Review: Multiscale Thermal Modeling in Nanoelectronics

S. Sinha^{*} & K. E. Goodson

Thermosciences Division, Mechanical Engineering Department Stanford University, CA 94305-3030

ABSTRACT

Subcontinuum phonon conduction phenomena impede the cooling of field-effect transistors with gate lengths less than 100 nm, which degrades their performance and reliability. Thermal modeling of these nanodevices requires attention to a broad range of length scales and physical phenomena, ranging from continuum heat diffusion to atomic-scale interactions and phonon confinement. This review describes the state of the art in subcontinuum thermal modeling. Although the focus is on the silicon field-effect transistor, the models are general enough to apply to other semiconductor devices as well. Special attention is given to the recent advances in applying statistical and atomistic simulation methods to thermal transport.

*Corresponding author: sanjiv@stanford.edu

1. INTRODUCTION

Ever since Dennard et al. [1] demonstrated that faster semiconductor devices could be fabricated by reducing device dimensions such that the electric field remained constant, device scaling has been the driver in VLSI technology. In keeping with Gordon Moore's 1965 forecast [2], semiconductor devices have continuously diminished in size over the past four decades yielding larger scale integration and faster computing. However, traditional device scaling faces new challenges in the near future as the minimum device feature length decreases below 100 nm. Besides the complexity of designing devices at these length scales, the twin issues of power dissipation and heat removal emerge as dominant concerns. The fundamental unit of an integrated circuit, the metal-oxide semiconductor field-effect transistor (MOSFET) produces nearly 0.5 mW of Joule heat per unit micrometer width in the current 90 nm gate Heat generation occurs length technology. within the inversion layer at the interface of The f and g processes are shown in a (110)-plane to identify the wave vectors of the involved phonons. The f-process phonon has the wave vector \mathbf{q}_f and the *g*-process phonon has the wave vector \mathbf{q}_{q} purely from geometry considerations.the channel region and the drain terminal, over a length of around 20 nm, with the peak power density approaching 5 W/ μ m³ [3]. At this length scale, the assumption of continuum heat conduction breaks down. Instead, conduction is described in terms of the energy quanta of lattice vibrations called phonons. Understanding thermal transport at the device level is becoming increasingly important as future technology evolves toward devices with larger internal thermal resistances and higher heat generation per unit volume. Figure 1 shows a schematic of the traditional bulk silicon MOSFET as well as various exploratory candidates for future commercial nanotransistors. Compared to a bulk device, novel devices such as the strained silicon FET and the various silicon-on-insulator (SOI) [4,5] devices offer better electrical performance, though at the cost of a higher thermal resistance. This is more so in the case of more revolutionary designs such as the three-dimensional FINFET [6].

A digital circuit is typically designed such that the peak transistor temperature does not exceed 80-100°C [7]. In current bulk silicon devices, the temperature rise is assumed to be largely dictated by heat transfer outside the transistor. The thermal resistances due to the silicon substrate, the thermal interface material (TIM), the external heat sink, and convection to the ambient add up to about 0.5 K/W [8]. However, for nanotransistors with dimensions as small as several tens of nanometers, thermal resistances arising from subcontinuum phenomena inside the device become equally important. By using an undoped silicon film as the transistor channel, SOI devices such as the dual gate transistor [4], the ultrathin-body transistor [5], and the FINFET [6] offer better current control due to improved electrostatics and reduced short-channel effects. However, the buried oxide layer increases the thermal resistance of the device, resulting in a relatively higher temperature rise compared to a bulk device for the same power dissipated. Further, the spatial extent of the heat source can be as small as 4 nm (full width at half maximum) in a nanotransistor, as shown in Fig. 2, and the peak power density may approach 60 W/ μ m³. Thermal measurements near such highly localized heat sources [9] indicate that heat conduction is impeded in the vicinity of the source due to ballistic phonon transport. The heat source in a nanotransistor exhibits similar behavior by forming a localized phonon hotspot with a high thermal resistance. Thus, innovations on the electrical side invariably lead to a poor thermal design, augmenting the thermal resistance internal to the device. Performance optimization



FIGURE 1. The schematics show various FET device structures starting from the bulk silicon MOSFET on the top left. The locations of the source (S), drain (D) and gate (G) terminals are indicated as is location of the buried oxide (BOX) layer. The stained silicon FET uses a strained silicon layer as the channel for electron flow, which enhances the mobility of electrons. The use of dual gates and an ultrathin film silicon channel enables better current control in short-channel devices. The FINFET is a three-dimensional structure with the gate surrounding a vertically oriented fin that serves as the channel



FIGURE 2. The location of the heat source is shown in an ultrathin body SOI transistor with a gate length of 18 nm [7]. The contours correspond to the spatial distribution of the heat generated inside the device and are equispaced at $0.5 \text{ W/}\mu\text{m}^3$ with a peak value of $60 \text{ W/}\mu\text{m}^3$ at the center

requires electrothermal models that incorporate microscopic electron-phonon and phononphonon interactions.

In this paper, we review heat conduction modeling in semiconductor devices at multiple length scales, from the lattice spacing at a few tenths of a nanometer to the phonon mean free path at a few hundred nanometers. The focus is largely on research from the past decade on silicon nanoelectronics, though thermal modeling in gallium arsenide and other optoelectronic devices is also referenced where appropriate. In particular, we have chosen to ignore more exotic nanoelectronic devices such as molecular devices, quantum-effect devices and single electron devices [10], the reason being that thermal aspects of these devices remain of secondary concern and have received very little attention. We begin with an overview of subcontinuum heat conduction in Section 2. In Section 3, we categorize different approaches to modeling heat transport in nanostructures. Investigations of electrothermal transport in semiconductor devices are reviewed in Section 4. The issue of modeling coupled length scales is discussed briefly in Section 5. The review concludes in Section 6 with a summary of some of the outstanding issues in modeling thermal transport at nanoscales, as well as some of the new issues posed specifically by research on nanoscale semiconductor devices.

2. LATTICE HEAT CONDUCTION

Atoms located at the lattice sites of a dielectric crystal undergo small oscillations about their equilibrium positions at every temperature. The resulting atomic displacement field in the crystal stores and transports energy. Classically, this is described by a displacement vector obeying a linear, homogeneous wave equation of second order in space and time. Thus, the field may be described as superpositions of plane waves, with the frequency and wave vector obeying a nonlinear dispersion relationship. These so-called normal modes of the crystal can have multiple polarizations, depending on whether the displacement is perpendicular (transverse) or parallel (longitudinal) to the wave vector. In the long wavelength limit, the longitudinal vibrations are essentially identical to sound waves in a solid. Quantum mechanically, the atomic displacement field may be described either as an infinite number of distinguishable, quantized oscillators, or as a gas of indistinguishable particles called phonons. The particle description is particularly useful in treating interactions with other systems such as electrons. In thermal equilibrium, phonons are described by the Planck distribution, which applies to bosons without rest mass and without conservation of particle number.

The calculation of the heat capacity and the thermal conductivity of a solid requires knowledge of the phonon dispersion relationship. The formalism for obtaining the dispersion relationship employs the concept of a crystal potential function to describe the chemical binding of atoms. The justification for this was first provided by Born and Oppenheimer in the context of molecular motion and is referred to as the adiabatic approximation [11]. The argument is based on the fact that the frequencies of atomic motions are much smaller than the electronic transition frequencies. The electronic system continuously adjusts itself without changing its quantum number as atoms vibrate about their mean positions. This allows the use of an effective crystal potential without solving for the electron system. In addition to the adiabatic approximation, a second approximation known as the harmonic approximation is made in computing the dispersion relationship. The crystal potential is expanded in terms of the powers of the amplitudes of the vibrations and all terms beyond the second order are neglected, in effect making the interatomic "spring" harmonic. The atomic motions are now described by an eigensystem whose eigenvalues are related to the phonon frequencies. Solution of this eigensystem constitutes the problem of lattice dynamics [12]. A bulk crystal may be assumed to be infinite in extent and the Born-von Karman periodic boundary conditions can be applied to further simplify the equations. Figure 3 shows the phonon dispersion in silicon along the directions of high symmetry in the crystal. Silicon has a diatomic basis which leads to phonon modes in which atoms in a primitive cell vibrate out of phase. These are referred to as optical phonons. The vibrations in which the basis atoms move in phase are referred to as acoustic phonons. Transverse (T) and longitudinal (L) polarizations of optical (O) and acoustic (A) phonon branches are indicated in the figure.

The harmonic approximation, though useful in calculating the phonon dispersion relationship, does not capture the physics of phononphonon interactions and scattering. If the crystal potential were indeed harmonic, phonons in an infinite and perfect crystal would have unlimited lifetimes, resulting in infinite thermal conductivity. We know conclusively that this is not true even at very low temperatures. Additionally, a harmonic crystal would not show any thermal expansion. In a real crystal, atomic interactions are anharmonic, causing the phonons to scatter with each other, even in a defect-free crystal. The lowest-order scattering process involves three phonons, as shown in Fig. 4. A phonon traveling in the crystal can annihilated to create two other phonons. be The phonon is described in the schematic by its angular frequency ω , which is a function of its wave vector q and its polarization s. An inverse process is equally possible from the principle of microscopic reversibility. Higherorder processes, such as one involving four phonons, are also possible. All such processes must ensure microscopic conservation of en-



FIGURE 3. The phonon dispersion relationship is shown along directions of high symmetry in silicon. Longitudinal and transverse polarizations are indicated by L and T respectively. The letters O and A refer to optical and acoustic branches, respectively. The slope of the dispersion curve gives the group velocity of the particular phonon mode



FIGURE 4. Anharmonic interactions between phonons leads to phonon-phonon scattering. The lowest-order process involves three phonons as shown schematically on top. Higher-order processes, such as one involving four phonons (bottom), are also possible

ergy and crystal momentum. Anharmonicity is treated as a small perturbation to the harmonic Hamiltonian for the crystal, and the lifetimes of phonons are computed using time-dependent quantum mechanical perturbation theory [13].

The level of detail that a heat conduction model must include depends on the length scale of the problem and the temperature. The phonon distribution obeys quantum statistics at low temperatures, but approaches the classical limit as the lattice temperature approaches the Debye temperature θ_D . Thus, quantum mechanical details that must be accounted for at very low temperatures become less important

at higher temperatures. The length scale of the problem determines whether phonons may be modeled as particles or whether the wave nature of phonons is important. We note here that the latter does not necessarily imply a quantum mechanical description. In fact, the classical description of crystal vibrations involves lattice waves. The particle behavior comes from the quantum nature of these waves. The important length scales for comparison are the phonon mean free path Λ , the phonon wavelength λ , and the lattice spacing a_0 . The validity of various modeling regimes is summarized in Fig. 5. The diffusion length is included for



FIGURE 5. Subcontinuum thermal modeling depends on the dimension of the system and the temperature. At dimensions comparable to the phonon wavelength and temperatures much smaller than the Debye temperature (bottom left), quantum mechanical nature is strongly manifest. At larger dimensions and room temperatures, a semiclassical approach is more pragmatic. PT in the figure stands for perturbation theory and α is the thermal diffusivity

the sake of completeness. Typical silicon semiconductor device dimensions lie in between the phonon mean-free path and the dominant phonon wavelength in room-temperature silicon. As shown in the figure, semiclassical or classical models are applicable in this regime. When domain dimensions are comparable to the phonon wavelength and the temperature is small compared to the Debye temperature, a quantum mechanical model is necessary. Finally, the extent of the crystal can alter the phonon dispersion relation [14] and the density of phonon states. As dimensions of the crystal are reduced, surface effects become important and the assumption of an infinite crystal does not hold along one or more dimensions. We review the different approaches indicated in Fig. 5 in the following sections.

3. MODELING PHONON TRANSPORT IN NANOSTRUCTURES

At nanoscale dimensions, it becomes possible to engineer electron and phonon transport by controlling the physical dimensions of the structures, which influences available states. Examples of nanostructures include quantum wells, quantum dots, and superlattices, which are common in optoelectronics and are currently being investigated for thermoelectric applications [15]. Silicon transistors, in contrast, did not start out by exploiting quantum mechanical length scales but are nevertheless approaching such dimensions through device scaling [7]. Since phonon wavelengths dominating heat conduction in silicon are approximately 1-2 nm at room temperature, a full quantum treatment of phonons is rarely warranted. In this section, we review general semiclassical and classical methods for modeling phonon transport. Semiclassical modeling involves either the use of the phonon Boltzmann transport equation (BTE) [16] with varying details of the phonon dispersion relationship, or the Monte Carlo (MC) [17] technique to statistically simulate phonon transport. Another choice is molecular dynamics [18], which is classical in origin and involves the use of statistical mechanics to compute transport coefficients. Molecular dynamics does not assume the validity of the BTE, but can be used to compute physical parameters needed by the BTE. Figure 6 summarizes the hierarchy of semiclassical thermal modeling and highlights the main features of each approach. We focus on the methods in this section, and discuss their application to semiconductor device problems in the next section.

3.1 Boltzmann Transport Equation

The phonon Boltzmann transport equation describes the rate of change of a statistical distribution function for phonons. The fundamental assumption in deriving the phonon BTE is that there exists a distribution function, $N_{\mathbf{q},s}(\mathbf{r}, t)$, which describes the average occupation of the phonon mode (\mathbf{q}, s) in the neighborhood of a location r at a time t. The equation assumes the simultaneous prescription of phonon position and momentum with arbitrary precision. However, in quantum mechanics these quantities correspond to noncommuting operators and, hence, obey the uncertainty principle. The semiclassical BTE for phonons resolves this contradiction by treating phonons as wave packets [19] (see Fig. 7 for example) formed from a superposition of normal modes of the crystal. Each wave packet has a small spread δq in the wave vector and is localized in space in a region of size δr such that $\delta q \delta r \sim 1$. The wave packet is constructed such that its group velocity corresponds to the slope of the phonon dispersion relationship. For a detailed discussion on the validity of the wave packet picture, we refer the reader to Ref. [19]. In the semiclassical model, the number of phonons becomes a function not only of the wave vector q and po-



FIGURE 6. Sub-continuum thermal modeling uses either the phonon BTE or the MC/MD methods. Though all approaches include the role of phonon dispersion, the exact dispersion relationship taken into account differs between various approaches



FIGURE 7. A phonon wave packet [3] created from a linear combination of the eigenmodes is shown. The longitudinal displacements of individual atoms at different axial locations are shown on the left and the corresponding energy distribution is shown on the right

larization *s*, but also of the spatial coordinate **r**. The distribution function evolves in time in a six-dimensional phase space due to motion in real space, due to anharmonic scattering with other phonons, and due to scattering with impurities and boundaries. The BTE is formally written as

$$\frac{\partial N_{\mathbf{q},s}}{\partial t} + \mathbf{v}_{\mathbf{q},s} \cdot \nabla N_{\mathbf{q},s} = \frac{\partial N_{\mathbf{q},s}}{\partial t} \bigg|_{c} \qquad (1)$$

where (q,s) refers to the phonon mode with wave vector q and polarization s, **v** is the group velocity, N is the phonon occupation number, and the term on the right-hand side is the rate of change due to collisions.

The solution to the phonon BTE requires evaluation of the collision term. Although Peierls [19] obtained the formal expression as an integral in wave vector space over all allowed scattering processes, it is too complicated to be directly evaluated in practice. Instead, the relaxation time approximation has been widely used to model the collision term. Under this approximation, the BTE is rewritten as

$$\frac{\partial N_{\mathbf{q},s}}{\partial t} + \mathbf{v}_{\mathbf{q},s} \cdot \nabla N_{\mathbf{q},s} = -\frac{N_{\mathbf{q},s} - \overline{N}}{\tau_{\mathbf{q},s}} \qquad (2)$$

where \overline{N} is the Planck distribution at the local temperature and τ is the relaxation time of the mode (q,s). Thermal conductivity models based on the solution to the phonon BTE under the relaxation time approximation are available in the literature [20–23]. In a more rigorous approach, Guyer and Krumhansl [24] used collision operators to represent normal (elastic) and umklapp (inelastic) scattering and obtained the formal solution for thermal conductivity in terms of the matrix elements of the operators. The relaxation times derived in Refs. [20–23] appear as special limiting cases in the full operator solution. However, the computation

of operators is cumbersome and has not been adopted in solving device problems so far. The relaxation time approximation has emerged as the default choice in almost all of the literature on microscale heat conduction. The derivations of the relaxation times involve time-dependent perturbation theory from quantum mechanics (see Ref. [13], for example). Klemens [20] and others [22,23] have developed semiempirical expressions for the relaxation rates that model thermal conductivity data over a wide range of temperatures. It is important to note that in the original work [23], the group velocity has been used interchangeably with the phase velocity. The expressions must be corrected when using a realistic nonlinear dispersion relationship [25,26].

A key conceptual problem in using the relaxation time approximation is the requirement for a thermodynamic temperature that governs the scattering rate. Since phonons are not in an equilibrium distribution, there is no temperature to strictly speak of. The usual practice in such nonequilibrium problems is to define an *ad hoc* "equivalent" temperature based on the local energy [27–30]. We note that a Planck distribution based on the equivalent temperature does not represent the local phonon population in most cases.

3.1.1 Equation of Phonon Radiative Transfer (EPRT)

The phonon occupation can be solved to first order by linearizing the gradient term in the BTE. The solution is of the form

$$N = \overline{N}(T) - \tau_{\mathbf{q},s} \mathbf{v}_{\mathbf{q},s} \cdot \nabla T \frac{\partial \overline{N}}{\partial T}$$
(3)

The second term in the solution is the firstorder departure from equilibrium and conforms to the Fourier heat flux law. The above first-order approximation works well, provided the temperature gradient is small enough so that the temperature does not vary appreciably over the relaxation length [20]. This is formally expressed as

$$\frac{\partial T}{\partial x}v\tau \ll T \tag{4}$$

and is usually valid except at length scales comparable to the phonon mean free path. Majumdar [27] pointed out the limited usability of the above solution in modeling transport in thin films and instead proposed the equation of phonon radiative transfer (EPRT). The EPRT is written in terms of the phonon intensity I_{ω} with the one-dimensional form being

$$\frac{1}{v}\frac{\partial I_{\omega}}{\partial t} + \mu \frac{\partial I_{\omega}}{\partial x} = -\frac{I_{\omega} - I_{\omega}^{0}(T(x))}{v\tau}$$
(5)

where μ is the direction cosine and v is the Debye velocity. The phonon intensity is given as

$$I_{\omega} = \sum_{s} v N(\omega(\mathbf{q}, s), x, t) \hbar \omega g(\omega)$$
 (6)

where $g(\omega)$ is the phonon density of states. In order to rigorously derive the EPRT from Eq. (2), the nonlinear phonon dispersion relationship of Fig. 3 must be replaced by the linear Debye model. The EPRT reproduces the expected radiative behavior in the acoustically thin limit and conforms to the Fourier heat flux law in the continuum limit.

Besides thin films [27,31,32], radiative phonon transport has been used to investigate interfacial transport [33], subcontinuum heat sources [34], and ballistic conduction from nanoparticles [28]. Prasher [35] recently extended the EPRT to include phase information in the in-scattering term on the right-hand side. However, the use of the Debye model limits the accuracy of the EPRT. While the Debye model is very accurate in predicting the specific heat capacity of a crystal, it is less accurate in describing transport. Nonlinearity in the phonon dispersion relationship of real crystals plays an important role in determining thermal conductivity.

3.1.2 Energy Moment of the Phonon BTE

An energy moment formulation of the BTE [30] involves taking a density-of-states-weighted frequency moment of the BTE. Using the Debye model, the moment equation is written as

$$\frac{\partial e''}{\partial t} + \mathbf{v} \cdot \nabla e'' = -\frac{e'' - e''_{EQ}}{\tau} \tag{7}$$

where e''_{EQ} and e'' are the equilibrium energy and the excess energy, respectively, per unit volume per unit solid angle. The excess energy is given by

$$e'' = \int \hbar\omega \left[N - \overline{N}(T_{\text{ref}}) \right] g(\omega) d\omega \qquad (8)$$

Equation 7 can be further simplified by assuming an isotropic distribution and integrating over the solid angle. The resulting local energy density can be expressed in terms of the volumetric heat capacity C and a nonequilibrium lattice temperature T_L , such that Eq. (7) reduces to [36,37]

$$\frac{\partial T_L}{\partial t} + \frac{1}{C} \mathbf{v} \cdot \nabla T_L = -\frac{T_L - T_{EQ}}{\tau} \qquad (9)$$

3.1.3 Ballistic-Diffusive Equations (BDE)

The ballistic-diffusive equations (BDE) [29] are based on the phonon BTE and involve an *a priori* description of the phonon distribution function. The BDE divides the distribution function N at any point into a ballistic, N_b , and a diffusive part, N_m . The ballistic part originates from the boundary and experiences outscattering only while the diffusive part originates inside the domain and evolves through in-scattering or through a source term. The equations are formally written as

$$\frac{1}{|\mathbf{v}|} \frac{\partial N_b}{\partial t} + \hat{\mathbf{\Omega}} \cdot \nabla N_b = -\frac{N_b}{|\mathbf{v}|\tau}$$
$$\frac{\partial N_m}{\partial t} + \mathbf{v} \cdot \nabla N_m = -\frac{N_m - \overline{N}}{\tau} \quad (10)$$

where Ω is the unit vector along the direction of propagation.

The BDE has been shown to reasonably approximate the phonon BTE while requiring less computations. However, the comparisons were made using an energy moment formulation under a gray-body approximation for phonons. It will be interesting to compare the BDE with the phonon BTE for a realistic phonon dispersion relationship. Another formulation that is conceptually close to the BDE but involves the nonlocal formulation described below, is the split-flux model [38]. This is discussed in Section 4.

3.1.4 Nonlocal Formulation

The nonlocal formulation [39,40] was developed for heat conduction problems where the phonon mean free path is comparable to domain dimensions or where the time scale of interest is comparable to the phonon relaxation times. Under such conditions, Eq. (4) is not satisfied and it is necessary to retain higher-order terms in the departure function. For simple domains, the departure function can be solved by direct integration of the BTE under the relaxation time approximation. The solution involves a nonlocal integral in space and a retardation function in time. The nonlocal formulation has the advantage that it does not require simplifying assumptions about the phonon dispersion relationship. The departure function can be integrated with full phonon dispersion information to yield the heat flux, the local energy density, and a nonequilibrium equivalent temperature field. Figure 8 compares the temperature fields computed for a one-dimensional



FIGURE 8. The temperature rise near a 40 nm long sheetlike symmetric heat source [38] is computed using the nonlocal formulation and compared to that from heat diffusion using Fourier law. The difference in the peak temperatures scales as the power density at the source

nanoscale phonon source [38] using the nonlocal formulation and the heat diffusion equation, respectively. While the nonlocal formulation provides more information than moments of the BTE by solving directly for the phonon distribution function, it becomes computationally expensive if the geometry is complex. It may be advantageous in such cases to opt for a Monte Carlo solution instead. This is discussed in the next subsection.

3.2 Monte Carlo Simulations

The Monte Carlo (MC) method is a statistical sampling technique [17] that can be applied to a wide variety of problems, from simulating neutron trajectories in a reactor to solving an integration in multidimensional space. In the MC method, a physical process may be simulated directly without requiring knowledge of the underlying governing equations. The sole requirement is that the process be describable in terms of probability density functions (PDFs). The simulation proceeds by sampling from the PDFs using a computer generated sequence of random numbers. The MC technique can be used to simulate thermal transport by a distribution of phonons. The method involves tracking the trajectories of individual phonons in phase space, given the PDFs associated with various scattering events.

While there has been rapid development in the MC method for electron transport in semiconductor devices, phonon MC is still in its infancy. This may be attributed partly to a lack of a compelling need to model subcontinuum transport in engineering problems. Another explanation is the difficulty of using MC to simulate particles whose number is not conserved. With the first reason fast vanishing, phonon MC is likely to draw more attention in the near future. Klistner et al. [41] employed the MC method to study radiative phonon transport in a crystal. Phonons were assumed to travel ballistically between the crystal surfaces and all scattering occurred at the interfaces. Peterson [42] reported phonon MC studies that included scattering within the crystal, but used a linear Debye dispersion.

Mazumder and Majumdar [43] developed a more comprehensive MC scheme for acoustic phonons that includes isotropic phonon dispersion. In this scheme, phonons are introduced into the crystal based on a probability dictated by the product of the phonon density of states and the Planck distribution function at a local temperature. Heat conduction in the crystal occurs due to the motion of the phonons at their individual group velocities. A boundary may specularly or diffusely scatter phonons as well as radiatively emit or absorb phonons. Phonon-phonon and phonon-impurity scattering within the bulk can annihilate a phonon. Energy conservation is the weak point in the scheme since details of microscopic interactions are overlooked. In the absence of microscopic energy conservation, an *ad hoc* global conservation scheme is used. In this scheme, the energy lost due to phonon annihilation is recovered by reintroducing phonons sampled from a Planck distribution based on the local energy density before annihilation. Despite this shortcoming, the scheme shows good agreement with data for thin-film thermal conductivity and is a very promising start to more detailed MC simulations. The power of the method lies in its ability to study the underlying physics of phonon transport, which cannot be accessed through the BTE methods described above.

Phonon-phonon scattering events are particularly difficult to model rigorously in a phonon MC scheme. To appreciate the complexity, one may consider the probability for the simplest phonon-phonon scattering event involving three phonons. Assuming a momentum conserving normal process [16], a phonon with frequency ω , wave vector **q**, and occupation N can be created from a combination of two phonons with frequencies ω' , ω'' ; wave vectors **q**', **q** – **q**'; and occupations N', N'', respectively. Using the time-dependent perturbation theory from quantum mechanics, the rate of transition from the initial to the final state is

$$W = \frac{2\pi}{\hbar} \left| \langle f | \Phi_3 | i \rangle \right|^2 \delta \left(E_f - E_i \right) \tag{11}$$

where *f* represents the final state, *i* represents the initial state, *E* is the energy, and Φ_3 is the third-order term in the expansion of the crystal potential. The sum total of all three-phonon interactions yields an expression for the net transition rate as follows [44]:

$$W = \frac{\hbar}{128\pi^2 M^3} \sum_{j'j''s} \int d\mathbf{q} \frac{\left| b_{jj'j''}(\mathbf{q}, \mathbf{q}', \mathbf{q} - \mathbf{q}') \right|^2}{\omega \omega' \omega''} \\ \times \left\{ \delta(\omega - \omega' - \omega'') \left[(N+1)N'N'' - N(N'+1)(N''+1) \right] + 2\delta(\omega + \omega' - \omega'') \right] \\ \left[(N+1)(N'+1)N'' - NN'(N''+1) \right] \right\}$$
(12)

where *M* is the atomic mass, j, j', j'' represent indices for the three phonons, and *b* is a matrix

element obtained from the third-order derivative of the potential. The integration is carried out over the Brillouin zone of the crystal. To use a rigorous probability for scattering that includes the microscopic conservation of energy and crystal momentum, transition rates such as Eq. (12) need to be evaluated at each time step in an MC calculation, which is clearly infeasible. The use of semiempirical relaxation rates [22,23] derived from thermal conductivity modeling is the current method in phonon MC. However, as pointed out in Section 2, these rates need to be carefully reexamined to remove some of the simplifying assumptions regarding the phonon dispersion relationship. An alternate method for evaluating relaxation rates is to use detailed molecular dynamics simulations. This is discussed in the next subsection.

3.3 Molecular Dynamics Simulations

Molecular dynamics (MD) refers to the solution of classical equations of motion (Newton's laws) for a set of molecules. In the context of phonon transport in solids, this amounts to solving the dynamics of atoms in a crystal lattice. As described in Section 2, atoms at a lattice site are free to undergo small oscillations driven by interatomic forces. For a system of p atoms interacting through a potential V, the equation of motion for the *i*-th atom is

$$m_i \ddot{\mathbf{r}}_i = \mathbf{f}_i = -\nabla_{\mathbf{r}_i} V \tag{13}$$

where **r** is the position, *m* is the mass, and **f** is the force on the atom. The interatomic potential has the general form

$$V = \sum_{i} v_1(\mathbf{r}_i) + \sum_{i} \sum_{j>i} v_2(\mathbf{r}_i, \mathbf{r}_j)$$

+
$$\sum_{i} \sum_{j>i} \sum_{k>j>i} v_3(\mathbf{r}_i, \mathbf{r}_j, \mathbf{r}_k) + \dots (14)$$

where v_1 is the self potential, v_2 is the contribution from pairs of atoms, and v_3 is the contribution from triplets of atoms. A potential that includes terms only up to v_2 is called a pair potential. A three-body potential includes terms up to v_3 . Finite difference algorithms [18] such as the Verlet or the Gear predictor-corrector are used to solve the system at each time step, which is usually less than a femtosecond. Computational constraints dictate that a trade-off be made between the size of the system and the time for the simulation. Since the maximum phonon wavelength is given by the size of the system, long wavelength phonons are cut off in an MD simulation. Use of periodic boundary conditions and a careful choice of the system size offsets this limitation. Parallelization permits simulations of systems comprising tens of thousands of atoms. To provide an idea of computational requirements, a 50 ps simulation of a 10,000 atom silicon system with a three-body potential requires approximately two days of computer time on 48 nodes of a Pentium-class cluster.

At the core of the MD calculation is the interatomic potential used. The nature of the chemical bonds decides the type of forces that must be taken into account by the potential. The simplest potential is the well-known Lennard-Jones potential that models weak dipole-dipole van der Waals bonding. Silicon, however, has highly directional covalent bonds and cannot be described by a pair potential. Three-body potentials [45,46] account for bond directionality by including the effect of bond bending in addition to bond stretching, and can describe a silicon crystal with reasonable accuracy. For an excellent comparison of silicon interatomic potentials, we refer the reader to Ref. [47].

The literature on MD simulations of heat transport may be divided into two broad categories: the calculation of thermal conductivity and the calculation of phonon relaxation times. The output from an MD simulation provides the position and momenta of the atoms at each time step. A thermodynamic property may be obtained from this information using statistical mechanics [18]. Calculation of a transport coefficient such as the thermal conductivity requires use of the Green-Kubo relation derived from the linear response theory [48]. Using linear response theory, the thermal conductivity can be expressed as an autocorrelation function of the heat current [48,18] as follows

$$K = \frac{V_0}{k_B T^2} \int_0^\infty dt \, \langle j_\alpha(t) j_\alpha(0) \rangle \qquad (15)$$

where $j_{\alpha}(t)$ is the heat current at time t (the subscript refers to a Cartesian component) and V_0 is the volume. Thus, thermal conductivity can be calculated from the fluctuations in the thermal current in a system under thermal equilibrium. The nonequilibrium method involves imposing a temperature gradient on the system and computing the resulting heat current. In an alternate approach, the temperature gradient may also be computed in response to an imposed heat current [49]. For a comprehensive comparison of equilibrium and nonequilibrium methods to compute the thermal conductivity, we refer the reader to Ref. [50].

The literature on thermal conductivity calculation using MD covers a variety of materials. We point out studies on materials and structures of current interest: thin films [51], nanowires [52], superlattices [53,54], strained materials [55,56], diamond [57], and carbon nanotubes [58-61]. The use of MD to study phonon lifetimes and phonon-phonon scattering is much more limited. The MD method has been used to study the lifetimes of vibrons (analogous to phonons) in glasses by Fabian and Allen [62], and later by Bickham and Feldman [63]. Oligschleger and Schön [64] have investigated the decay of single phonon modes in selenium and quartz. Ladd et al. [65] have compared MD calculations with classical perturbation theory calculations to show good agreement in thermal conductivity and phonon lifetimes. McGaughey and Kaviany [66] have used a Lennard-Jone argon crystal to investigate the validity of single-mode relaxation time (SMRT) approximation, commonly used to solve the phonon BTE. They compute the thermal conductivity of the crystal using the BTE, with the relaxation times derived from MD simulations instead of empirical fits. In comparing these calculations with those based on the Green-Kubo formulation, they found that the agreement depended on the degree to which dispersion is taken into account. The general applicability of the semiempirical relaxation time models [22,23], which employ several fitting parameters, is thus doubtful. Molecular dynamics simulations can greatly help in this case by providing a methodology to extract relaxation rates for any novel material in a systematic manner. More general topics on MD simulations are covered in a recent review by Poulikakos et al. [67].

4. ELECTROTHERMAL TRANSPORT

Thermal transport in a semiconductor device is strongly coupled with charge transport through electron-phonon interactions. The electrons in the channel gain energy from the applied electric field and subsequently lose it to the lattice while restoring thermodynamic equilibrium through scattering. Electrons may also absorb phonons, thereby gaining energy from the lattice. This occurs near the device source where phonon absorption enables the less energetic electrons to overcome the potential barrier. At the drain terminal, electrons shed energy gained from the field by emitting phonons. In addition, charge carriers also interact with photons during generation and recombination. The overall thermodynamic system comprising electrons, holes, phonons, and photons is shown schematically in Fig. 9, along with the paths for energy exchange.



FIGURE 9. The schematic, adapted from Ref. [70], shows the different thermodynamic systems in a semiconductor device and their energy interaction pathways

Electrothermal modeling in devices may be split into a two-step problem. The first step involves modeling the transport of charge carriers to extract the heat source distribution in the device. The second step is to model the transport of phonons and calculate the resulting temperature distribution in the lattice. The interaction of charge carriers and phonons is a many-body problem in quantum mechanics, and a holistic treatment ensuring rigorous selfconsistency is not practical. We classify electrothermal modeling in literature as belonging to one of two categories. The first develops fully coupled models, but makes simplifying assumptions regarding transport by electrons and phonons and their interactions. The second philosophy is to consider heat generation and transport as two separate problems and treat each as rigorously as possible.

In this section, we begin with a brief overview of electron transport modeling in the context of a transistor. Although we do not discuss hole transport explicitly, the approach for holes is similar to that for electrons. We next discuss the physics of heat generation through electron-phonon scattering. Finally, we review the literature on the two classes of electrothermal models as discussed above.

4.1 Modeling of Charge Transport

Semiclassical electron transport in a device is governed primarily by three sets of equations [68]. The first is a set of equations for conservation of particles. The second set relates the particle currents to the field and the particle gradients. Finally, the Poisson equation provides the electrostatic potential inside the device due to charged immobile dopant atoms and mobile carriers. This basic description comprises what is known as the driftdiffusion model [68,69]. The drift-diffusion model does not include an energy conservation equation, but instead assumes isothermal conditions. This leads to inaccuracies in the diffusion term. The next model in the hierarchy is the hydrodynamic model, which includes an energy equation and models temperaturedependent diffusion in the equation for the particle current. In both the drift-diffusion and the hydrodynamic models, the constitutive equations for charge continuity and particle momentum are derived by taking moments of the Boltzmann equation for electrons. In the hydrodynamic model, heat generation is treated rigorously using irreversible thermodynamics [70,71] and Onsager's relations. However, the energy transport equation does not include a subcontinuum description of heat flow.

Higher in the hierarchy of electron transport models are Monte Carlo (MC) device simulations [72–74], which track the trajectory of electrons inside the device. The MC method includes a quantum mechanical description of electron-phonon scattering, but the electrons behave as classical particles, obeying Newton's laws of motion, in between the collisions. This semiclassical picture in effect treats the periodic field of the crystal quantum mechanically and the externally imposed field classically. Hence, the electronic states are quantum mechanical, but the motion of electrons in response to the field is classical. For a rigorous justification of the semiclassical model in the context of electron transport, we refer the reader to Ref. [75]. In devices with ultrathin bodies such as the one shown in Fig. 2, long-range Coulomb interactions and quantum effects become important. The MC scheme must be modified in this case to include a charge distribution based on the electronic wave function, which is given by the Schrodinger equation. [76,77]. For a more comprehensive review of charge transport in devices, we refer the reader to Refs. [78] and [79].

4.2 Electron-Phonon Scattering and Heat Generation

Whatever model one may choose to describe electron transport, the influence of the lattice cannot be ignored. Electron-phonon scattering limits the electron's velocity, leading to velocity saturation [68]. A rigorous treatment of electron-phonon scattering is too complicated to implement in practice. A common approach is to make the rigid ion approximation, in which the potential field surrounding each ion is assumed to be rigidly attached to it. The change in potential due to the motion of ions is then described in terms of a deformation potential that is related to the volume change of each cell. For a discussion on the limitations of this model as well as a comparison with other models, we refer the reader to Ref. [16]. We now discuss electron-phonon interactions in silicon from the viewpoint of momentum conservation. Electron transport in solids is analyzed in the wave vector or the reciprocal space of the crystal. In n-type silicon, the electrons involved in transport are located at the bottom of the conduction band [11]. The constant-energy surface at the bottom of the conduction bands in silicon consists of six equivalent ellipsoids along the $\langle 100 \rangle$ directions, as shown in the Brillouin zone representation of Fig. 10. The conduction band minima occur at about 0.85 of the distance from the zone center to the zone edge along these directions [80]. Electron-phonon scattering can be categorized as intervalley and intravalley, depending on whether the scattering moves the electrons within a valley or from one valley to another in wave vector space. Intravalley scattering in silicon is entirely due to acoustic phonons, since optical phonons are forbidden from energy and symmetry considerations [81]. Intervalley scattering, however, involves mainly optical and zone-edge acoustic phonons. Energetic electrons tend to scatter more with optical phonons, since such processes result in a higher energy loss [82]. An electron must have energy in excess of about 51 meV, the minimum energy of optical phonons in silicon, in order to emit an optical phonon.

Phonon modes that conserve crystal momentum in intervalley processes have been calculated by Long [83]. Intervalley scattering requires a large change in the wave vector as shown by the arrows in Fig. 10. The two types of electron transitions, labeled *g*- and *f*scattering, involve scattering between valleys along the same axis and along different axes, respectively. In the reduced zone scheme, both are umklapp processes as shown in Fig 11. The *g* process requires a phonon with wave vector $0.3 \times \left(\frac{2\pi}{a}\right) \langle 001 \rangle$. The *f* process requires a



FIGURE 10. The *f* and *g* intervalley electronic transitions [83] are shown in the Brillouin zone of the crystal



FIGURE 11. The *f* and *g* processes are shown in a (110)-plane to identify the wave vectors of the involved phonons. The *f*-process phonon has the wave vector \mathbf{q}_f and the *g*-process phonon has the wave vector \mathbf{q}_g purely from geometry considerations

phonon with wave vector that is 11° off $\langle 100 \rangle$. In the sketch of Fig. 11, the wave vector is approximately $\left(\frac{2\pi}{a}\right) \left[\overline{0.15}, \overline{0.15}, \overline{1}\right]$. Further, symmetry considerations [84] based on a zeroth order expansion dictate that only LO phonons are allowed in a *g* process and LA and TO are allowed in an *f* process. However, it has been shown by a first-order expansion that low-frequency TA and LA phonons are also involved in the *g* process [81]. In fact, scattering with these modes must be considered in order to match the mobility data in silicon [80]. The *f*-and *g*-transition picture holds up to moderately high electric fields (< $\approx 10^5$ V/cm).

Details of electron-phonon scattering, as described above, are not captured by the driftdiffusion and hydrodynamic models. The sophistication of a Monte Carlo simulation is needed to model such microscopic physics. Pop *et al.* [82] have computed the spectral distribution of phonons emitted by hot electrons at different electric fields. Figure 12 shows the distribution for a field of 4 MV/m. This detailed information is particularly useful in describing phonon transport, since it can be used to specify the source term in a phonon BTE or



FIGURE 12. Phonon emission from electron scattering at 4 MV/m [82] is shown. The source term in the BTE may be extracted from this emission spectrum

MC framework. We discuss this approach in greater detail in the subsection on detailed thermal models.

4.3 Coupled Subcontinuum Electrothermal Modeling

The first of the two categories of electrothermal simulations discussed above uses the hydrodynamic model for electron transport and an energy moment of the BTE for phonons. Fushinobo et al. [36,85] calculated the electron and the phonon temperatures in submicrometer gallium arsenide devices. They used a system of coupled energy equations for the lattice. Treatment of phonon dispersion was limited to distinguishing between longitudinal optical and acoustic phonons, with the optical phonons assumed to be purely capacitive and all heat being conducted by the acoustic phonons alone. The nonequilibrium phonon distribution was assumed to be isotropic and the model equations resembled Eq. (9). Each of the three systems, electrons, LO phonons, and acoustic phonons was assumed to remain in local thermodynamic equilibrium with well-defined individual temperatures. Further, electrons were assumed to couple only with longitudinal optical (LO) phonons over a single relaxation time of 100 fs and the energy in LO phonons was subsequently passed onto a gray distribution of acoustic phonons over a relaxation period of 10 ps.

Device simulations showed that while the electron temperature in the device reached more than 1000 K, the lattice temperature rise was only about 10 K. Further, the difference between the LO and the acoustic temperature was less than 10% of the LO temperature. An interesting observation was that if the gate voltage was modulated on a time scale of 10 ps, electron-LO relaxation would be unaffected but LO-acoustic relaxation would be retarded. This would lead to an energy bottleneck in the LO phonons. Lai and Majumdar [37] adapted the above model for a bulk submicrometer silicon device. While their study did not show any significant rise in the lattice temperature, the authors found that altering the substrate temperature by 100 K decreased the drain current by 17% and the electron temperature by 8%.

4.4 Detailed Thermal Modeling

In the coupled treatment described above, the complexity in the phonon dispersion relationship and the anisotropy in the nonequilibrium distribution are ignored. However, both of these are crucial in accurately modeling phonon transport. In this subsection, we review electrothermal models where phonon transport is solved in isolation from electron transport. In order to extract the heat source term, electron transport is modeled during the dynamic period lasting tens of picoseconds, when the device is being switched from the onstate to the off-state and vice versa. Thermal transport is modeled during the off-state when electrical conditions are static. The focus is on how the phonon distribution emitted by hot electrons in the on-state evolves during the offstate. The thermal problem remains coupled to electron transport in the sense that the obtained temperature or phonon distribution is fed back to the electron transport problem to estimate its impact on device characteristics.

4.4.1 Phonon Hot Spot in a Device

A detailed consideration of phonon transport in a device using the BTE-based models described in Section 3 shows that the heat generation region in the device behaves as a phonon hot spot. A phonon hot spot is defined as a localized region with a nonequilibrium population of high-frequency phonons [86]. The heat source is spatially localized to within 20 nm in a bulk device and to nearly 4 nm in a nanotransistor, as shown in Fig. 2. Additionally, the relaxation rate is approximately 0.1 ps for electrons and on the order of 10 ps for optical phonons [81,87]. This disparity, coupled with a strong excitation of optical phonon modes, results in the heat source having a power density on the order of 5 W/ μ m³ in a bulk device. The small extent and the high rate of phonon generation in the higher frequencies combine to give rise to hot spot effects. These are described below.

Chen [28] used the equation for phonon radiative transfer under a gray-body approximation to show that the thermal conductivity in the region surrounding a nanoscale heat source decreases as the ratio of the source size to the phonon mean free path. The magnitude of this size effect depends on the degree to which thermal transport in the device is truly ballistic. The spread in phonon free paths in room temperature silicon is shown in Fig. 13 as a function of phonon frequency and polarization. The meanfree paths for phonons emitted at the hot spot are also provided. Since the mean-free paths of the emitted optical and acoustic phonons are



FIGURE 13. The distribution of free paths in room-temperature silicon as a function of phonon frequency and polarization is shown. The mean-free path of phonons emitted by hot electrons in a device are also given for comparison. The use of a gray-body approximation for the heat source leads to large errors in predicting the magnitude of the hot spot effect

comparable to the source size, transport is ballistic in the vicinity of the hot spot. In addition to the size, the power density at the hot spot is also important. If the rate of generation of phonons that have high frequencies but relatively low group velocities exceeds the rate of their decay into phonons with lower frequencies but higher group velocities, there is a bottleneck to energy propagation from the hot spot. This leads to severe nonequilibrium in the high frequency modes. Sinha et al. [38] have shown that this effect is more important in nanoscale devices than the above-mentioned size effect. Since the equivalent temperatures are locally high in the vicinity of the hot spot, the size effect is reduced due to the shorter mean free path of phonons outside the hot spot. Finally, we note that though we have discussed hot spots only in the context of silicon, they are likely to be created in any ballistic device. Electrons traversing the channel of such a device will scatter inside the contact and create a hot spot. However, if the contacts are metallic, such as in carbon nanotube FETs, thermal transport from the hot spot will be dominated by electrons instead of phonons. The impact of the hot spot will partly depend on the current density in the device, which can be quite high in nanotube FETs [88]. This needs further investigation.

4.4.2 Phonon BTE Modeling

Of the three different semiclassical models described in Section 3, phonon BTE-based models have been exclusively used in modeling transport in devices. The reason stems from the relatively low computational cost of the BTE compared to MC or MD. Sverdrup et al. [30] solved a steady-state two-step energy moment of the BTE for a 400 nm gate length siliconon-insulator (SOI) device. Phonons were separated into two fluids, one serving as a purely capacitive reservoir in which all electronic energy was dissipated, and the other solely responsible for heat conduction. The two-fluid model indicated that the temperature rise in the device was 160% higher as compared to that predicted by solving the heat diffusion equation. Narumanchi et al. [34] considered the unsteady problem under a gray-body approximation to show that the difference between the solutions to the BTE and the heat diffusion equation depended strongly on the boundary conditions. Diffuse scattering at the boundaries caused a larger deviation from the Fourier law. In a separate article, Narumanchi et al. [89] extended the energy moment formulation to include phonon dispersion and polarization in the acoustic branches. As in the above models, the optical branch was assumed to have a zero group velocity. Although the model accounted for selection rules that govern phonon-phonon scattering, it did not ensure microscopic energy and momentum conservation. A macroscopic energy conservation equation was developed in terms of equivalent phonon temperatures. The model exhibited expected behavior in the ballistic and the continuum limits and was able to predict bulk silicon thermal conductivity above 300 K with reasonable accuracy.

All of the models described above are based on the energy moment of the BTE and do not include a phonon polarization and frequencydependent heat source term. Electron-opticalphonon coupling is assumed to dominate and the heat source term is included only in the equation for optical phonons. Since the models further assume optical phonons to be purely capacitive, they tend to exaggerate the temperature rise in the device. Recently, Sinha et al. [38] have proposed a split-flux BTE model that solves directly for phonon occupation instead of phonon energy density. The general approach is similar to a nonlocal formulation with the phonon departure from equilibrium having a second-order term. An isotropic dispersion relationship that includes all branches and polarizations is taken into account. The source term is obtained through detailed Monte Carlo simulations [82]. The temperature rise computed from this model is much lower than the temperature rise from previous BTE models. This is attributed to the dominance of phonons with nonzero group velocities in the emission spectrum. This reduces the thermal resistance arising from ballistic transport. The flux due to various phonon branches is shown in Fig. 14. We note that like previous BTE models, the split-flux model does not enforce microscopic energy conservation.

Although the Monte Carlo method is yet to be used for phonons in silicon devices, Lugli and coworkers [90–92] have performed selfconsistent electron-phonon Monte Carlo simulations in bulk GaAs, bulk InP, and GaAs- $Al_xGa_{1-x}As$ heterostructures. However, since



FIGURE 14. A branchwise breakup of the farfrom-equilibrium heat flux near a hot spot in a bulk device [38] shows the dominance of LO phonons

the simulations were for polar semiconductors, only LO phonons were simulated. Their results indicated that that the phonon population is driven into severe nonequilibrium during photoexcitation and impedes the relaxation of electrons and holes.

4.4.3 Evaluation of Phonon Relaxation Times Using MD

A critical parameter in all BTE models is the relaxation time for phonon-phonon scattering. Since the contribution of optical phonons to thermal conductivity is much less than that of acoustic phonons, the evaluation of relaxation rates for optical phonons has not received much attention in the literature. However, these rates are critical in modeling transport near device hot spots since the hot spot has a higher population of optical phonons. Sinha et al. [3] have used parallel MD to investigate anharmonic scattering of optical phonons at devicelike hot spots. The hot spot is modeled as a wave packet, shown in Fig. 7. Figures 15 and 16 show the evolution of hot spots phonons observed using MD. The energy in different phonon modes may be computed by Fourier analysis of the atomic displacements at any given time. This method allows identification of decay channels for hot spot phonons and the computation of relaxation rates. In another recent MD study, McGaughey and Kaviany [66] used a solid argon system to investigate the validity of the single mode relaxation time approximation used in BTE models. They showed



FIGURE 15. Snapshots of anharmonic phonon scattering at a 2 eV hot spot observed in parallelized MD simulations are shown. The axial displacements are plotted on the left and the corresponding energy distribution in real space is plotted on the right. The *x*-axis shows the axial location. The snapshot at the top is taken just before scattering is initiated. As scattering proceeds, phonons of various modes are created, corresponding to the lumps on the left and the peaks on the right, and propagate from the hotspot at different group velocities



FIGURE 16. The distribution of energy in the phonon spectrum is shown at various times during the decay of phonons at a hot spot with 50 eV energy

that relaxation rates may be directly obtained from MD calculations without requiring fitting parameters used in thermal conductivity modeling.

4.5 Outstanding Issues in Electrothermal Modeling

Compared to the maturity of electron transport modeling in devices, phonon transport modeling is still in the process of development. Key contributions in this area are summarized in Table 1. Despite the progress made in the past decade or so, there remain numerous outstanding issues in electrothermal modeling, especially in the context of nanotransistors. The first concerns the use of realistic phonon dispersion relationships in BTE-based models and phonon MC. The phonon dispersion relationship in strained materials and ultrathin films [14] has been shown to differ from that of the bulk. The impact of phonon confinement in thin films and nanowires [93] on electrothermal transport needs further investigation. Another important issue is the impact of surface phonon modes or surfons [94] in highly confined device geomeTABLE 1. Key contributions on subcontinuum electrothermal simulations are summarized. The nomenclature used is as follows. BC: Boundary condition; *T*: Temperature; QW: Quantum well; Combination 1: Adiabatic top surface and isothermal substrate; Combination 2: Layout specific heat loss to contacts and substrate; *h*: heat transfer coefficient

Approach	Dispersion Approx.	Device/Material	Key Result
Sample Ref.	Thermal BCs	Electron Transport	
Energy moment	: Two-fluid	0.2 μm GaAs MESFET	$T_{ m optical} pprox T_{ m acoustic}$
BTE [85]	$h=10^{3} \text{ W/m}^{2} \text{K}$	Coupled, hydrodynamic	$T_{ m electron}$ weakly influenced by h
Energy moment	Two-fluid	0.3 µm bulk Si	Increase in substrate temperature
BTE [37]	Combination 1	Coupled, hydrodynamic	reduces drain current
Energy moment	Two-fluid	0.4 μm SOI	$T_{\text{lattice, peak}}$ is 160% higher
BTE [30]	Combination 1	Decoupled, hydrodynamic	than $T_{\text{diffusion, peak}}$
Non-local	Isotropic	90 nm bulk Si	LO phonons with non-zero group velocity
BTE [38]	Combination 2	Decoupled, MC	dominate transport near hot spot
Monte Carlo	Gray LO	GaAs, InP, QWs	Non-equilibrium phonon distribution
[91, 92]	_	Coupled, MC	impedes cooling of photoexcited electrons

tries such as the FINFET. Interfacial phonon transport [95,96] remains another area of concern, especially in three-dimensional structures with numerous interfaces between novel materials.

Quantum interference effects become important at the nanometer length scale. As shown in Fig. 5, transport in such mesoscopic systems is modeled in terms of either the Landauer formalism [97] or using the nonequilibrium Green's function (NEGF) [98] approach. The former approach applies to ballistic transport between two reservoirs, which results in a quantized thermal conductance. The more general NEGF formalism includes scattering and may be thought of as the quantum analogue of the BTE. There is very little literature on heat transport in mesoscopic systems, and this is an important area for future research. As a start, the Landauer and the NEGF formalisms have been used to investigate heat transport in mesoscopic systems in Refs. [99] and [100], respectively.

5. LINKING MULTIPLE SCALES IN THERMAL TRANSPORT

While thermal transport in devices and nanostructures involves multiple length scales, the coupling of length scales has not received much attention in electrothermal modeling. The focus has been mostly on the development of methods for subcontinuum length scales. The boundary conditions in electrothermal simulations are typically obtained using simplifying assumptions about the geometry for heat conduction outside the device [101]. However, boundary conditions have been shown to strongly affect the lattice temperature inside the device [37,38]. An accurate treatment of boundary conditions requires linking subcontinuum thermal transport by phonons inside the device to the continuum heat diffusion outside the device. A multiscale approach that links a finite volume or finite element method with the phonon BTE or MC needs development in the near future. Such a framework already exists for MD in the context of mechanics problems such as crack propagation [102]. The extension

of methods such as the coupling of length scales (CLS) [103] and the coarse-grained MD [104] to thermal transport needs investigation.

6. CONCLUDING REMARKS

In this paper, we have reviewed the emerging field of subcontinuum thermal simulations of semiconductor devices. As device dimensions are increasingly reduced below 100 nm in the coming decade, thermal modeling will assume more importance in device design. Starting from a basic overview of heat conduction by phonons, we have discussed three principal methods for investigating phonon transport: the phonon Boltzmann transport equation, statistical Monte Carlo methods, and molecular Most of the past modeling efdynamics. fort at the device level has revolved around the phonon BTE. Device simulations using the phonon BTE show that the small size of the heat source and the high rate of volumetric energy dissipation in the lattice leads to the formation of a phonon hot spot. Transport in the vicinity of the hot spot is ballistic, leading to impeded heat conduction. The resulting temperature rise exceeds that calculated from the heat diffusion equation. However, the deviation between the two depends on the degree to which the phonon dispersion relationship is modeled. Energy moment formulations that assume all heat to be generated as optical phonons with zero group velocities predict the difference to be as high as 160%. However, detailed Monte Carlo simulations of electron-phonon scattering show that the hot spot is dominated by optical phonons with nonzero group velocities. If this nonequilibrium source distribution is included in a BTE formulation, LO phonons are found to dominate the heat flux close to the hot spot and the deviation is reduced to as low as 13% in a bulk device.

While past effort focused on the phonon BTE, molecular dynamics and Monte Carlo meth-

ods are fast gaining importance. These approaches avoid the many simplifying assumptions in a BTE formulation, but are computationally prohibitive for current devices. As devices continue to diminish in size in future, these methods will become increasingly feasible. While the MD method is fairly well developed, phonon MC needs more theoretical development, particularly in terms of improved treatment of phonon-phonon scattering and energy/momentum conservation schemes. An important assumption in all thermal modeling of semiconductor devices so far has been the use of a bulk dispersion relationship. Clearly, in devices whose body thickness is only several atomic layers, phonon dispersion changes dramatically [14] and surface effects become important. This poses a significant challenge, since data on phonon dispersion in silicon thin films is not available in the literature. Measurement of phonon dispersion in ultrathin films is, thus, an important step for the future. Additionally, lattice dynamical models for ultrathin films need to be developed in parallel.

ACKNOWLEDGMENTS

S.S. acknowledges financial support from the Stanford Graduate and the Intel Fellowship programs. This work was supported through Semiconducting Research Corporation Task 1043. S.S. thanks Eric Pop for providing data for Fig. 12 and for many helpful discussions.

REFERENCES

- Dennard, R. H., Gaensslen, F. H., Yu, H. W., Rideout, V. L., Bassous, E., and LeBlanc, A. R., Design of ion-implanted mosfets with very small physical dimensions, *IEEE J. Solid-State Circuits* SC-9:256–268, 1974.
- 2. Moore, G. E., Cramming more components onto integrated circuits, *Electronics* **38**:114, 1965.
- 3. Sinha, S., Schelling, P. K., Phillpot, S. R., and Goodson, K. E., Atomistic simulations of g-

type phonons in silicon devices, In *ASME Heat Transfer Fluids Engineering Summer Conference*, Charlotte, NC, ASME, New York, 2004.

- Taur, Y., Analytic solutions of charge and capacitance in symmetric and asymmetric double-gate mosfets, *IEEE Trans. Electron. Dev.* 48:2861–2869, 2001.
- Yeo, Y. C., Subramanian, V., Kedzierski, J., Xuan, P., King, T-.J., Bokor, J., and Hu, C., Nanoscale ultra-thin-body silicon-on-insulator p-mosfet with a sige/si heterostructure channel, *IEEE Electron. Dev. Lett.* 21:161–163, 2000.
- Hisamoto, D., Lee, W. C., Kedzierski, J., Takeuchi, H., Asano, K., Kuo, C., Anderson, E., King, T. J., Bokor, J., and Hu, C., Finfet-a selfaligned double-gate mosfet scalable to 20 nm, *IEEE Trans. Electron. Dev.* 12:2320–2325, 2000.
- 7. International Technology Roadmap for Semiconductors (ITRS). *http://public.itrs.net*, 2003.
- Mahajan, R., Nair, R., Wakharkar, V., Swan, J., Tang, J., and Vandentop, G., Emerging directions for packaging technologies, *Intel Technol. J.* 6:62, 2002.
- Sverdrup, P. G., Sinha, S., Asheghi, M., Srinivasan, U., and Goodson, K. E., Measurement of ballistic phonon conduction near hotspots in silicon, *Appl. Phys. Lett.* **78**:3331–3333, 2001.
- Goldhaber-Gordon, D., Montemerlo, M. S., Love, J. C., Opiteck, G. J., and Ellenbogen, J. C., Overview of nanoelectronic devices, *Proc. IEEE* 85:521–540, 1997.
- 11. Ashcroft, N. W., and Mermin, N. D., *Solid State Physics*, Brooks/Cole, Belmont, MA, 1976.
- Venkataraman, G., Feldkamp, L. A., and Sahni, V. C., *Dynamics of Perfect Crystals*, MIT Press, Cambridge, MA, 1975.
- 13. Kroemer, H., *Quantum Mechanics for Engineering, Materials Science and Applied Physics*, Prentice Hall, Englewood Cliffs, NJ, 1994.
- 14. Balandin, A., and Wang, K. L., Significant decrease of the lattice thermal conductivity due to phonon confinement in a free-standing semiconductor quantum well, *Phys. Rev. B* **58**:1544–1549, 1998.
- 15. Venkatasubramanium, R., Lattice thermal conductivity reduction and phonon localizationlike behavior in superlattice structures, *Phys.*

Rev. B **61**:3091–3097, 2000.

- 16. Ziman, J. M., *Electrons and Phonons*, Oxford University, Press, Oxford, UK, 1960.
- 17. Liu, J. S., Monte Carlo Strategies in Scientific Computing, Springer-Verlag, New York, 2001.
- 18. Allen, M. P., and Tildesley, D. J., *Computer Simulation of Liquids*, Clarendon Press, Oxford, 1987.
- 19. Peierls, R. E., *Quantum Theory of Solids*, Clarendon Press, Oxford, 1955.
- 20. Klemens, P. G., The thermal conductivity of dielectric solids at low temperatures - theoretical, *Proc. Roy. Soc. A* **208**:108–133, 1951.
- Carruthers, P., Theory of thermal conductivity of solids at low temperatures, *Rev. Mod. Phys.* 33:92–138, 1961.
- 22. Callaway, J., Model for lattice thermal conductivity at low temperatures, *Phys. Rev.* **113**:1046– 1051, 1959.
- Holland, M. G., Analysis of lattice thermal conductivity, *Phys. Rev.* 132:2461–2471, 1963.
- 24. Guyer, R. A., and Krumhansl, J. A., Solution of the linearized phonon boltzmann equation, *Phys. Rev.* **148**:766–778, 1966.
- Piekos, E., Atomistic simulations of phonon hotspots in silicon devices, In *Materials Research Society Spring Meeting - Symposium S*, San Francisco, April, Materials Res. Soc., Pittsburgh, 21– 25, 2003.
- Chung, J. D., McGaughey, A. J. H., and Kaviany, M., Role of phonon dispersion in lattice thermal conductivity modeling, *ASME J. Heat Transfer* 126:367–380, 2004.
- Majumdar, A., Microscale heat conduction in dielectric thin films, *ASME J. Heat Transfer* 115:7–16, 1993.
- 28. Chen, G., Nonlocal and nonequilibrium heat conduction in the vicinity of nanoparticles, *ASME J. Heat Transfer* **118**:539–545, 1996.
- 29. Chen, G., Ballistic-diffusive heat-conduction equations, *Phys. Rev. Lett.* **86**:2297–2300, 2001.
- Sverdrup, P. G., Ju, S. Y., and Goodson, K. E., Sub-continuum simulations of heat conduction in silicon-on-insulator transistors, *ASME J. Heat Transfer* 123:130–137, 2001.
- 31. Joshi, A. A., and Majumdar, A., Transient ballistic and diőusive phonon heat transport in thin

films, J. Appl. Phys. 74:31-39, 1993.

- 32. Zeng, T., and Chen, G., Phonon heat conduction in thin films: Impacts of thermal boundary resistance and internal heat generation, *ASME J. Heat Transfer* **123**:340–347, 2001.
- 33. Majumdar, A., Effect of interfacial roughness on phonon radiative heat conduction, *ASME J. Heat Transfer* **113**:797–805, 1991.
- 34. Narumanchi, S. V. J., Murthy, J. Y., and Amon, C. H., Simulation of unsteady small heat source effects in sub-micron heat conduction, *J. Heat Transfer* **125**:896–903, 2003.
- 35. Prasher, R., Generalized equation of phonon radiative transport, *Appl. Phys. Lett.* **83**:48–50, 2003.
- Fushinobo, K., Majumdar, A., and Hijikata, K., Heat generation and transport in submicron semiconductor devices, *J. Heat Transfer* 117:25– 31, 1995.
- Lai, J., and Majumdar, A., Concurrent thermal and electrical modeling of submicrometer silicon devices, *J. Appl. Phys.* **79**:7353–7363, 1996.
- Sinha, S., Pop, E., and Goodson, K. E., A splitflux model for phonon transport near hotspots, in *Int. Mechanical Engineering Congress and Exposition* 2004, Anaheim, November, ASME, New York, 13–19, 2004.
- 39. Mahan, G. D., and Claro, F., Nonlocal theory of thermal conductivity, *Phys. Rev. B* **38**:1963–1969, 1988.
- 40. Claro, F., and Mahan, G. D., Transient heat transport in solids, *J. Appl. Phys.* **66**:4213–4217, 1989.
- 41. Klistner, T., VanCleve, J. E., Fischer, H. E., and Pohl, R. O., Phonon radiative heat transfer and surface scattering, *Phys. Rev. B* **38**:7576–7594, 1988.
- 42. Peterson, R. B., Direct simulation of phononmediated heat transfer in a debye crystal, *ASME J. Heat Transfer* **116**:815–822, 1994.
- 43. Mazumder, S., and Majumdar, A., Monte carlo study of phonon transport in solid thin films including dispersion and polarization, *ASME J. Heat Transfer* **123**:749–759, 2001.
- 44. Callaway, J., *Quantum Theory of the Solid State*, Academic Press, San Diego, 1991.

- 45. Stillinger, F. H., and Weber, T. A., Computer simulation of local order in condensed phases of silicon, *Phys. Rev. B* **31**:5262–5271, 1985.
- 46. Tersoff, J., Empirical interatomic potential for carbon, with applications to amorphous carbon, *Phys. Rev. Lett.* **61**:2879–2882, 1988.
- 47. Balamane, H., Halicioglu, T., and Tiller, W. A., Comparative study of silicon empirical interatomic potentials, *Phys. Rev. B* **46**:2250–2279, 1992.
- 48. Kubo, R., Fluctuation-dissipation theorem, *Rep. Prog. Phys.*, **29**:255, 1966.
- 49. Muller Plathe, F., A simple nonequilibrium molecular dynamics method for calculating the thermal conductivity, *J. Chem. Phys.* **106**:6082–6085, 1997.
- 50. Schelling, P. K., Phillpot, S. R., and Keblinski, P., Comparison of atomic-level simulation methods for computing thermal conductivity, *Phys. Rev. B* **65**:144306, 2002.
- Lukes, J. R., Li, D. Y., Liang, X. G., and Tien, C. L., Molecular dynamics study of solid thinfilm thermal conductivity, *ASME J. Heat Transfer* 122:536–543, 2000.
- 52. Volz, S. G., and Chen, G., Molecular dynamics simulation of thermal conductivity of silicon nanowires, *Appl. Phys. Lett.* **75**:2056–2058, 1999.
- 53. Volz, S., Saulnier, J. B., Chen, G., and Beauchamp, P., Computation of thermal conductivity of si/ge superlattices by molecular dynamics techniques, *Microelectron. J.* **31**:815– 819, 2000.
- Imamura, K., Tanaka, Y., Nishiguchi, N., Tamura, S., and Maris, H. J., Lattice thermal conductivity in superlattices: Molecular dynamics calculations with a heat reservoir method, *J. Phys.: Cond. Matter* **15**:8679–8690, 2003.
- 55. Abramson, A. R., Tien, C. L., and Majumdar, A., Interface and strain effects on the thermal conductivity of heterostructures: a molecular dynamics study, *ASME J. Heat Transfer* **124**:963– 970, 2002.
- Picu, R. C., Borca-Tasciuc, T., and Pavel, M. C., Strain and size eőects on heat transport in nanostructures, *J. Appl. Phys.* 93:3535–3539, 2003.

- 57. Che, J. W., Cagin, J. W., Deng, W., and Goddard, W. A., Thermal conductivity of diamond and related materials from molecular dynamics simulations, *J. Chem. Phys.* **113**:6888–6900, 2000.
- 58. Berber, S., Kwon, Y. K., and Tomanek, D., Unusually high thermal conductivity of carbon nanotubes, *Phys. Rev. Lett.* **84**:4613–4616, 2000.
- Che, J. W., Cagin, T., Deng, W., and Goddard, W. A., Thermal conductivity of carbon nanotubes, *Nanotechnology* 11:65–69, 2000.
- Osman, M. A., and Srivastava, D., Temperature dependence of the thermal conductivity of single-wall carbon nanotubes, *Nanotechnology* 12:21–24, 2001.
- 61. Maruyama, S., A molecular dynamics simulation of heat conduction in finite length swnts, *Physica B* **323**:193–195, 2002.
- 62. Fabian, J., and Allen, P. B., Anharmonic decay of vibrational states in amorphous silicon, *Phys. Rev. Lett.* **77**:3839–3842, 1996.
- 63. Bickham, S. R., and Feldman, J. L., Calculation of vibrational lifetimes in amorphous silicon using molecular dynamics simulations, *Phys. Rev. B* **57**:12234–12238, 1998.
- 64. Oligschleger, C., and Schön, J. C., Simulation of thermal conductivity and heat transport in solids, *Phys. Rev. B* **59**:4125–4133, 1999.
- 65. Ladd, A. J. C., Moran, B., and Hoover, W. G., Lattice thermal conductivity: A comparison of molecular dynamics and anharmonic lattice dynamics, *Phys. Rev. B* **34**:5058–5064, 1986.
- McGaughey, A. J. H., and Kaviany, M., Quantitative validation of the boltzmann transport equation phonon thermal conductivity model under the single-mode relaxation time approximation, *Phys. Rev. B* 69:94303, 2004.
- Poulikakos, D., Arcidiacono, S., and Maruyama, S., Molecular dynamics simulation in nanoscale heat transfer: A review, *Microscale Thermophys. Eng.* 7:181–206, 2003.
- 68. Muller, R. S., and Kamins, T. I., *Device Electronics for Integrated Circuits*, Wiley, New York, 1977.
- 69. Bløtekjær, K., Transport equations for electrons in two-valley semiconductors, *IEEE Trans. Electron. Dev.* **ED-17**:38–47, 1970.
- 70. Wachutka, G. K., Rigorous thermodynamic

treatment of heat generation and conduction in semiconductor device modeling, *IEEE Trans. Computer-Aided Design* **9**:1141–1149, 1990.

- 71. Lindefelt, U., Heat generation in semiconductor devices, J. Appl. Phys. **75**:942–957, 1994.
- 72. Fischetti, M. V., Dimaria, D. J., Bronson, S. D., Theis, T. N., and Kirtley, J. R., Theory of highfield electron transport in silicon dioxide, *Phys. Rev. B* **31**:8124–8142, 1983.
- Fischetti, M. V., and Laux, S. E., Monte carlo study of electron transport in silicon inversion layers, *Phys. Rev. B* 48:2244–2274, 1993.
- 74. Lundstrom, M., *Fundamentals of Carrier Transport*, Cambridge Univ. Press, Cambridge, UK, 2000.
- 75. Zak, J., Dynamics of electrons in solids in external fields, *Phys. Rev.* **168**:686–695, 1968.
- Laux, S. E., Kumar, A., and Fischetti, M. V., Analysis of quantum ballistic electron transport in ultrasmall silicon devices including space-charge and geometric effects, *J. Appl. Phys.* 95:5545–5582, 2004.
- Frank, D. J., Laux, S. E., and Fischetti, M. V., Monte carlo simulations of p- and nchannel dual-gate si mosfet's at the limits of scaling, *IEEE Trans. Electron. Dev.* 40:2103, 1993.
- 78. Grasser, T., Tang, T. W., Kosina, H., and Selberherr, S., A review of hydrodynamic and energy-transport models for semiconductor device simulation, *Proc. IEEE* **91**:251–274, 2002.
- 79. Kosina, H., Nedjalkov, M., and Selberherr, S., Theory of the monte carlo method for semiconductor device simulation, *Proc. IEEE Trans. Electron. Dev.* **47**:1898–1908, 2000.
- Ferry, D. K., First-order optical and intervalley scattering in semiconductors, *Phys. Rev. B* 14:1605–1609, 1976.
- 81. Ferry, D. K., *Semiconductor Transport*, Taylor & Francis, New York, 2000.
- Pop, E., Dutton, R. W., and Goodson, K. E., Detailed heat generation simulations via the monte carlo method, In *SISPAD*, Boston, Sept., IEEE, New York, 3–5, 2003.
- Long, D., Scattering of conduction electrons by lattice vibrations in silicon, *Phys. Rev.* 120:2024– 2032, 1960.

- 84. Ridley, B. K., *Quantum Processes in Semiconductors*, Clarendon Press, Oxford, 1999.
- 85. Fushinobo, K., Majumdar, A., and Hijikata, K., Effect of gate voltage on hot-electron and hotphonon interaction and transport in a submicrometer transistor, *J. Appl. Phys.* **77**:6686-6694, 1995.
- 86. Kazakovtsev, D. V., and Levinson, I. B., Formation, dynamics, and explosion of a phonon hot spot, *Sov. Phys. JETP* **61**:1318–1326, 1985.
- 87. Menèndez, J., and Cardona, M., Temperature dependence of the first-order raman scattering by phonons in si, ge, and alpha -sn: Anharmonic effects, *Phys. Rev. B* **29**:2051–2059, 1984.
- Javey, A., Wang, Q., Ural, A., Li, Y., and Dai, H., Carbon nanotube transistor arrays for multistage complementary logic and ring oscillators, *Nano Lett.* 2:929–932, 2002.
- Narumanchi, S. V. J., Murthy, J. Y., and Amon, C. H., Computations of sub-micron heat transport in silicon accounting for phonon dispersion, in *ASME Summer Heat Transfer Conference*, Las Vegas, July, ASME, New York, 21–23, 2003.
- Lugli, P., Jacoboni, C., and Reggiani, L., Monte carlo algorithm for hot phonons in polar semiconductors, *Appl. Phys. Lett.* 50:1251–1253, 1987.
- Lugli, P., Bordone, P., Reggiani, L., Rieger, M., Kocevar, P., and Goodnick, S. M., Monte carlo studies of nonequilibrium phonon effects in polar semiconductors and quantum wells. i. laser photoexcitation, *Phys. Rev. B* 39:7852– 7865, 1989.
- Lugli, P., Bordone, P., Reggiani, L., Rieger, M., Kocevar, P., and Goodnick, S. M., Monte carlo studies of nonequilibrium phonon effects in polar semiconductors and quantum wells. ii. nonohmic transport in n-type gallium arsenide, *Phys. Rev. B* 39:7866–7875, 1989.
- 93. Mingo, N., Calculation of si nanowire thermal conductivity using complete phonon disper-

sion relations, Phys. Rev. B 68:113308, 1997.

- Knäbchen, A., Transport relaxation rate of a two-dimensional electron gas: Surface acoustic-phonon contribution, *Phys. Rev. B* 55:6701–6703, 1997.
- 95. Swartz, E. T., and Pohl, R. O., Thermal boundary resistance, *Rev. Mod. Phys.* **61**:605–668, 1989.
- Cahill, D. G., Ford, W. K., Goodson, K. E., Mahan, G. D., Majumdar, A., Maris, H. J., Merlin, R., and Phillpot, S. R., Nanoscale thermal transport, *J. Appl. Phys.* **93**:793–818, 2003.
- 97. Landauer, R., Spatial variation of currents and fields due to localized scatterers in metallic conduction, *IBM J. Res. Dev.* **32**:306–316, 1988.
- Datta, S., Electron Transport in Mesoscopic Systems, Cambridge Univ. Press, Cambridge, UK, 1995.
- Angelescu, D. E., Cross, M. C., and Roukes, M. L., Heat transport in mesoscopic systems, *Superlattices Microstruct.* 23:673–689, 1998.
- 100. Lake, R., and Datta, S., Energy balance and heat exchange in mesoscopic systems, *Phys. Rev. B* **46**:4757–4763, 1992.
- Goodson, K. E., and Flik, M. I., Effect of microscale thermal conduction on the packing limit of silicon-on-insulator electronic devices, *IEEE Trans. Compon. Hybrids Manuf. Technol.* 15:715–722, 1992.
- 102. Rudd, R. E., and Broughton, J. Q., Concurrent coupling of length scales in solid state systems, *Physica Stat. Solidi B* **217**:251–291, 2000.
- 103. Broughton, J. Q., Meli, C. A., Vashishta, P., and Kalia, R. K., Direct atomistic simulation of quartz crystal oscillators: bulk properties and nanoscale devices, *Phys. Rev. B* **56**:611–618, 1997.
- 104. Rudd, R. E., Broughton, J. Q., Coarse-grained molecular dynamics and the atomic limit of finite elements, *Phys. Rev. B* 58:R5893–R5896, 1998.