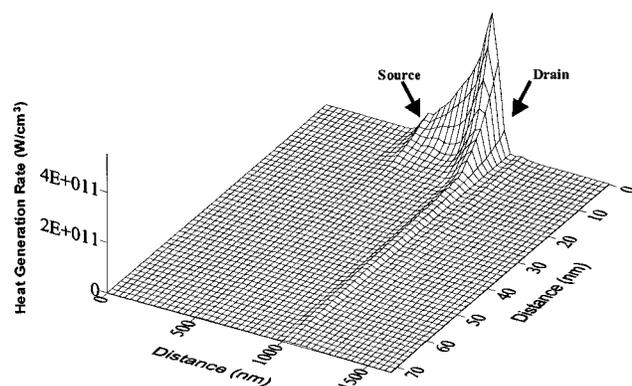


Fig. 5 SOI device heat generation rate caused by strong electron-phonon scattering. This quantity is extracted from an electrical simulation in MEDICI.

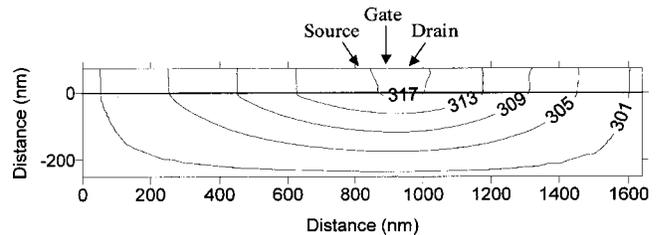


Fig. 6 Temperature contours in a SOI device calculated using the heat diffusion equation and bulk properties in MEDICI

channel length, L_c , is approximately 400 nm. Each of these dimensions is comparable with the mean free path in silicon, indicating that boundary scattering effects will be important.

This SOI device structure is then imported into the semiconductor device simulator, MEDICI. This simulator couples electrical transport with the heat equation to model both electrical and thermal characteristics. For this simulation, the gate and drain of the SOI transistor are biased at $V_G=2$ V and $V_D=2$ V, respectively, allowing current to flow through a thin conducting channel between the source and drain of the transistor. The electric field peaks near the drain junction of the device and imparts a large thermal energy to electrons in this region. These hot electrons collide with the lattice through electron-phonon scattering and transfer their energy into the phonon system, which conducts the heat out of the device. In the steady state, the heat generation rate (W/m^3) can be extracted as the dot product of the current density, $J_{current}$, and the electric field, E_{field}

$$q''' = q'_{\text{electron-phonon}} = \vec{J}_{\text{current}} \cdot \vec{E}_{\text{field}} \quad (26)$$

The extracted heat generation rate is shown in Fig. 5. The heat generation is peaked on the drain side of the transistor. The conducting channel through which the electrical current flows is limited to the first few nanometers of silicon below the gate electrode. Consequently, the region of heat generation is also confined to this region, yielding a localized heat source, which has characteristic dimensions much smaller than the mean free path of phonon energy carriers. The small dimensions of the heat source are expected to increase the temperature rise over that predicted by diffusion theory [11].

Figure 6 plots the temperature contours in the silicon layer and buried oxide layer obtained from the heat diffusion equation and bulk thermal properties in MEDICI. The peak temperature in the device is 317.8 K and is located near the drain junction of the device. The isotherms in the silicon layer are nearly perpendicular to the silicon/oxide interface indicating that the silicon layer is nearly isothermal in the out-of-plane direction and that heat is predominantly transported along the plane of the silicon layer. This is expected because of the relatively large thermal impedance of the buried oxide layer.

The temperature contours within the SOI device from the coupled BTE/heat equation code are presented in Fig. 7. The temperature contours have the same characteristic shape, indicating

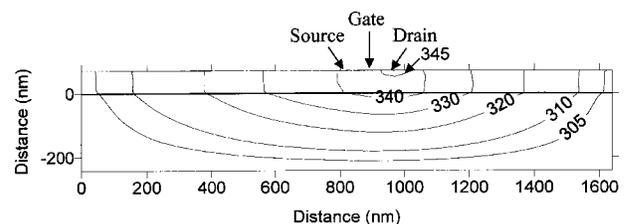


Fig. 7 Temperature contours in a SOI device calculated using the coupled BTE/heat diffusion equation code

that heat is flowing laterally out of the drain junction region. However, the boundary scattering and the effect of a heat source that is smaller than the phonon mean free path result in a peak device temperature of 346.0 K, an increase in the temperature rise of 159 percent over the value predicted by MEDICI. The entire channel region has a significantly higher temperature rise, which will affect electrical transport in the device.

The 317 K temperature contour in the MEDICI solution of Fig. 6 is 0.8 K from the peak temperature in the device and consists of straight lines that extend across the entire silicon layer thickness. In contrast, the 345 K temperature contour in the BTE solution is more curved and does not reach the oxide even though it is 1 K from the peak device temperature. This indicates that there is more temperature uniformity in the vicinity of the heat source region in the diffusion solution as compared to the more localized hot spot of the BTE solution. The increased temperature rise is caused by the fact that the heat source is smaller than the phonon mean free path. This effect is confined to a distance approximately two mean free paths from the edge of the heat source. This can be understood with the help of the survival function from kinetic theory

$$N = N_o \exp\left(-\frac{x}{\Lambda}\right) \quad (27)$$

which states that only a small fraction, N/N_o , of phonons generated in the heated region can traverse a distance much larger than the mean free path without having a collision. This small heat source effect is expected to have a large impact on short time-scale transient temperature rise under brief electrical stressing.

As stated in the introduction, Sondheimer [10] has developed an analytical solution to the BTE that can account for the reduction in thermal conductivity by boundary scattering. This effective thermal conductivity is appropriate for SOI devices because Sondheimer's assumptions for the electrical problem are equivalent to assuming that heat flows only laterally within the silicon layer. The reduced thermal conductivity given by Sondheimer's relation was inserted into MEDICI, and the resulting peak temperature rise was 339.6 K. This temperature is still 16.3 percent smaller than the peak temperature rise calculated using the coupled BTE/diffusion code because it does not account for the small dimension of the heat source.

The solutions from the BTE/diffusion code can also be compared to experimentally measured temperatures in SOI devices in order to explain the observed thermal behavior. Precise comparison of temperature distributions with experimental devices is not possible at present because experimental techniques do not exist for measuring temperature distributions in sub-micron transistors. Average temperature rise in the channel region can be measured and can be used for comparison. Goodson et al. [5] measured the average channel temperature in SOI field-effect transistors using electrical resistance thermometry. The device structure differs from that simulated in this paper; consequently, direct comparisons are not possible for an additional reason not related to experimental methods. The processing conditions and doping profiles used in the study of Goodson et al. [5] are unknown, making it difficult to construct an accurate representation of the device in a process simulator. However, the experimental results for a 78 nm thick silicon layer fall within the temperature rise range observed in this numerical study at the simulated power of 0.6 mW/ μm . The study of Goodson et al. [5] yielded temperature rise data that were somewhat larger than their predictions, which were calculated using diffusion theory. This discrepancy could possibly be explained by the phonon non-equilibrium effect examined in this paper and by phonon boundary scattering, but might also result from the uncertainties in the dimensions and impurity concentrations for those devices, which influence predictions. More research needs to examine the phonon non-equilibrium effect in dedicated experimental structures, which resemble transistors and

provide the opportunity to diminish the uncertainty in measurements of the temperature rise, the device dimensions, and the impurity concentrations.

6 Summary and Conclusions

This work presents a simulation approach for sub-continuum heat transfer in a SOI device geometry. The simulation couples the phonon BTE and the heat diffusion equation. The solution of the phonon BTE captures the impact of boundary scattering and small dimensions of heat sources on the peak temperature rise in semiconductor devices. The results from the coupled BTE/heat diffusion equation simulation are compared with temperature distributions obtained using the existing heat diffusion capabilities of the commercial device simulator MEDICI. For a SOI device in which the silicon layer thickness, d_{sl} , is 72 nm, the buried oxide layer, d_{ox} , is 243 nm, and the channel length, L_c , is approximately 400 nm, there is a 159 percent discrepancy in the peak steady-state temperature rise in the device between calculations made with the heat diffusion equation and calculations that consider micro-scale heat transfer phenomena. New thermal conductivity models must be developed that can account for both boundary scattering and the small dimensions of the heat source with the diffusion equation. More detailed calculations accounting for the combined behavior of optical, transverse, and longitudinal acoustic phonons can be considered in future work. Optical and high frequency transverse phonons will be grouped into a reservoir or capacitive mode with zero group velocity, and the longitudinal acoustic phonons will be characterized as the propagating mode. Interactions between the two modes will yield solutions appropriate for short transient high current electrical pulses. The solution method developed here for the two-dimensional BTE will be used for studies of the thermal characteristics of short time-scale electrostatic discharge events in which accurate prediction of temperature fields is currently not possible.

Acknowledgments

The authors appreciate technical input and comments from Tristan Burton, Mehdi Asheghi, William King, Robert Jacobs, and Younes Shabany, all from Stanford University. The authors also appreciate financial support from Texas Instruments through a fellowship for Per Sverdrup and from the Semiconductor Research Corporation through contract SJ-461.

Nomenclature

- C = volumetric heat capacity, $\text{J}/\text{m}^3\text{K}$
- $D(\omega)$ = phonon density of states per unit volume, $1/\text{m}^3$
- d_{ox} = buried oxide layer thickness, m
- d_{sl} = silicon layer thickness, m
- E = phonon energy, J
- E_b = blackbody energy flux, W/m^2
- E_{field} = electric field, V/m
- e'' = phonon energy per unit volume per unit solid angle, $\text{J}/(\text{m}^3\text{-sr})$
- e''_{eq} = equilibrium phonon energy per unit volume per unit solid angle, $\text{J}/(\text{m}^3\text{-sr})$
- e'_{eq} = equilibrium phonon energy per unit volume, J/m^3
- f = phonon distribution function
- f_{eq} = equilibrium phonon distribution function, Planck distribution
- G = irradiation of silicon/oxide interface from silicon, W/m^2
- \hbar = Planck's constant divided by $2\pi = 1.054e^{-34}$ Js
- I = phonon intensity, W/m^2
- J = flux of energy leaving silicon/oxide interface into silicon, W/m^2
- J_{current} = current density, A/m^2

k_b = Boltzmann constant = $1.38e^{-23}$ J/K
 k_{ox} = oxide thermal conductivity, W/(m·K)
 L_c = transistor channel length, m
 N = number of surviving phonons
 N_0 = initial number of phonons leaving a point in space
 q = energy flux into oxide, W/m²
 q''' = heat generation rate, W/m³
 $q'_{\text{electron-phonon}}$ = source representing electron-phonon scattering, W/m³
 $q_{\text{electron-phonon}}$ = source representing electron-phonon scattering, 1/s
 T = temperature, K
 T_{ox} = oxide interface temperature, K
 T_{ref} = reference temperature, K
 t = time, s
 V_D = transistor drain voltage, V
 V_G = transistor gate voltage, V
 v = phonon velocity, m/s
 v_{gr} = phonon group velocity, m/s
 v_{ph} = phonon phase velocity, m/s
 w_i = discrete ordinate weights, steradians
 x, y = coordinate, m

Greek Symbols

α = transmission coefficient
 α_t = thermal diffusivity, m²/s
 ε = emissivity
 Λ = frequency-averaged phonon mean free path, m
 μ, η = direction cosine
 Ω = solid angle, steradians
 ρ = reflectivity
 τ_{phonon} = phonon relaxation time
 ω_d = phonon cutoff frequency, rad/s

Subscripts

b = blackbody
 c = channel
 D = transistor drain
 d = Debye
 eq = equilibrium
 G = transistor gate
 gr = group
 i = index number
 o = initial value
 ox = oxide
 ph = phase
 $prop$ = propagating mode
 ref = reference value
 si = silicon
 t = thermal
 η^+ = positive y-direction, into silicon
 η^- = negative y-direction, into oxide

References

- [1] Majumdar, A., 1993, "Microscale Heat Conduction in Dielectric Thin Films," *ASME J. Heat Transfer*, **115**, pp. 7–16.
- [2] Ju, Y. S., and Goodson, K. E., 1999, "Phonon Scattering in Silicon Films with Thickness of Order 100 nm," *Appl. Phys. Lett.*, **74**, pp. 3005–3007.
- [3] Lai, J., and Majumdar, A., 1996, "Concurrent Thermal and Electrical Modeling of Sub-Micrometer Silicon Devices," *J. Appl. Phys.*, **79**, pp. 7353–7361.
- [4] Raha, P., Ramaswamy, S., and Rosenbaum, E., 1997, "Heat Flow Analysis for EOS/ESD Protection Device Design in SOI Technology," *IEEE Trans. Electron Devices*, **44**, pp. 464–471.
- [5] Goodson, K. E., Flik, M. I., Su, L. T., Antoniadis, D. A., 1995, "Prediction

- and Measurement of Temperature Fields in Silicon-on-Insulator Electronic Circuits," *ASME J. Heat Transfer*, **117**, pp. 574–581.
- [6] Wolbert, P. B., Wachutka, G. K., Krabbenborg, B. H., and Mouthaan, T. J., 1994, "Nonisothermal Device Simulation Using the 2-D Numerical Process/Device Simulator TRENDY and Application to SOI-Devices," *IEEE Trans. Comput.-Aided Des.*, **13**, pp. 293–302.
 - [7] Yu, Z., Chen, D., So, L., and Dutton, R. W., 1994, "PISCES-2ET—Two Dimensional Device Simulation for Silicon and Heterostructures—Device Simulation Manual," Stanford University, Stanford, CA.
 - [8] Chen, G., 1997, "Size and Interface Effects on Thermal Conductivity of Superlattices and Periodic Thin-Film Structures," *ASME J. Heat Transfer*, **119**, pp. 220–9.
 - [9] Goodson, K. E., 1996, "Thermal Conduction in Nonhomogeneous CVD Diamond Layers in Electronic Microstructures," *ASME J. Heat Transfer*, **118**, pp. 279–286.
 - [10] Sondheimer, E. H., 1952, "The Mean Free Path of Electrons in Metals," *Adv. Phys.*, **1**, No. 1.
 - [11] Chen, G., 1996, "Nonlocal and Nonequilibrium Heat Conduction in the Vicinity of Nanoparticles," *ASME J. Heat Transfer*, **118**, pp. 539–545.
 - [12] Cahill, D. G., 1997, "Heat Transport in Dielectric Thin Films and at Solid-Solid Interfaces," *Microscale Thermophys. Eng.*, **1**, pp. 85–109.
 - [13] Avant! Corporation, 1999, *Medici: Two-Dimensional Semiconductor Device Simulation and TSUPREM-4: Two-Dimensional Semiconductor Process Simulation*, Fremont, CA.
 - [14] Duvvury, C., Amerasekera, A., Joyner, K., Ramaswamy, S., and Young, S., 1996, "ESD Design For Deep Submicron SOI Technology," 1996 Symposium on VLSI Technology, pp. 194–195.
 - [15] Lundstrom, M., 1998, "Scattering Theory of Carrier Transport in Semiconductor Devices," *SISPAD 1998*, p. 215–222.
 - [16] Blotekjaer, K., 1970, "Transport Equations for Electrons in Two-Valley Semiconductors," *IEEE Trans. Electron Devices*, **ED-17**, No. 1, pp. 38–47.
 - [17] Buot, F. A., and Slegler, K. J., 1984, "Numerical Simulation of Hot-Electron Effects on Source-Drain Burnout Characteristics of GaAs Power FETs," *Solid-State Electron.*, **27**, No. 12, pp. 1067–1081.
 - [18] Mogilestue, C., Buot, F. A., and Anderson, W. T., 1995, "Ensemble Monte Carlo Particle Investigation of Hot Electron Induced Source-Drain Burnout Characteristics of GaAs Field Effect Transistors," *J. Appl. Phys.*, **78**, No. 4, pp. 2343–2348.
 - [19] Majumdar, A., 1998, *Microscale Energy Transport*, C.-L. Tien et al., eds., Taylor & Francis, New York.
 - [20] Wachutka, G. K., 1990, "Rigorous Thermodynamic Treatment of Heat Generation and Conduction in Semiconductor Device Modeling," *IEEE Trans. Comput.-Aided Des.*, **9**, pp. 1141–1149.
 - [21] Ju, Y. S., 1999, "Microscale Heat Conduction in Integrated Circuits and their Constituent Thin Films," Ph.D. thesis, Stanford University, Stanford, CA.
 - [22] Joshi, A., and Majumdar, A., 1993, "Transient Ballistic and Diffusive Phonon Heat Transport in Thin Films," *J. Appl. Phys.*, **74**, pp. 31–39.
 - [23] Armstrong, B., 1981, "Two-Fluid Theory of Thermal Conductivity of Dielectric Crystals," *Phys. Rev. B*, **23**, No. 2, pp. 883–899.
 - [24] Ju, Y. S., and Goodson, K. E., 1997, "Impact of Phonon Dispersion upon the Size Effect on Thermal Conduction along Thin Semiconductor Films," *Proc. ASME International Mechanical Engineering Congress and Exposition*, Dallas, TX, **DSC-Vol. 62**, pp. 181–190.
 - [25] Lewis, E., 1984, *Computational Methods of Neutron Transport*, Wiley, New York.
 - [26] Fiveland, W. A., 1991, "The Selection of Discrete Ordinate Quadrature Sets for Anisotropic Scattering," *Fundamentals of Radiation Heat Transfer*, HTD-Vol 160, AMSE New York, NY, pp. 89–96.
 - [27] Siegel, R., and Howell, J., 1992, *Thermal Radiation Heat Transfer*, 3rd ed., Taylor & Francis, Washington, D.C.
 - [28] Glass, D., Ozisik, M., and McRae, D., 1985, "On the Numerical Solution of Hyperbolic Heat Conduction," *Numer. Heat Transfer*, **8**, pp. 497–504.
 - [29] Holland, M. G., 1963, "Analysis of Lattice Thermal Conductivity," *Phys. Rev.*, **132**, pp. 2461–2471.
 - [30] Kreysig, E., 1993, *Advanced Engineering Mathematics*, 7th ed., Wiley, New York, pp. 1061–1063.
 - [31] Little, W. A., 1959, "The Transport of Heat Between Dissimilar Solids at Low Temperatures," *Can. J. Phys.*, **37**, pp. 334–349.
 - [32] Swartz, E. T., and Pohl, R. O., 1989, "Thermal Boundary Resistance," *Rev. Mod. Phys.*, **61**, pp. 605–668.
 - [33] Phelan, P. E., 1998, "Application of Diffuse Mismatch Theory to the Prediction of Thermal Boundary Resistance in Thin-Film High-T_c Superconductors," *ASME J. Heat Transfer*, **120**, pp. 37–43.
 - [34] Stephens, R. B., 1973, "Low-Temperature Specific Heat and Thermal Conductivity of Noncrystalline Solids," *Phys. Rev. B*, **8**, pp. 2896–2905.
 - [35] Kittel, C., 1996, *Introduction to Solid State Physics*, Wiley, New York.
 - [36] Heaslet, M. A., and Warming, R. F., 1965, "Radiative Transport and Wall Temperature Slip in an Absorbing Planar Medium," *Int. J. Heat Mass Transf.*, **8**, pp. 979–984.